NEW ENGLAND LAW REVIEW

VOLUME 36

2001-2002

NEW ENGLAND LAW REVIEW
VOLUME 36 Fall 2001 NUMBER 1

CONTENTS

Symposium: Transgressing Borders: Women's Bodies, Identities and Families

ARTICLES

Mary Joe Frug's *Postmodern Feminist Legal Manifesto* Ten Years Later: Reflections on the State of Feminism Today
Regina Austin and Elizabeth M. Schneider.....1

The Law Room: Hyperrealist Jurisprudence & Postmodern Politics
Maria Grahn-Farley.....29

Foster Care: The Border of Family Identity Maintaining, (Re)creating, Destroying
Zanita E. Fenton.....59

Transgressing the Border Between Protection and Empowerment for Domestic Violence Victims and Older Children: Empowerment as Protection in the Foster Care System
Susan Vivian Mangold.....69

Take the Child and Run: Tales From the Age of ASFA
Richard Wexler.....129

Juvenile Delinquent Rehabilitation: Placement of Juveniles Beyond Their Communities as a Detriment to Inner-City Youths
Theresa A. Hughes.....153

COMMENT

PGA v. ADA: A Skins Match, Winner Take All in the Battle for Equality for Disabled Individuals in Professional Golf
Damien Bielli.....181

NOTES

The Evolutionary War on First Amendment Rights and Abortion Clinic Demonstration
Jessica Wainwright.....231

Mandatory Rent Escrow: A Bill Relative to Rent Withholding
Serena D. Madar.....283

Concurrent Protection of Products By Patent and Trade Dress: Use of the Functionality Doctrine in *Marketing Displays, Inc. v. TrafFix Devices, Inc.*
Alison Marcotte.....327

NEW ENGLAND LAW REVIEW
VOLUME 36 Winter 2002 NUMBER 2

CONTENTS

ARTICLES

Literal Interpretation and "Absurd" Results: *Commonwealth v. Wallace* and "Trial on the Merits"
Francis D. Doucette.....373

Is There Anything "Legal" About Extralegal Action? The Debate Over Dorr's Rebellion
Paul M. Thompson.....385

COMMENTS

John v. Baker & Child Custody: A Battle Between Parents and Courts
Alison DeCoste Bancroft.....433

Ensuring Justice Prevails in the Wake of *Theresa Canavan's Case*: A Proposal for Reform
Maryellen Ryan.....479

IMS Technology, Inc. v. Haas Automation, Inc.: Does the Next Mutation of Means-Plus-Function Claim Infringement Analysis Comport with the Purpose Envisioned by the Drafters of Section 112, Paragraph 6?
Shenanne Tucker.....525

NEW ENGLAND LAW REVIEW
VOLUME 36 Spring 2002 NUMBER 3

CONTENTS

Symposium: Translating Science into Law: Lessons from Doctors, Judges, and Lawyers about the Use of Medical Evidence in the Courtroom

Introduction 573

Science in the Courts
Robert L. Park, Ph.D. 575

Using Court-Appointed Experts
The Honorable Timothy Hillman 587

The Expert Witness
Brian Holmgren 593

Evidence-Based Medicine
Dan Mayer, M.D. 601

Medical Evidence in the Context of Child Abuse Litigation
Robert Reece, M.D. 607

Judicial Perspective on the Presentation of Medical Evidence
The Honorable Peter Lauriat 615

COMMENT

Stenberg v. Carhart: Have the States Lost Their Power to Regulate Abortion?
Aimée M. Gauthier 625

NOTE

Reasoned Compulsory Licensing: Applying U.S. Antitrust's "Rule of Reason" to TRIP's Compulsory Licensing Provision
Dora Kripapuri 669

NEW ENGLAND LAW REVIEW
VOLUME 36 Summer 2002 NUMBER 4

CONTENTS

Symposium: Responding to Rogue Regimes: From Smart Bombs to Smart Sanctions

Welcome Address
Dean John F. O'Brien 707

Report of the Committee of Experts on Nation Rebuilding in Afghanistan
Michael P. Scharf and Paul R. Williams 709

After September 11th
Ruth Wedgwood 725

International Law and Rogue States: The Failure of the Charter Framework
Anthony Clark Arend 735

Lessons from the 1986 Libya Airstrike
W. Hays Parks 755

The Iraqi Quagmire: Enforcing the No-Fly Zones
Scott L. Silliman 767

Humanitarian Intervention and International Law in the Case of Kosovo
Major General William Moorman 775

Intentional Targeting of Regime Elites: The Legal and Policy Debate
Robert F. Turner 785

National Missile Defense: A Retreat from Dr. Strangelove or How I Learn to Stop Worrying and Love MAD
Dr. Willie Curtis 795

National Missile Defense and the 1972 ABM Treaty
Robert F. Turner 807

Rogue Regimes and the Individualization of International Law
Anne-Marie Slaughter 815

Coercive Appeasement: The Flawed International Response to the Serbian Rogue Regime
Paul R. Williams and Karina M. Waller 825

The Bush Administration View of International Accountability
Ambassador Pierre-Richard Prosper and Michael A. Newton 891

Playing Games with Terrorists
Abraham D. Sofaer 903

The Case for an International Trial of the al-Qaeda and Taliban Perpetrators of the 9/11 Attacks
Michael P. Scharf 911

Accountability Without Hypocrisy: Consistent Standards, Honest History
Beth Stephens 919

Terrorism Supported by Rogue States: Some Foreign Policy Questions Created by Involving U.S. Courts
Barry E. Carter 933

The United States Sanctions Response to the Attacks of September 11, 2001: A Synopsis of Remarks at the NESL Rogue Regimes Conference
Himamauli Das 943

Managing "Smart Sanctions" Against Terrorism Wisely
Peter L. Fitzgerald 957

NOTES

Light Pollution in the United States: An Overview of the Inadequacies of the Common Law and State and Local Regulation
Kristen M. Ploetz 985

Sign on the Dotted Line: Enforceability of Signed Agreements, upon Divorce of the Married Couple, Concerning the Disposition of Their Frozen Preembryos
Marysol Rosado 1041

SYMPOSIUM

Mary Joe Frug's *Postmodern Feminist Legal Manifesto* Ten Years Later: Reflections on the State of Feminism Today*

Regina Austin**

Elizabeth M. Schneider***

INTRODUCTION

Liz: On Saturday, March 31, 2001, a conference was held at New England School of Law in memory of Mary Joe Frug on the tenth anniversary of her death.[1] The theme of the conference was "Transgressing Borders: Women's Bodies, Identities and Families."[2] Regina and I spoke at the

---

* © Regina Austin and Elizabeth M. Schneider, 2001.

** William A. Schnader Professor of Law, University of Pennsylvania Law School. I wish to thank Manthia Diawara, Deborah Post and Barbara Lewis for their helpful comments. I especially want to thank Liz Schneider for undertaking what has proven to be, for me, a most generous, stimulating and productive collaboration.

*** Rose L. Hoffer Professor of Law, Brooklyn Law School. This article is a revised version of the opening plenary session at the conference, "Transgressing Borders: Women's Bodies, Identities and Families," held at New England Law School on March 31, 2001. Thanks to Regina Austin for a very special experience of stimulating dialogue, collaboration and friendship that carried on Mary Joe's legacy. Thanks also to Judi Greenberg, Nan Hunter, Sylvia Law and Martha Minow for helpful comments, to Linda Gordon and Elinor Langer for useful conversation and to Lisa Pepe for helpful research assistance. The Brooklyn Law School Faculty Research Fund generously supported my work on this article.

1. Mary Joe Frug was a professor at New England School of Law from 1981 to 1991. On April 4, 1991, she was murdered on the streets of Cambridge, Massachusetts, not far from the home that she shared with her husband Jerry and their children Stephen and Emily. Though her death was tragic, we believe that how she lived her life is more important than how she died.

2. This conference was jointly planned by Professors Judi Greenberg, New

opening plenary session. In preparation for the conference, we reread Mary Joe's work, particularly her posthumously published Harvard Law Review article, *A Postmodern Feminist Legal Manifesto*,[3] and talked with each other about her and her ideas. We had some wonderful conversations generated by our memories of Mary Joe and her words. At the conference, we focused on our favorite quotes from the *Manifesto* (which we flashed on a screen for the audience to read) and offered our reactions and thoughts on them.

Our conversations continued after the conference. We discussed a range of topics: anti-feminism, differences among feminists, style, sexuality, middle age, the material world, activism and law reform. Our musings about Mary Joe's *Manifesto* provoked reflections, improvisations against the backdrop of a common riff, on the contemporary state of feminism, which we share with you here.

Regina: The conference was a homage to Mary Joe Frug's capacities as an academic impresario. She was instrumental in precipitating the breakaway moves of the FemCrits from the white male-dominated Conference on Critical Legal Studies, and in organizing FemCrit meetings both in Boston and around the country. Among the most pleasant memories I have of the year I spent in Cambridge as a visitor at Harvard Law School are of the dinners Mary Joe fed us and the conversations that erupted around her dining room table about everything under the feminist sun.

Mary Joe is vivid in our memories because she had the capacity to focus on you, *just you*, and make *you* feel as if *you* were the most special person on earth at that particular moment. But as alive as she remains in our hearts as a leader, organizer, confidant, and girlfriend, Liz and I feared that our recollection of her as a thinker and intellectual nudge was fading. So we thought that we would return to what she said in her *Postmodern Feminist Legal Manifesto*, and thereby, in essence, give her the chance to set the agenda for our dialogue, to call us to account for ourselves and what we have been thinking and doing these past ten years, and to force us to consider what remains of the postmodern legal feminism she envisioned and the challenges that lie ahead.

Liz and I came to the task from different places and arrive at somewhat different conclusions. Liz is an expert, activist, teacher, and theorist in the

---

England Law School, Martha Minow, Harvard Law School, Clare Dalton, Northeastern University Law School and Liz Schneider, Brooklyn Law School. Special thanks to Judi Greenberg for her leadership and work in making the conference happen and for making it a wonderful event. There were many stimulating panels at the conference on topics such as welfare, employment, domestic violence, women's international human rights and reproductive rights; only a few of the presentations are included as articles in this symposium.

3. Mary Joe Frug, *A Postmodern Feminist Legal Manifesto (An Unfinished Draft)*, 105 HARV. L. REV. 1045 (1992).

area of women's rights and domestic violence whose commitment to achieving women's equality as a lived experience has been unwavering.[4] I, on the other hand, immersed in researching the problem of antiblack economic discrimination, had let my "feminist intellectual" credentials lapse. Revisiting Mary Joe's ideas caused me to embark on a scholarly pilgrimage to locate contemporary race feminism. Along the way, I stumbled on the work of young, liberal, so-called Third Wave feminists.[5] The ideas and the energy exuded by these young, female writers of all colors left me reservedly optimistic. Liz, from her different vantage point, has more questions and concerns about some of what passes for feminism today. I see more subversive possibilities, more likelihood of producing changes in the real economic and social status of women, in the cultural/ideological take of young feminists than Liz does. Both of us, however, are concerned about the tendency of postmodern feminism to pay too little attention to the material world. We also agree that differences between and among female activists and theorists committed to the cause of women's liberation are important to the vibrancy of the Feminist Movement. Mary Joe passionately advocated this view in her *Manifesto*. Only time will tell whether we and other contemporary feminists are worthy successors of Mary Joe Frug.

TOTAL THEORY VERSUS LOCAL DISRUPTIONS

- *I am in favor of localized disruptions.*

- *I am against totalizing theory.*[6]

Regina: Local disruptions take advantage of narrative, contextualization, and attention to detail. Totalizing theories tend to be arrogant and elitist. As Third World Feminists have warned, feminism must avoid "becoming another homogenizing Western master discourse that can be imported, like any commodity or development program, to 'modernize' the developing world."[7] A "feminist" theory that works to liberate one group of women

---

4. *See generally* ELIZABETH M. SCHNEIDER, BATTERED WOMEN AND FEMINIST LAWMAKING (2000).

5. *See, e.g.*, JENNIFER BAUMGARDNER & AMY RICHARDS, MANIFESTA: YOUNG WOMEN, FEMINISM AND THE FUTURE (2000); TO BE REAL: TELLING THE TRUTH AND CHANGING THE FACE OF FEMINISM (Rebecca Walker ed., 1995); THIRD WAVE AGENDA: BEING FEMINIST, DOING FEMINISM (Leslie Heywood & Jennifer Drake eds., 1997); LISTEN UP: VOICES FROM THE NEXT FEMINIST GENERATION (Barbara Findlen ed., 1995).

6. *See* Frug, *supra* note 3, at 1046.

7. Edna Acosta-Belén & Christine E. Bose, *U.S. Latina and Latin American Feminisms: Hemispheric Encounters, in* FEMINISMS AT A MILLENNIUM 117, 120 (Judith A. Howard & Carolyn Allen eds., 2000).

(Western, bourgeois professional women, for example) may result in the oppression of another (poor immigrant domestic workers of color, for example).[8] The condescension of white bourgeois feminists pushing totalizing theories that seem culturally insensitive and materially naive has deterred many women of color from identifying themselves as feminists. The reality today is that totalizing feminism has given way to a plethora of feminisms that address gender inequality within contexts of racism, ethnocentrism, nationalism, religious fundamentalism, homophobia, class exploitation, ageism, etc., *i.e.*, whatever the situation calls for. In any event, women will not be liberated by theory alone; there must also be disruptions, local or otherwise, the more, the better.

Liz: I am sympathetic to the impulse behind both statements, but I am concerned that "localized disruptions" and "totalizing theory" seem to stand in opposition to each other. Setting up a dichotomy between "localized disruption" and "totalizing theory" is problematic. I think the best feminist work starts with the local, the experiential, the "particular," then moves into the realm of the "general" and then moves back to the local, the "particular." Indeed much of my work has focused on the need to challenge the dichotomy of theory and practice, to understand that localized disruptions can be a rich source of theory, and to embrace the notion that theory need not be totalizing.[9] Mary Joe may have thought that postmodern theory was the only theory that was not "totalizing" and true to "localized disruption." If so, was she wrong?

Martha Fineman's observation more than ten years ago that feminist theory "drifts between the extremes of 'grand theory' which is totalizing in its scope and ambitions, and personal narratives, which begin and end with the presentation of one individual's unique experience,"[10] seems relevant here. She argued, and I agree, that "[b]etween these extremes, in that space between something so exclusively personal as to be beyond generalization or political content, and something so general and abstract as to be removed from the everyday realities of women's lives, lies fertile ground for feminist methodology."[11] I want to see the connections between "localized disrup-

---

8. *See* Peggie R. Smith, *Regulating Paid Household Work: Class, Gender, Race and Agendas of Reform*, 48 AM. U. LAW REV. 851, 853 (1999); Regina Austin, *Of False Teeth and Biting Critiques:* Jones v. Fisher in Context, 15 TOURO L. REV. 389, 397-99 (1999).

9. *See generally* SCHNEIDER, *supra* note 4; Elizabeth M. Schneider, *The Dialectic of Rights and Politics: Perspectives from the Women's Movement*, 61 N.Y.U. L. REV. 589 (1986); Elizabeth M. Schneider, *Particularity and Generality: Challenges of Feminist Theory and Practice in Work on Women-Abuse*, 67 N.Y.U. L. REV. 520 (1992).

10. Martha L. Fineman, *Challenging Law, Establishing Differences: The Future of Feminist Legal Scholarship*, 42 FLA. L. REV. 25, 25 (1990).

11. *Id.* at 25-26.

tion" and theory. I want theory to emerge from "localized disruption." I want disruption to emerge from theory. I also want disruption to move beyond the "local." Neither "localized disruption" nor "totalizing theory" will ever be enough to make for substantial change.

STYLE AND SUBSTANCE

- *Style is important in postmodern work. The medium is the message, in some cases - although by no means all.*

- *When style is salient, it is characterized by irony and wordplay that is often dazzlingly funny, smart, and irreverent.*

- *Things aren't just what they seem.*[12]

- *[T]he oppositional character of the [postmodern] style arguably coincides with the oppositional spirit of feminism.*[13]

Regina: In her *Manifesto*, Mary Joe states that she was uncertain that she had mastered postmodern literary style, but she produced other texts in which her mastery of the postmodern came through loud and clear, other texts that amplify the words of her *Manifesto*.

I told my husband so much about Mary Joe that he felt as if he had known her. My husband is a professor of literature and film. Here is his reading of Mary Joe as an archetypical postmodern feminist:

> To participate in style, or to accept the hetero-normative construction of woman, or to accept the stereotype of yourself for the visual pleasure of the male gaze (as with black people playing minstrels) is, according to traditional feminist discourse, dangerous. Feminism seemingly constructed a new woman who, instead of being fully emancipated, was a prisoner of feminist discourse – anti-feminine, strong, against pleasure, aesthetics, and seduction.
>
> Mary Joe was complex because she believed in agency, subjecthood and their opposites, objecthood and being beautiful for herself and others. She believed in self-fashioning herself, objectifying herself, and being the artificial material girl. Perhaps that also made her vulnerable.
>
> Ultimately she was not playing it safe. Her body was one of her texts; her style constituted for us the resources of the postmodern feminism she was trying to develop. Our feminism made us believe that we could not have our cake and eat it too. For Mary Joe, true liberation and emancipation consisted of having them both. She was a feminist with a notion of pleasure, a feminist beyond binary oppositions. In legal terms, she sought women's well being

12. Frug, *supra* note 3, at 1047.
13. *Id.*

and the good life through full citizenship with a cosmopolitan flair.[14]

Has he read her right? Has he interpreted the style that Mary Joe deployed? If so, young advocates of Third Wave feminism[15] should find much to appreciate in Mary Joe's postmodernism. The Third Wave, in some ways, represents the legacy of her kind of styling and theorizing. Third Wavers eschew victimhood and anger; they extol the virtues of hybridity, flexibility, contradictions, and coalitions. To engage them is to engage Mary Joe intellectually.

Liz: Mary Joe lived out the dimensions of style, both intellectually and personally. She exemplified the interrelationship between style and substance in her persona. As a person, she was, to use her own words, "dazzlingly funny, smart and irreverent."[16] She also dressed that way: conscious of fashion, explicitly sexy, wearing short leather skirts and tight sweaters at a time when this was not the feminist norm. She was very consciously creating herself and using her body as a postmodern text. She dared to challenge feminist norms all the time.

There is a strong link between Mary Joe's postmodernism and Third Wave feminism. Perhaps the most startling aspect of reading Third Wave feminist literature was to see how much of it resonates with Mary Joe's work, ideas and persona.[17] The emphasis on culture, on using historic terms of derision such as "girl" affirmatively, on the sense of play and sexuality seems so directly linked to Mary Joe's postmodernism. At the same time, I have concerns. Style can be, but is not necessarily, related to substance. Oppositional style does not inevitably reflect or lead to oppositional substance. I certainly have reacted this way in the past to much of the Third Wave's emphasis on culture and style. Yet when I think of style in connection with Mary Joe, I have a renewed respect for its subversive potential.

THE F-WORD

•*[T]he emergence of different feminisms . . . [has] demonstrated that women stand in a multitude of places, depending on time and geographical location, on race, age, sexual preference, health, class status, religion, and other factors.*[18]

Regina: Unfortunately, these differences have led some of my sisters to

---

14. These words were written by Manthia Diawara, Professor, New York University, to Regina Austin. They were included among comments, long ago discarded, on an early draft of this work.

15. *See supra* note 5 and accompanying text.

16. Frug, *supra* note 3, at 1047.

17. *See supra* note 5.

18. Frug, *supra* note 3, at 1049.

stand totally outside of feminism. A case in point.

The students who run the multicultural umbrella organization at Penn Law School where I teach took it into their heads to commemorate Malcolm X's death and Women's History Month at the same time. (Malcolm X's choice of the remarkable Dr. Betty Shabazz as his spouse perhaps makes this less of a stretch than one might initially imagine.) They invited as a speaker a woman who was a poet, an actress, a professor, an entrepreneur, and a recent law school graduate. She had presence, she had style, and she commanded attention with her voice and her Afrocentric mud cloth hat. She read a poem entitled, *And Who Are We Not to Remember*, that even I, who hate poetry, found moving. She went on about how we needed to earn our X's by vexing and perplexing and excelling and exhaling or whatever. Then a member of the audience asked her about the linkage between women, Malcolm X, and Islam and the speaker turned into the human equivalent of a motor boat engine, sputtering and stuttering about how she was not a feminist and how misogyny was something foreign to black people that we experienced only because we had been set down, against our will, in a foreign Western culture. Ergo, there is no real imperative for black women to be feminists.

Feminism. The F-word. A foreign (read white racist) concept. Was I hearing a feminism that denies that it is feminism? Is that postmodern or what? I doubt it. Such an evasion of the power of the critique of the material world, which is one of the things feminism is, has its source in both a romantic nostalgia for a past that never existed and an unrealistic optimism about a future that will not become a reality without meaningful struggle. The speaker's sentiments in particular are more than likely the product of a misguided notion of African exceptionalism. Whatever or whomever we black Americans might have been if our ancestors had not been brought to America, we're here now and, like most Africans and other peoples of the African Diaspora, we are trying to hold our own in the Western-dominated modern/postmodern globalizing world! This cannot be done without women's full participation. We need feminism in some form more than ever!

Liz: Postmodernism opens up the possibility of exploring, acknowledging, and recognizing difference, and has the potential to promote anti-essentialism. At the same time, though, postmodernism does not give us a way of getting beyond anti-essentialism and of confronting anti-feminism.

As Regina suggests, the F-word continues to be a tremendous obstacle for many women, particularly younger women.[19] The wording of the dedication of the recent Third Wave book, *Manifesta, Young Women, Feminism and*

19. See for example, Mary C. Dunlap, *The "F" Word: Mainstreaming and Marginalizing Feminism*, 4 BERKELEY WOMEN'S L.J. 251 (1989-90), for an early discussion of this problem.

the Future,[20] reflects this equivocal attitude towards the word "feminist." It reads: "To feminists everywhere–including those of our generation who say, 'I'm not a feminist, but...' and others who say, 'I am a feminist, but'–with the faith that young women will transform the world in ways we haven't yet imagined." [21]

In one of my favorite New York Times pieces, entitled *And Now, Babe Feminism*, Anna Quindlen criticized the "do me feminists" with their "agenda heavy on sex when and how they want it, with no guilt, no regrets."[22] She goes on to say, and I love this part,

> it's important to remember that feminism is no longer a group of organizations or leaders. It's the expectations that parents have for their daughters and their sons too. It's the way we talk about and treat one another. It's who makes the money and who makes the compromises and who makes the dinner. It's a state of mind. It's the way we live now. [23]

Third Wave feminism operates in a historical time in which women's equality is, in a sense, taken for granted. Yet I think often there is a profound lack of a sense of history, a lack of race and class consciousness, and a cultural myopia in this assumption. Ahistorical tendencies within feminism are not unique to the Third Wave.[24] But history generally, and Second Wave history in particular, can so easily be suppressed and distorted.

Two recent wonderful books on Second Wave feminism, the feminism of my generation, provide a remedial historical account. Rachel Blau DuPlessis and Ann Snitow's edited collection of essays, *The Feminist Memoir Project, Voices from Women's Liberation*[25] and Rosalyn Baxandall and Linda Gordon's edited collection of original documents, *Dear Sisters*,[26] detail the richness and diversity of Second Wave activist work. In their Introduction, Rachel Blau DuPlessis and Ann Snitow poignantly describe this process of suppression and distortion with respect to Second Wave feminism:

> As we write in 1997, celebration of a feminist past is entirely out of style. Old rhetoric is like last decade's clothes: nothing dates worse. Feminism is not only repressed as an obstreperous act by one's parents, it is also com-

---

20. BAUMGARDNER & RICHARDS, *supra* note 5.

21. *See id.*

22. Anna Quindlen, *And Now, Babe Feminism*, N.Y. TIMES, Jan. 19, 1994, at A21.

23. *Id.*

24. Elizabeth M. Schneider, *Feminist Lawmaking and Historical Consciousness: Bringing The Past Into The Future*, 2 VA. J. SOC. POL'Y & L. 1 (1994).

25. THE FEMINIST MEMOIR PROJECT: VOICES FROM WOMEN'S LIBERATION (Rachel Blau DuPlessis & Ann Snitow eds.,1998) [hereinafter DUPLESSIS & SNITOW].

26. DEAR SISTERS: DISPATCHES FROM THE WOMEN'S LIBERATION MOVEMENT (Rosalyn Baxandall & Linda Gordon eds., 2000) [hereinafter BAXANDALL & GORDON].

> pressed into a narrow set of gestures. Ambivalent or hostile acts of memory can give rise to caricature as feminism ages in a culture that detests age, receding fast into what Henry Louis Gates, Jr. calls a recent bygone era.[27]

Yet DuPlessis and Snitow do not want to extol the past, indeed they argue that extolling the past is not enough. They say, "... we want more from the past than a safe sense of its limitations. The task is to entertain doubts and develop new formulations without censoring memory or deprecating an earlier self."[28] This is such an important statement. The development of new feminist formulations *without censoring memory* is a challenge for all of us. We cannot move forward without understanding the strengths and limitations of the past. But it is a special task for a new generation. Rosalyn Baxandall and Linda Gordon express this idea in dedicating their book "[t]o the feminists of the future, may you learn from our achievements and our mistakes."[29] Critical memory of both Second Wave achievements and mistakes is necessary to the development of new feminist formulations.

THE BODY

- *Despite . . . [significant challenges to the concept of a singular feminine identity], there remains a common residue of meaning that seems affixed, as if by nature, to the female body. Law participates in creating that meaning.*[30]

- *Legal discourse . . . explains and rationalizes the [ ]meanings [of the female body] by an appeal to the "natural" differences between the sexes, differences that the rules . . . help to produce.*[31]

Legal rules and discourse permit and sometimes mandate:

- *The* terrorization *of the female body.*

- *The* maternalization *of the female body.*

- *The* sexualization *of the female body.*[32]

Regina: I view what the law does to women's bodies differently now that I am older. More literally, less figuratively. Whereas the postmodernists were very much taken with the idea of the body as metaphor, I find myself

---

27. DUPLESSIS & SNITOW, *supra* note 25, at 22.
28. *Id.*
29. BAXANDALL & GORDON, *supra* note 26.
30. Frug, *supra* note 3, at 1049.
31. *Id.*
32. *Id.* at 1049-50.

compelled to address the subject of the body as the material, concrete embodiment of the person. Legally sanctioned sexism, racism, ethnocentrism, ageism, homophobia, and religious intolerance leave their marks all over women's bodies, like a plague of poxes. The bodies of women of color especially bear evidence of a legacy of forced excessive childbirth which supported various systems of slavery whose perpetuation was dependent upon the reproduction of children. The bodies of women of color also bear evidence of their involvement in the hard labor of subsistence agriculture (which made those of their children who survived infancy–and many did not–valuable as field hands). Finally, the bodies of women of color bear evidence of the most ubiquitous form of terror women experience, the terror that facilitates the exploitation of their labor power, that is the terror of material poverty. The lupus, the diabetes, the hypertension, the heart disease, the alcoholism, the virulent cancers, and the AIDS minority women suffer from are the marks of a history of oppression and degradation inscribed on their bodies.

The law is not there for women of color. It does not adequately police the medical profession, call discriminatory employers and insurers to account, penalize polluters in the name of environmental justice, or liberate drugs developed in part with government funding from the clutches of greedy pharmaceutical companies. In this regard, unlike ten years ago, many of us no longer draw a distinction between women at home and abroad. If the law tells white, Western, and/or privileged woman that they must be weak, nurturing, and sexy, it has been telling colored, non-Western, and/or unprivileged women something different and it has taken a tremendous toll on their bodies. There is nothing natural about that.

Liz: Understanding the body as the concrete, material embodiment of the person is crucial. Third Wave feminism has tended to focus on sexuality, agency and pleasure without a fuller analysis of other aspects of the body, reproduction, violence, childbirth, menopause, and aging, and the way these other aspects impact on sexuality, agency and pleasure. I do think age comes in here, or at least a perspective on women's experiences of our bodies over time. The body is not just theoretical but very concrete. You experience the body differently when you are young and healthy and have a sense of omnipotence of your own life and health. The material dimensions of women's bodies, of health, of reproduction, of sexual assault, pregnancy, AIDS, menopause: these are concrete problems that women face throughout our lives. I probably didn't think about these issues in the same way when I was twenty-five. Sexuality and sexual pleasure are very important, and certainly they are not just the province of the young, but these broader concrete, material aspects of women's seem missing.

In Mary Joe's work the themes of the body had all these various forms. I think that she understood the body in all of these dimensions. Her exploration of issues of maternalization and sexualization were very developed,

particularly on issues concerning women and work, pornography and prostitution. But the terrorization aspect is what I want to talk about. In the *Manifesto,* she says "[o]ne meaning of 'female body,' then, is a body that is 'in terror,' a body that has learned to scurry, to cringe and to submit."[33] This is such a vivid sentence; the words "scurry," "cringe" and "submit" convey the experience of terror so graphically.

Terrorization was the subject of my last conversation with Mary Joe, the morning that she was killed. I was in Cambridge that academic year, visiting at Harvard Law School and teaching a course on Battered Women and the Law for the first time. We spent a lot of time together. We went to her gym that morning and we talked about many things, including the course. Students at New England School of Law had asked Mary Joe if she would be interested in teaching a course on domestic violence, and we made arrangements to talk about it at a date and time that tragically ended up being the date and time of her memorial service. This conversation about domestic violence, and this theme of terrorization, has also remained, sadly, linked in my mind to the terrible circumstances of her death. But happily, Judi Greenberg, Mary Joe's close friend and colleague at New England School of Law, has carried on Mary Joe's legacy and taken on the work that Mary Joe wanted to move on to, teaching a course and supervising students in work on domestic violence.[34]

THE IMPORTANCE OF LAW REFORM PROJECTS

- *[L]egal discourse should be recognized as a site of political struggle over sex differences.*[35]

[tab]- *[C]ontinuous interpretive struggles over the meaning of sex differences can have an impact on patriarchal legal power.*[36]

- *What law (at least in part) constructs, law reform projects can reconstruct or alter.*[37]

Liz: In the *Manifesto,* Mary Joe is most optimistic about the potential for

33. *Id.* at 1049-50.

34. At New England School of Law, Judi Greenberg teaches a seminar on domestic violence and supervises an externship program in which students work with legal services offices, battered women's shelters, child advocacy organizations, and occasionally District Attorney's offices and the courts. She has also taught a course, "Comparative Approaches to Protecting Victims of Intimate Violence: India, South Africa, Ireland and the U.S.," in a summer law program in Ireland.

35. *See* Frug, *supra* note 3, at 1046.

36. *Id.*

37. *Id.* at 1048.

law reform. Much of my own work over the last decade has dealt with the impact of feminist law reform on social change. The question that I continually ask and explore in different contexts is how much reconstruction or alteration of sex differences is possible through law. If anything, experience over the last ten years has made me more skeptical about the role of law in social change, for there has been a tremendous backlash, a fierce resistance to change, and a sense in many areas of feminist lawmaking of the need to "reinvent the wheel."[38] Just as we think we have moved forward in one area of the law, judicial interpretation of doctrine on that front or on another front thwarts progressive development. One step forward, two steps backward is the name of the game. But law is always a mirror of social ambivalence, so it is to be expected. Ten years later, I think the reconstruction and alteration process of lawmaking is even more complex.

I have focused on this problem in the specific context of domestic violence in my book, *Battered Women and Feminist Lawmaking*.[39] But I think we see this in other areas of law as well. I'm particularly struck by this development in the law of sexual harassment. We have seen dramatic progress with regard to the recognition of sexual harassment as a legal wrong. Then the courts have begun to whittle away the protections, particularly on procedural issues. Recent judicial decisions have made it harder to prove sexual harassment.[40] There is tremendous resistance from employers and deep social ambivalence about what it is, and whether or how often it happens. Evidence of the minimization of the problem can be found both in the culture and among younger women who are just entering the work force. Some number of younger women and men in my gender classes are shocked that sexual harassment exists, while others are shocked at the pervasiveness of the sexual harassment that they have already experienced, particularly in legal workplaces.

Regina: The social ambivalence of which Liz speaks is as much the product of the proliferation of feminisms as it is of a generalized hostility to feminism. There are many younger women who came to feminism, not through activism, but by virtue of growing up in a time graced by the freedoms feminists past have wrought. As these younger women attempt to create a feminist identity of their own, they may be reluctant to accept without question the manifestoes of those who came before them, however much they may enjoy their fruits.

---

38. *See generally* SCHNEIDER, *supra* note 4.

39. *Id.*

40. *See* Meritor v. Vinson, 477 U.S. 57 (1986) (recognizing legal harm of sexual harassment); Harris v. Forklift Systems, 510 U.S. 17 (1993) (recognizing the harm of hostile work environment claim of sexual harassment); Faragher v. City of Boca Raton, 524 U.S. 775 (1998); Burlington Industries v. Ellerin, 524 U.S. 742 (1998) (limiting employer liability in sexual harassment cases).

Consider the case of Rebecca Walker. In the introduction to *To Be Real*, Rebecca Walker (daughter of Alice, author of *The Color Purple* and *Possessing the Secret of Joy*) reveals the misgivings she had about collecting and editing the essays included in her book. By Walker's own admission, the book is not "filled with the incisive critique of . . . patriarchy, plenty of young women from every background fighting against all manner of oppression, and inspirational rhetorical prose meant to uplift, empower, and motivate." [41] Rather, as Angela Davis says, *To Be Real* is "a gathering of 'introspective' voices examining their own personal histories for signs of feminist influence and evaluating the usefulness of what they identify as 'feminism.'"[42] This equivocal focus prompted Walker to ask the following rhetorical questions about herself:

> Am I being a bad feminist by making a book that isn't about welfare reform, environmental racism, and RU486? What about the politics? What about the activism that people need to hear about? What about the expectations that this book will be an embodiment of what I seem to represent to the outside world: young feminist activism and organizing? On some level, an internal voice is still saying: how dare you make a book that is about ideas and not about problems, how dare you be a feminist in this way and not that way?[43]

Can we forgive her? Can we forgive Rebecca Walker for editing such a book? Can we forgive her for suggesting that feminism is victim-fixated or narrowly action-oriented?

Perhaps. At least she is being honest. Anyway, she is not a lawyer and there may be something that law reformers can learn from her ambivalent feminism.

Rebecca Walker's qualms about her book project, when viewed most generously, are implicitly addressed by Alain Locke in his essay, *The New Negro*, the first selection in the historic volume edited by Locke and bearing the same title.[44] *The New Negro* is considered "the definitive text" of the Harlem Renaissance.[45] In his essay, which was written in 1925, Locke suggests that prior to that time, "[t]he chief bond between [the American Negro had] been that of a common condition rather than a common consciousness; a common problem rather than a life in common."[46] But he found abundant

---

41. Rebecca Walker, *Being Real: An Introduction, in* TO BE REAL, *supra* note 5, at xxix, xxxix.

42. Angela Davis, *Afterword, in* TO BE REAL, *supra* note 5, at 279, 281.

43. Walker, *supra* note 5, at xxxix.

44. Alain Leroy Locke, *The New Negro, in* THE NEW NEGRO 3 (Alain Leroy Locke ed., with Introduction by Arnold Rampersad, 1992) (1925).

45. Arnold Rampersad, *Introduction, in* THE NEW NEGRO, *supra* note 44, at ix, ix.

46. Locke, *supra* note 44, at 5.

evidence of a "New Negro," however, not only "in the latest phases of social change and progress, but still more in the internal world of the Negro mind and spirit."[47] Capitalizing on the economic and political gains that black Americans had achieved in the post-World War I North, the artists and writers of the Harlem Renaissance embarked on a course of "self-expression," "self-determination," and "self-portraiture"[48] that produced an extraordinarily rich cultural movement that exemplified the intelligence, creativity, the very humanity of a whole people.

I do not mean to suggest that Rebecca Walker is part of a wave of feminist writers and cultural workers whose achievements will match those of the Harlem Renaissance, but there are parallels between our time and that. Now, like then, social problems seemed to have receded in importance, and mass protests have died down; as a result, the culture wars have heated up. The fiercest race and gender battles being fought today are occurring along the cultural front lines. Women of all colors and minorities of various colors and creeds are making strides in the worlds of music, art, drama, cinema, and literature as opposed to the social or political spheres. They are capitalizing on the material gains and social tolerance won by identity politics and struggling to find a way to combine group consciousness with self-acceptance or individuality, as well to pursue a political commitment that is more of an integral part of their everyday existences than the occasion for extraordinary sacrifices of exceptional (but fleeting) intensity or all-consuming dedication that is never to be requited. The fuzzy focus of these cultural warriors may seem self-indulgent and a copout to feminist activists and Black Power militants of a bygone era, but culture is where it's at today.

Nonetheless, there are limits to how far young feminists can go in questioning or doubting the efficacy of the Second Wave's action agenda. If they think that they have a great deal of rope to play with, I suspect that they are sadly mistaken. The material conditions of women and/or minorities have not improved that much. Title VII and Title IX mean something even to young women who are not lawyers. The protections of these laws have to be defended. Indeed, there are Third Wavers who are activists in the more traditional sense, in the social and political spheres. Jennifer Baumgardner and Amy Richards, authors of *Manifesta: Young Women, Feminism and the Future*,[49] have much to say about activism. They say that activism requires a "clear intention, a realistic plan, and an identifiable constituency."[50] Among the topics that they believe merit activism are the following: the exclusion of women in the military from combat; the exclusion of gender persecution

47. Alain Leroy Locke, *Forward, in* THE NEW NEGRO, *supra* note 44, at xxv.
48. *Id.*
49. BAUMGARDNER & RICHARDS, *supra* note 5.
50. *Id.* at 295.

as a basis for granting political asylum; the motion picture industry's practice of according lower ratings to films depicting female sexual pleasure as opposed to male sexual pleasure; the lack of provision of reproductive health care services for female prisoners; disparities in the reimbursement by insurance companies for female gynecological and obstetrical services as compared with urological services obtained by men; and the practice of the National Honor Society of denying membership to expectant females as opposed to expectant males.[51] Lawyers should have something to contribute to the debate of any one of these items.

Mary Joe admonished her readers to be cognizant of the way in which the law enters into debates over sex differences and urged us to pursue legal reforms that challenge the law's role in their construction. Feminist lawyers should not be strangers to the gender and race culture wars being fought by our younger, more equivocal sisters who may be blind to the law's complicity in their struggles. Their concerns may provide a road map as to where legal activism is needed. If their analysis is insufficiently critical, they should be called on it.

POLITICS AND THE ELUSIVE "WOMAN"

- *If, or when, the social construction thesis seems about to deconstruct the basic category of woman, its usefulness to feminism is problematized.*

- *How can we build a political coalition to advance the position of women in law if the subject that drives our efforts is 'indeterminate,' 'incoherent,' or 'contingent?'*[52]

Liz: To me, this last question is one of the most important underlying concerns that critics have raised about postmodernism. While postmodernism opens up the possibility of appreciating variance and difference, there is a certain necessity for "strategic essentialism" in order to facilitate political mobilization and coalition building. I suspect that herein lies one of the real dangers of Third Wave feminism. The more fractured, hybrid, perhaps postmodern dimensions of Third Wave feminism do not lend themselves to an embrace of an activist women's agenda, with a notion of a struggle for women's rights at the core. The reluctance to identify with feminism and with a notion of women's rights seems to me to limit the possibilities of activism and organizing.

---

51. *See id.* at 299-302.
52. Frug, *supra* note 3, at 1051.

FEMINIST CLEAVAGES, HEALTHY DIFFERENCES

- *Although a powerful, broad, and coherent political community is critical to feminist law reform projects, I believe it is a mistake to fear or avoid or condemn differences among feminists as we pursue these projects.*

- *Accepting . . . our differences, in my view, is a critical component of challenging the ideology of gender difference, which includes the assumption that there is a feminine essence that unalterably unites women, binding us together under the generic category 'woman.'*[53]

Liz: The dilemma that Mary Joe identifies is how to acknowledge difference without "fracturing community." She talks about this in the *Manifesto* in a number of different legal contexts: differences among women based on race, class, ethnicity, age, and sexual orientation. She particularly focuses on this in the discussion of the antipornography campaigns, where she manages to be critical of the way in which these campaigns did "fracture community," while also seeing how the campaign opened up new ideas.

As I began to explore these issues of Third Wave feminism as a reflection of Mary Joe's work, I began to think of a new dimension of this issue of "fracturing community" -- namely, intergenerational differences. I was struck by the complex navigation of intergenerational differences many of these Third Wave books describe. Regina's quote from Rebecca Walker very much sums this up. Acknowledgment of, and respect for, intergenerational differences is a critical aspect of where we are now.

Regina: I recently came across a case involving a black woman who was turned away from the Stadium Club, a restaurant at Yankee Stadium, because she was wearing a "tank top."[54] Tank tops, thongs, and any other abbreviated attire were strictly forbidden according to the dress code printed on the back of the admission pass to the Stadium Club.[55] Unfortunately for the Yankees, she was not the only woman wearing a "tank top" that day; for one, her white friend was also wearing a "tank top." Additionally, there were other white women present who were wearing skimpier tops than the

---

53. Frug, *supra* note 3, at 1070.

54. Joseph v. New York Yankees Partnership, No. 00 Civ. 2275 (SHS), 2000 WL 1559019, at 1 (S.D.N.Y. Oct. 19, 2000).

55. *Id.* at 1. Shorts and t-shirts were also prohibited except during day games, while "quality SNEAKERS and JEANS" were permitted at all times. Exhibit A to Affidavit of David M. Bernstein in support of Defendant's Motion for Judgment on the Pleadings or Summary Judgment, Joseph v. New York Yankees Partnership, No. 00 CIV. 2275 (SHS) (S.D.N.Y. Oct. 19, 2000).

black woman's.[56] After discussions with Stadium Club and Yankee personnel, the black woman went to her car, put on a T-shirt, and was admitted to the restaurant.[57] She capitulated because she wanted to make the outing the best possible one for her kids.[58] *Black* Mom, apple pie, and *white* baseball!

The black woman sued and defeated the Yankees' motion for summary judgment. The case was ultimately settled. Her lawyer was not at liberty to speak freely about the case; and there were many questions I could not expect him to answer, such as: When did tank tops make the transition from underwear to a garment suitable for leisure activities outside the home? How old was the plaintiff? What did she look like? How skimpy was that tank top? Doesn't she know that the sun is bad for your skin and that mosquito bites carry the threat of the West Nile virus? Whatever happened to black women's quest for respectability?

Black women's bodies are dangerous because they allegedly bespeak excessive sexuality, immodesty, and aggression. For whatever reason, they make white people feel uncomfortable. Perhaps the plaintiff was saying in essence: "What if I do? That's their problem!" Would that have been a black feminist response? I think so, but I am not sure. Black women's expressions of sexuality are so often misread and our ability to control the misreading is so very limited. Repression and suppression, however, are hardly formulas for challenging stereotypes. Still, sometimes it is hard for me to discern when young black women are being sexy in defiance of external control of their sexuality or when they are being sexy as a manifestation of their submission to the demands of a racist misogyny. Which was this?

A review of the official court record in *Joseph v.N.Y. Yankees* together with a few discrete inquiries put many of my factual concerns to rest. The events took place on a very hot day in July toward the tail end of the first game of a doubleheader.[59] The plaintiff got the tickets from her best friend's husband, an employee of Major League Baseball, and they included passes to the Stadium Club.[60] The Stadium Club is reputedly a stodgy, overpriced, pretentious place that is primarily for season-ticket holders. It was originally a white male leisure space and the practices to which Ms. Joseph was subjected may reflect vestiges of that. The dress code, which is as readily appli-

---

56. *Joseph*, 2000 WL 1559019, at 1, 5.

57. *Id.* at 1.

58. Deposition of V. Whitney Joseph, July 25, 2000, at 56, *Joseph*, 2000 WL 1559019, (excerpted and attached to the Reply Affidavit of Thomas F. Wong, Jr., Defendant's Attorney).

59. Deposition of V. Whitney Joseph, at 26-31, *Joseph*, 2000 WL 1559019, (excerpted and attached as Exhibit A to the Affidavit of Aaron J. Schindel, Defendant's Attorney).

60. *Id.* at 22-23.

cable to men as women, likely originated at a time in the not so distant past when men wore coats and ties to daytime ball games.[61] In any event, minority attendance at professional baseball games is generally considered poor, although statistics for the Yankees are unavailable. Blacks' presence in special seating or concession areas like the Stadium Club is probably equally rare.[62]

There is nothing per se illegal about restaurant dress codes.[63] Such codes generally reflect bourgeois values regarding the maintenance of order and decorum in public places and the impact one's physical appearance can have on the good opinion, comfort, and enjoyment of others. Good taste/Tastes good? Dress codes tend to elevate the social status of a venue through an appeal to fashion snobbery or by creating a certain visual aesthetic.[64] Prohibiting some forms of dress may even promote civility and public peace.[65] Dress codes become legally problematic, however, when their enforcement becomes the opportunity for the exercise of invidious discrimination. Given that blacks are viewed as being bad for business be-

---

61. *See* Charlie LeDuff, *Office Dress: Baseball Casual,* N.Y. TIMES, July 27, 2001, at A17 (reporting on the impact of casual dress in the office on the attire of daytime baseball game patrons).

62. Ken Shropshire, a black colleague of mine who teaches at the Wharton School and is an expert on sports law, suggested that black attendance at sporting events, especially major league baseball, is disproportionately low. As a result, gatekeepers and ticket takers may not see many blacks and may accord them the type of treatment that can flow from a lack of interaction. When he sits in the better seats at an event (i.e., on the floor or in a luxury box), Shropshire says that he on his guard in the same way that he would be if he walked into a jewelry store or was driving an expensive car and being following by the police. E-mail from Kenneth L. Shropshire to Regina Austin (May 23, 2001) (on file with the author).

63. *See* Moolenaar v. Atlas Motor Inns, Inc., 616 F.2d 87 (1980) (mere ejection of a patron wearing "open-toed 'dress' sandals sans socks" by a nightclub requiring "elegantly casual" dress did not support a claim for intentional infliction of emotional harm).

64. *See* JOANNE FINKELSTEIN, DINING OUT: A SOCIOLOGY OF MODERN MANNERS, 81, 87, 97 (1989) (comparing the dress requirements of a bistro mondain, a café mundane, and an ethnic restaurant). A male friend told me that he went to a restaurant that required that gentlemen wear a coat only while walking to and from their tables; they were free to remove their jackets after they were seated. The restaurant floor was in essence hyper-public while the customers' tables were quasi-private. Proper patron attire was an integral aspect of the ceremony of escorting customers to their tables.

65. For example, in Hessians Motorcycle Club v. J.C. Flanagans, 86 Cal. App. 4th 833, 103 Cal. Rptr.2d 552 (2001), a sports bar was found not to have violated California's public accommodations law by denying admission to anyone wearing "colors," i.e., patches indicating membership in a particular motorcycle club. The policy was intended to prevent fights between members of rival motorcycle gangs and was enforced without regard to any of the ascriptive characteristics (like sex, race or color) addressed by the law.

cause they may inject a note of tension and disharmony into a white-dominated leisure venue by their mere presence,[66] black restaurant patrons are very likely prime targets for biased invocation of dress codes. Indeed, the plaintiff claimed that she was not the only black person harassed that day at the Stadium Club. She spoke with a scantily-clad white female who had no trouble entering the restaurant, but was nonetheless very disturbed because her black male companion was admitted only after he borrowed as pair of socks to wear with sandals.[67]

Finally, the plaintiff and her friend were not technically wearing "tank tops." A tank top is similar to a man's undershirt; it is sleeveless and has "deep armholes shaped toward the shoulder to form narrow straps" and a scooped or U-shaped neckline.[68] The plaintiff, on the other hand, had on a knit top manufactured by Inc. that was cut almost like a halter at the sleeves; the neckline in front was a shallow V; and unlike a halter, the top fully covered her back except around the shoulders.[69] Her friend was wearing a DKNY sleeveless cable knit swallow V-neck sweater.[70]

Stylish young women today are wearing classic white ribbed tank tops, which they reportedly refer to as "wifebeaters." That term is said to have "caught on less because of its crude shock value than for its jokey irony as a combination of put-down and form of appropriation."[71] It supposedly calls to mind a stereotype of "lumpen brutes" like Stanley Kowalski of *A Streetcar Named Desire* and Ralph Kramden of television's *The Honeymooners*.[72] For young women to wear in public the underwear of the hypermasculine, aggressive working-class male represents a turning of the tables of sorts. Their use of the term "wifebeaters" perhaps also suggests a belief in their own invulnerability ("too hip to be hurt"). Some feminists might object to

---

66. *See* Regina Austin, "Not *Just for the Fun of It": Governmental Restraints on Black Leisure, Social Inequality, and the Privatization of Public Space*, 71 S. CAL. L. REV. 667 (1998) (describing how and why whites attempt to privatize public spaces so as to exclude blacks) [hereinafter, Austin, *Governmental Restraints on Black Leisure*]; Regina Austin, *"Bad for Business": Contextual Analysis, Race Discrimination, and Fast Food*, 34 J.MARSHALL L. REV. 207, 230-42 (2000) (discussing the construction of blacks as bad customers in the fast food context).

67. *See* Deposition of V. Whitney Joseph, *supra* note 58, at 53-55.

68. *See* CHARLOTTE MANKEY CALASIBETTA, FAIRCHILD'S DICTIONARY OF FASHION 582-83 (2d rev. ed. 1998).

69. Ex. F, Reply Aff. of Thomas F. Wong, Jr., Defendant's Attorney, *Joseph*, 2000 WL 1559019.

70. Ex. G, Reply Aff. of Thomas F. Wong, Jr., Defendant's Attorney, *Joseph*, 2000 WL 1559019,.

71. Tom Feran, *Getting the Hang of Slangy Topper*, THE PLAIN DEALER, May 6, 2001, at 1L.

72. *See* Elizabeth Hayt, *An Undershirt Named . . . What?*, N.Y. TIMES, Apr. 22, 2001,at 1.

the term for making light of the pain and vulnerability of battered women as well as underestimating the pervasiveness of domestic violence among young people.[73] But public campaigns and legal efforts, much of which are documented in Liz's excellent book on the domestic violence movement,[74] have made domestic violence very much a part of the public discourse and generated weapons with which to combat it. Fashion and style are tools or ammunition that facilitate the struggle. Purity of purpose and point of view cannot always be expected where culture is put to the service of a good cause.[75]

Feminists who prize hip hop confront somewhat the same problem in dealing with a male-dominated art form that can be ironic and playful, but that, as a matter of course, refers to women as "bitches" and "hoes" and otherwise puts them down. Liberal use of "nigger" and its variations also abound. Of course, with regard to "nigger," the hip hopsters can claim they are turning a term of derision into one of affection so as to reduce its sting and to put others on notice that blacks are on to their game. I see little of that ironic or preemptive spirit behind the derogatory references to black women in rap music.

Despite this, black feminist afficionados of hip hop defend their enthusiasm for the genre. For example, in an essay entitled, *Sexism and the Art of Feminist Hip-hop Maintenance*, Third Waver Eisa Davis asserts that the derision of women found in rap does not oppress her because she knows that it does not refer to her;[76] it is the product of an imaginary, reactionary performative bravado that produces lyrics that she sometimes finds humorous because of the possibility that they may be mistaken for reality.[77] She believes that women can resist being bound by hip-hop's sexist representations of them. She finds positive aspects in the internal debate about the sexism that is being led by strong female rap figures like Queen Latifah. As

---

73. *See* Jay G. Silverman et al., *Dating Violence Against Adolescent Girls and Associated Substance Use, Unhealthy Weight Control, Sexual Risk Behavior, Pregnancy, and Suicidality*, 286 JAMA 572, 572 (2001) (reporting that approximately twenty percent of female Massachusetts public high school students surveyed reported experiencing physical or sexual violence from dating partners); Kathryn E. Suarez, *Teenage Dating Violence: The Need for Expanded Awareness and Legislation*, 82 C. L. REV. 423 (1994) (describing the problems incurred by teenagers involved in battering relationships, though the victims are living at home with their parents).

74. *See generally* SCHNEIDER, *supra* note 4.

75. *See* Jennifer Reed, *Roseanne: A "Killer Bitch" for Generation X*, *in* THIRD WAVE AGENDA, *supra* note 5, at 122, 124 (arguing that the second wave of feminism revealed the limits of the politics of purity and the need to negotiate the desirable counterhegomonic and the inescapable hegemonic in culture and politics).

76. Eisa Davis, *Sexism and the Art of Feminist Hip-Hop Maintenance*, *in* TO BE REAL, *supra* note 5, at 127, 133-34.

77. *See id.* at 133-34, 137-38.

Davis puts it, the female artists should be given credit for "relentlessly transform[ing] the taboo into the cool because we all privilege from the truth."[78] Davis places her hopes on the capacity of women to engage in "dialogue [with hip-hop's misogynists and nihilists] that doesn't turn on the axis of anger, blame, or victimology."[79] If hip hop is defensible, then surely "wifebeaters" is too.

Finding safety in a closet of sort, many black feminists have chosen to be silent about black women's sexuality.[80] They feared that expressions of support of sexual freedom would be too easily misconstrued because of the power of stigmatizing stereotypes about black women's sexual promiscuity.[81] Black feminists tended to emphasize the restrictive, repressive, and dangerous aspects of sexuality rather than its association with black women's own pleasure and agency.[82]

In lieu of sexual freedom, traditional black bourgeois feminism promoted respectability. Respectability was both black women's primary weapon against social denigration and legal discrimination, as well as their most cherished possession or prize. It is hard to imagine sexuality replacing either. But consider the views of one of the shining examples of black Third Wave feminists:

> But while women today still experience sexism, we do so in markedly different ways. Many of us are empowered enough to combine our erotic power with resources that were unimaginable to our mothers–money, education, talent, drive, ambition, confidence, and the freedom to just "go for ours."
>
> We have the luxury of choosing both our battles and our artillery. We know that sometimes winning requires utilizing whatever confrontational measures are necessary. We're not afraid of lawsuits, boycotts, organized protests, or giving a deserving offender a good cussing out. But we also recognize that there are times when winning requires a lighter touch. And sometimes a short skirt and a bat of the eyes is not only easier but infinitely more effective.[83]

I cannot imagine the impact young black feminists' explicit quest for and

---

78. *Id.* at 140.

79. *Id.* at 138.

80. *See* Evelynn M. Hammonds, *Toward a Genealogy of Black Female Sexuality: The Problematic of Silence*, *in* FEMINIST THEORY AND THE BODY 93, 97 (Janet Price & Margrit Shildrick eds., 1999).

81. *See, e.g.*, Lori A. Tribbett, *Saying Nothing, Tallking Loud: Lil' Kim and Foxy Brown, Caricatures of African-American Womanhood*, 10 S. CAL. REV. L. & WOMEN'S STUD. 167, 198 (2000) (concluding that the force of stereotypes undermines the ability of two representatives of the newer generation of female hip-hop artists to define black womanhood through reliance on erotic personas).

82. *See* Hammonds, *supra* note 80, at 99.

83. JOAN MORGAN, WHEN CHICKEN-HEADS COME HOME TO ROOST: A HIP-HOP FEMINIST BREAKS IT DOWN 221-22 (1999).

utilization of greater sexual freedom will have on critical race feminism or its law reform agenda. Perhaps black feminists will be less fixated on notions of respectability or freer to explore new formulations of the concept. Non-bourgeois black women may find that more social and material resources will be put into the defense of their sexual freedom and reproductive health. The alarming rate of AIDS and HIV among poor heterosexual black women in the United States cries out for a frontal assault on ideological and material sources of their sexual and economic disempowerment.[84] Economic hard times may also send the Third Wavers into retreat. Of course, uneven development means that some segments of black womanhood have yet to benefit from the formal equality and legal entitlements that are associated with the Second Wave of feminism. A renewed commitment to the old agenda will work for them. The quest for black women's full political, social and material equality such as envisioned by the Third Wavers may move the locus of activism into areas that were once considered frivolous and unimportant because of their association with bodily pleasure and agency . . . like wearing a tank top at a baseball game!

While I admire the optimism young black feminists express about their ability to put their sexual freedom on the agenda or to use their sexuality as tool in the struggle for equality, in my view, they will not be able to project a more sexualized image of themselves and be taken seriously without gaining greater control over the interpretation of their images. The power to control the interpretation depends on their effectively amassing and welding economic, social, and political capital. Black women will not know that they have arrived at the point where they can be sexy for the sake of being sexy unless they have the resources to choose to be something else instead. It is dangerous to think that cultural expression can be wholly divorced from the material and political reality. By all means be sexy, but never take your eye off the material ball.

At base, I suspect that Ms. V. Whitney Joseph was more concerned about being cool and comfortable than sexy and stylish when she wore a "tank top" to the Yankee's Stadium Club, but what if she were not. Excerpts from her deposition reveal her to be a smart, persistent, down-to-earth, courageous, and savvy advocate of equal rights. She took on a bastion of white male privilege and won. How much has not been disclosed. She claimed an entitlement to enjoy a place of leisure, pleasure, and enjoyment on an equal basis with whites, an entitlement that I have elsewhere argued is too rarely claimed and fought for.[85] There was nothing objectionable about the manner

---

84. *See* Kevin Sack, *Epidemic Takes Toll on Black Women*, N.Y. TIMES, July 3, 2001, at A1.

85. *See generally* Austin, *Governmental Restraints on Black Leisure*, *supra* note 66 (describing the importance of black leisure and arguing that leisure restraints must be challenged in and out of courtrooms).

in which she objected to her treatment. She kept her emotional cool in circumstances in which many black women lose it (and not without justification). Though embarrassed and humiliated, Ms. Joseph moderated her tone during an encounter with a Yankee Stadium security guard; according to her:

> I was determined not to fulfill any low expectations that man had of me. I did not scream, I did not shout, I did not get ugly because I knew that that would be exactly what would justify the treatment that I was receiving in his mind.
>
> So I dug way down, found my self-reserve and used the best words from the education I've got to try to speak with that man.[86]

A black woman in shorts and a "tank top," she was a credit to her race and her gender. What more can we ask?

LET'S TALK ABOUT SEX

- *If women's oppression occurs through sex, then in order to end women's oppressions in its many manifestations the way people think and talk and act about sex must be changed.*[87]

Liz: I thought about Mary Joe recently, and the difficulty of changing the ways that people think and talk and act about sex, when I went to the Valentine's Day mega-benefit performance in Madison Square Garden of Eve Ensler's *The Vagina Monologues*. For several years now some friends and I have had season tickets for the WNBA's New York Liberty. As we would attend these games, with the Garden filled with women of every color, age and sexual orientation – all pumped up from watching women's basketball and rooting for an extraordinary group of women athletes – we would talk about how wonderful it would be if the Garden could be filled for a feminist political event.

Well this night it *was* filled with women for such an event-women thinking about sex and sexual power–thinking about their vaginas. It was V-Day!!! My friend Susan Herman and I went to the Garden with our early-twenties daughters looking forward to an intergenerational feminist evening! But it proved much more complicated than I expected. I had seen and enjoyed *The Vagina Monologues* before, and knew that it focused exclusively on women's vaginas and affirming the importance of women's sexuality, but seeing it again at the Garden in this mega-version, I was struck by the fact that the *Monologues* do not mention anything other than sexuality, such as jobs, education or health, or even the way in which women's sexuality impacts these areas. Sexiness was the answer to women's inequality;

---

86. Deposition of V. Whitney Joseph, *supra* note 58, at 61.

87. Frug, *supra* note 3, at 1073.

women's freedom was the freedom to flaunt our sexuality, like the riff of Calista Flockhart, TV's *Ally McBeal*, on the freedom to wear skirts as short as she wants. There *was* the ritual of linking women's sexuality to the situation of women in countries where clitorodectomies are performed and of women under the Taliban, with Oprah Winfrey dramatically lifting the burka off a woman from Afghanistan, who entered the Garden with lights and a drum roll, after a brief rap by Oprah on global feminism. But I came away with some very uneasy feelings about the degree to which the event had to do with feminism, and so did my companions. When I read a story in The Village Voice a few days later, written by a skeptical commentator relating the "cuntism" of the event to ambivalence about feminism, my discomfort was reinforced:[88]

> Last Saturday's V-Day celebration reached what you might call its climax with Glenn Close's reading of "Reclaiming Cunt." She savored the word, slipping her tongue around its letters (c as in "cute," u as in "urge") until, by its end (t as in "tangy"), she was down on her red pantsuited knees on the red stage at the center of Madison Square Garden, screaming, "Cunt! Cunt! Cunt!" at the top of her lungs. In a mass simultaneous experience, the audience of 18,000 was screaming with her. . . .
>
> But even as "hoochie," "pooter," "twat," and "coochi snorcher" rolled off celebrity tongues and Ensler waved a vibrator at her delighted fans, another word seemed to stick in the throats of even the most vagina-loving Hollywood icons: feminist. Ask about it at the Hammerstein Ballroom gala -- which, along with the performance, raised more than $2 million to fund antiviolence programs and more V-Day performances -- and it's as if you've defamed the clitoris or something.
>
> "I don't know about feminist," said Isabella Rossellini. "Is this about feminism really? Violence against women is a feminist issue? I don't think it is." OK, but does she consider herself a feminist? Rossellini looked as if she were smelling something unpleasant. "Well, I don't know what you mean. I would not label tonight a feminist night; it's a women's night. I mean, there are Republican women, there are Democratic women, there are feminist women, and women who don't define themselves, they're just women against violence."
>
> Ricki Lake . . . was similarly dodgy about the term. "I don't like labels," said Lake, adding that she does support women's rights. Marisa Tomei admitted to considering herself a feminist, though quickly pointed out that she's not sure what that is. "I always was a feminist, but without definition," said Tomei.
>
> Even the Queen of Cunt herself wouldn't own the F-word. "I have this cli-

88. Sharon Lerner, *Clit Club V-Day's Charismatic Cuntism Rocks the Garden*, VILLAGE VOICE, Feb. 20, 2001, at 60.

> ched image of what a feminist would be, and I don't want to be that way," said Glenn Close . . . . And what way are feminists perceived exactly? "They don't like men -- you know, kind of, um, butch," said the rabbit-killer of Fatal Attraction fame. So feminist stays in the verbal gutter, while Close carries cunt to redemption? "That, to me, is about humanity," the actress said . . . . "It's about something that's bigger than what I always kind of thought feminism was."
>
> Therein lies the marvel of this V-movement, which has turned the stuff of the old take-back-the-night rallies into a hot ticket. Rape, domestic violence, even homelessness, when it happens to women -- Ensler has transplanted these issues into a context that seems edgier and yet is somehow more palatable than the dread feminism. "Vaginism" doesn't get all mucked up in messy issues like abortion or unequal pay. And though it often references lesbians -- or their vaginas, anyway -- the V-movement doesn't get in the way of being attractive to men. . . .
>
> V-day, which the program refers to as "a movement . . . a vision . . . a spirit," is about seeing sexiness as the source of women's power. . . .[89]

"Cuntism"–having women thinking differently, and arguably more affirmatively, about their own sexuality–doesn't change much in terms of women's power; "cuntism" is simply not the same as feminism. I do agree with Mary Joe that, in order to change the condition of women, the way women talk and think about sex and sexuality must be different. But the V-Day story resonates for me because talking about sex or sexuality does not necessarily translate into feminism. This is very tricky stuff and goes back to the way in which sexuality can be simultaneously a tool of oppression and empowerment. Many Third Wavers seem to see sex, sexuality and sexual pleasure as primary. Women's interest in sexual freedom, agency and autonomy has the potential to, but does not necessarily, translate into a broader understanding of women's experience and oppression in the world. And sex and sexuality has to be understood in a larger material context.[90]

---

89. *Id.*

90 To be fair, V-Day event fundraising has raised nearly $7 million for educational programs on violence against women around the globe so Ensler has certainly attempted to give this work on women's sexuality a material impact. *See* Cindy Richards, *Ensler Asks That We Envision World Without Rape*, WOMEN'S E-NEWS (December 3, 2001) *at* http://www.womensenews.org/article.cfm/dyn/aid/741/context/archive.

MOVING FROM SEX! TO SEX?

- *Only when sex means more than male or female, only when the word "woman" cannot be coherently understood, will oppression by sex be fatally undermined.*[91]

Regina: Essentialism can be strategic. I don't know that we have to go the whole way in deconstructing womanhood. This may be an area where Mary Joe, the person, diverges from Mary Joe, the Manifesto writer. As hip-hop feminist Joan Morgan says in her book *When Chicken-Heads Come Home to Roost,*

> Am I no longer down for the cause if I admit that while total gender equality is an interesting intellectual concept, it doesn't do a damn thing for me erotically? That, truth be told, men with too many 'feminist' sensibilities have never made my panties wet, at least not like that reformed thug nigga who can make even the most chauvinistic of 'wassup, baby' feel like a sweet, wet tongue darting in and out of my ear." [92]

I think Mary Joe would have loved this.

My hormones, or the lack of them, tell me that women are both biologically *and* socially constructed. In any event, neither social construction, nor biology can be the basis for vibrant politics. Feminism is, at base, a range of social practices, cultural understandings, and political approaches. Unless it proposes to be a form of fascism, feminism must socialize; it must proselytize. Women are not born feminists. Though they may be born free, over time they come to know constraint and too many of them consent to its strictures. Liberation does not come without a struggle, without crossing boundaries, without opening up possibilities that once seemed naturally foreclosed. Feminism should not be another obstacle; it should grease the path.

That said, I can see how liberating it would be if critical race theory, for example, embraced the subtlety, generosity, and mischievousness that would come with fully acknowledging and challenging the role that sexuality plays in racial subordination. It makes me sad and then furious, but mostly sad, when I hear black women accept the foolish idea that sexual liberation is about sexual orientation. They cannot see that the varieties of sexual oppressions are largely indivisible, and that racial oppressions cannot be divorced from sexual oppressions.

I have changed and matured a great deal since Mary Joe died. I marvel at how much of Mary Joe's work seems predictive of the path my thinking has followed. I understand better now that the style she exhibited in word and

---

91. Frug, *supra* note 3, at 1075.
92. *See id.*

deed, and the praxis it promotes, would be beneficially applied across the range of struggles for human liberation.

Liz: It has been wonderful to reread the *Manifesto* and think about Mary Joe. Mary Joe's work addresses the important questions that we continue to grapple with–the role of theory, the possibility and limits of law and social change, the salience of the category "women" as the basis for political engagement. Yet rereading her work at this time, talking with Regina, and thinking about where are we today has forced me to confront my many complicated feelings about the contemporary state of feminism, shaped by my own locale, the law school.

On the one hand, there is the excitement and power of global connections, and the fact that there is a women's international human rights movement. On the other hand, I teach in law school classrooms that are filled with women, who sometimes compose more than half the class, yet so few of them understand that they would not be in law school were it not for the struggles of First and Second Wave feminists. On the one hand, there are the tremendous accomplishments directed at reducing and preventing domestic violence over the last thirty years. On the other hand, there are "wifebeater" t-shirts that legitimize and render ordinary woman-abuse although claiming to use the term subversively. On the one hand, there is the fact that courses and clinics on women's rights, domestic violence and advocacy work continue to attract an extraordinary group of smart, committed and diverse women who are desperate and hungry for this work (and for teachers and mentors who can guide them in this work) at every law school where I teach, whether Brooklyn, Harvard, or Columbia. On the other hand, it is only because of these courses that many students understand for the first time that issues of equal pay, discrimination and harassment in the workplace, and problems of work and family responsibilities continue to be material real-world problems that will affect their lives and must be fought. In effect, there is little knowledge of, and therefore little "memory" of First and Second Wave struggles. For many of these students, aspects of women's equality may have been absorbed "in the culture" in some sense. But many do not recognize or appreciate the importance of material struggles for women's equality both for themselves and for women around the globe.

Mary Joe's work focused on culture and its subversive possibilities, but it saw the material world, both the physicality of the body, and the structure of economic power, discrimination, wages, poverty, racism, economic inequality as the most potent source of feminism and the most important site of struggle.[93] For both Regina and I, rereading Mary Joe and reflecting on

93. *See, e.g.*, Mary Joe Frug, *Securing Job Equality for Women*, 59 B.U. L. REV. 55 (1979); *see also* Mary Joe's course materials for Women and the Law, now published as JUDITH G. GREENBERG, MARTHA L. MINOW & DOROTHY E. ROBERTS,

feminism has highlighted the importance of this material dimension and the need for this dimension to be developed in contemporary feminist struggle. As Regina eloquently puts it, "it is dangerous to think that cultural expression can be wholly divorced from material and political reality." Use of the term "wifebeater" cannot be ironic and subversive until the reality of violence against women has changed.

I too have changed since Mary Joe died. The world has changed. It has certainly dramatically changed since September 11. Engaging with the *Manifesto* has made me think more deeply about feminism today, if only because I see so much resonance of Mary Joe in what has happened. As a movement across generations, we need to "develop new formulations without censoring memory."[94] I hope that engaging with Mary Joe's work and thinking about these issues with Regina has made me more generous and open to the perspectives of many of the younger women that I teach and work with (who carry on her legacy, even if many of them don't know it). Remembering Mary Joe and her inimitable capacity to facilitate human and intellectual connection has made me want to forge deeper intergenerational links that are so important to our struggle.

MARY JOE FRUG'S WOMEN AND THE LAW (2d ed. 1998)

94. DUPLESSIS & SNITOW, *supra* note 25, at 22.

The Law Room: Hyperrealist Jurisprudence & Postmodern Politics©

Maria Grahn-Farley*

I. INTRODUCTION

I am in favor of localized disruptions; I am against totalizing theory.[1]
– Mary Joe Frug

The most "localized disruption" of all takes place within oneself. The self is a site from which to reject "totalizing theory." Totalizing theory takes away one's ability to experience disruption from within. Totalizing theory prevents one from being aware of one's own existence. The experience of internal disruption, of ambiguity, is necessary if one is to develop an awareness of one's own existence as a human being.

This Article is about the ways in which the system of marks replaces the internal ambiguities with a totalizing theory that fixes one's understanding of the world and the people in it along lines like sex and race. This Article

---

© Maria Grahn-Farley, New England Law Review 2001.

* Visiting Adjunct Professor, Golden Gate University School of Law (2001-2002); Co-Facilitator, Harvard Child Rights Working Group (1999-2001); former Member of the National Board of Rädda Barnen (Save the Children, Sweden). LL.M. Göteborg Universitet (Gothenburg University, School of Economics and Commercial Law, Sweden). This Article benefited from a presentation at "Transgressing Borders: Women's Bodies, Identities and Families," a conference in memory of postmodern feminist legal theorist Mary Joe Frug. I thank Clare Dalton, Judi Greenberg, Martha Minow, and Liz Schneider for their work in organizing the conference. I thank the New England School of Law for hosting the conference. I thank the law faculty of Göteborg Universitet in Sweden, especially Professor Eva-Maria Svensson. I thank Carl Lebeck at Stockholm Universitet and Sven Gustafsson at Göteborg Invandrarbyrå. I thank Daria Roithmayr for her insightful contribution to the conference panel, "Postmodern Feminism: Law's Bodies," of which this Article was part. I never met Mary Joe Frug but I have felt her presence in my friendship and intellectual exchanges with feminist legal theorist Phyllis Goldfarb: Thanks Phyllis. I also thank Ruth Arlene W. Howe for all her support. Finally, I thank my wonderful husband, Anthony Paul Farley, for his constant enthusiasm and encouragement of this and all of my projects.

1. MARY JOE FRUG, *A Postmodern Feminist Legal Manifesto*, *in* POSTMODERN LEGAL FEMINISM 125, 125 (Routledge 1992).

is about the ways in which people marked as *Woman* and people marked as *Colored* are produced in and through the system of law as *Woman* and *Colored*. Finally, this Article is about the way that the system of marks and the system of law together produce the hyperreal: the site of postmodern political struggle.

We must examine the conditions under which a person becomes a natural subject and, later, a legal subject, and, finally, hyperreal. This is an examination of the roles of the system of marks and the system of law in the creation of the hyperreal identities of the *Woman* and the *Colored*. The hyperreal operates through an essentialist politics of group hierarchy. The hyperreal is the supercession of the distinction between the true and the false, between the copy and the original. Postmodern identity politics is neither an examination of authenticity, nor a separation of the true from the false within the totalizing theories governing the *Woman* and the *Colored*. Rather, postmodern identity politics examines the uses of authenticity; it examines the authenticity of authenticity itself. It is a politics of the hyperreal. The space in which these examinations will be organized in this Article consists of two provisions of the Swedish criminal law: first, the provision regarding "Agitation Against an Ethnic Group;"[2] and second, the provision "Prohibiting Purchases of Sexual Services."[3]

The space of examination is the Swedish criminal law, but the thing to be examined is the "ethics of ambiguity." Specifically, the feeling of seeing oneself and others as human beings, a formulation and a feeling inspired by Simone de Beauvoir. The practice of experiencing others and oneself as human beings is the practice of ambiguity, an anti-essentialist practice opposed to totalizing theory. It is an ethics based on the willingness to experience localized disruptions within oneself. This Article will begin to develop an ethics of ambiguity in law.[4] This Article is intended for readers with or without knowledge of civil law systems like Sweden's.

---

2. "Agitation Against an Ethnic Group" The Penal Code Brottsbalken chapter 16 § 8 [hereinafter BrB 16:8].

3. "Law 1998:408 About Prohibiting Purchases of Sexual Services" [hereinafter Law (1998:408)].

4. *See generally* SIMONE DE BEAUVOIR, THE ETHICS OF AMBIGUITY (Bernard Frechtman trans., Citadel Press 1976) (1948).

II. TOTALIZING THEORY

From a circumstantial association between economic relations and physical traits was born a new type of mark ('colour'), which had great success. Later development turned it from the traditional status of a symbol to that of a sign of a specific nature of social actors.[5]

- Colette Guillaumin

The system of marks naturalizes our understanding of the people marked as *Women* and the people marked as *Colored.* It does so by guaranteeing that the material treatment of a person is also what the person is seen to be.[6] The system of marks is located within: it is an ideology of racism and sexism. The system of marks is a description of how to look at the society to whose ruling group one belongs. The system of marks comes into existence through the eyes and mind of the master, and is executed through the actions taken by the master. The system of marks is born in his mind when he connects the way people are being treated with an essential quality that can be traced through physical features.[7] The system of marks is a hierarchy fixed to bodily marks such as sex and race. The system of marks is born when people start believing that the way they see people being treated is the reflection of an internal essential quality rather than the imposition of an external social order. The system of marks depends on an imagined connection between the essence of a person and the marks on the person's body, a physical feature or set of features, such as the marks of race and sex.[8]

The ideology of racism and sexism is a system that only works in the closed space of a racist and sexist ideology. Race and sex are circular definitions that can only be explained through the ideology upon which they are based. *Woman* and *Colored* become "real" through social practices based on the circular definitions of sex and race. "Race" in this Arti-

---

5. COLETTE GUILLAUMIN, *Race and Nature: The System of Marks, in* RACISM, SEXISM, POWER AND IDEOLOGY 133, 140 (Mary Jo Lakeland trans., Routledge 1995).

6. COLETTE GUILLAUMIN, *The Practice of Power and Belief in Nature Part II The Naturalist Discourse, in* RACISM, SEXISM, POWER AND IDEOLOGY 211, 211 (Linda Murgatroyd trans., Routledge 1995).

7. My use of the pronominal adjective "he" and the genitive "his" is to be interpreted structurally and not literally. Just as the term "master" is a structural description, so too have I used the terms "he" and "his" as structural descriptions of power relations.

8. "The sexual identification of the dualisms has both a descriptive and a normative element." Frances Olsen, *The Sex of Law, in* THE POLITICS OF LAW A PROGRESSIVE CRITIQUE 691 (3rd. ed., David Kairys ed., Basic Books 1998).

cle refers to the "biological" belief that people are divided into "natural" groups based on "race." "Sex" in this Article refers to the "biological" belief that people are divided into "natural" groups based on "sex." *Colored* refers to the "social" process of translating the "biological" belief into a social practice. *Woman* refers to the "social" process of translating the "biological" belief into a social practice.

The fact that most people treated as prostitutes are *Women* does not mean that *Women* are essentially prostitutes. The fact that most people treated as mothers are *Women* does not mean that *Women* are essentially mothers. The fact that immigrants to Sweden are not often given employment does not mean that immigrants are essentially lazy or unemployable people. Nonetheless, *Women* are viewed as essentially prostitutes or mothers just as immigrants are viewed as essentially lazy and unemployable. Jean-Paul Sartre described this need to label people according to an imagined essence as both a passion and a concept of the world.[9] It is the passion of mediocrity in that it is a passion to remain the same, to remain without learning and developing. It is a passion to always look for the same, to always find the same, and to always be the same.[10] It is the passion of mediocrity in that it is based on the fear of being visible as an individual, of being seen, of standing out.

The passion of mediocrity is a need to be superior without standing out, a need to be superior without individuality. The need to categorize and place people into groups based on bodily marks of sex and race is a need to be superior. The need to categorize is a need to belong to groups designated as superior. The need to categorize is a need to blend in with other superiors without having to risk not qualifying for the group. The qualifying process is external; it is beyond the control of the individual. To qualify for the ruling group is to be unmarked by race and sex, it is to be undifferentiated. One is qualified simply by being unmarked.

The need to mark is based on a fear of standing out as an individual. To place oneself as one among many in a ruling group based on the absence of bodily marks is to justify one's own membership in the ruling group without having to risk not qualifying. One does not have to qualify for or deserve membership because it is "nature" that has placed one on the top, just as it is "nature" that has placed the marked ones on the bottom. Membership in the privileged group is also a way to escape being seen. Membership in the ruling group is an escape from individuality and from the fear of loneliness. One can always hide in the collective. To choose membership in the ruling group is to choose "the permanence and impenetra-

9. JEAN-PAUL SARTRE, ANTI-SEMITE AND JEW 17 (George J. Becker trans., Schocken Books 1976) (1948).

10. *Id.* at 18-19.

bility of stone."[11] To choose membership in the ruling group is to value the non-changing through rejecting what would help develop a sense of existence: the constant questioning of one's being.[12]

To talk about the natural is to hide the mechanisms that uphold and re-create social hierarchy.[13] The system of marks is an ideology of social hierarchy justified by regarding submission as something within the *Woman* and the *Colored.* Submission as an essential quality of the marked is held out as an explanation of what might otherwise be explained as oppression. The system of marks requires one to look at the "natural" instead of looking at the power relationship between the marker and the marked: an oppression founded in mutually reinforcing financial, social and legal statuses. When the relationship between the oppressed and the oppressor is made "natural," the possibility of a change, and a rejection of the existing hierarchy, is made a nullity.[14]

The system of marks does not choose the person. The person chooses the system of marks. It is a choice to make the system of marks into a totalizing theory of life, a theory that expands into every part of life.[15] The system of marks is unaware of its own existence and it makes him unaware of his own existence. The system of marks blinds him from seeing others as human beings, as existences that precede their essences. It is a choice to be a stone. It is a choice to value the permanence of a stone. It is a choice to not want to will the self and others free.[16]

Simone de Beauvoir wrote that "the drama of original choice is that it goes on moment by moment for an entire lifetime, that it occurs without reason, before any reason, that freedom is there as if it were present only in the form of contingency."[17]

The system of marks not only hides the power relationship between the

---

11. *Id.* at 53.

12. "[T]he [white] feminist route to fuller understanding [of people across racial categories] depends on developing as heightened a capacity as possible for empathetically imagining another's situation while remaining acutely aware of the difficulties and possible distortions of doing so." Phyllis Goldfarb, *From the Worlds of "Others": Minority and Feminist Responses to Critical Legal Studies*, 26 NEW ENG. L. REV. 683, 687 (1991) ("For Mary Joe Frug: A Symposium on Feminist Critical Legal Studies and Postmodernism"); *see also* Anthony Paul Farley, *All Flesh Shall See It Together*, 19 CHICANO-LATINO L. REV. 163 (1998).

13. COLETTE GUILLAUMIN, *Race and Nature: The System of Marks, in* RACISM, SEXISM, POWER AND IDEOLOGY, *supra* note 5, at 133, 137.

14. "In order to prevent this revolt, [one] of the ruses of oppression is to camoflage itself behind a natural situation since, after all, one can not revolt against nature." SIMONE DE BEAUVOIR, THE ETHICS OF AMBIGUITY, *supra* note 4, at 83; *see also* Anthony Paul Farley, *The Black Body as Fetish Object*, 76 OREGON L. REV. 457 (1997).

15. *See* JEAN-PAUL SARTRE, ANTI-SEMITE AND JEW, *supra* note 10, at 17.

16. *See id.* at 53.

17. SIMONE DE BEAUVOIR, THE ETHICS OF AMBIGUITY, *supra* note 4, at 40-41.

marker and the marked; it also prevents one from seeing oneself and others as human beings. It is a system that objectifies the self and others as fixed sub-categories of human beings. It is through totalizing theory that one is robbed of the experience, of the feeling, of seeing oneself and others as human beings, instead of fixed sub-categories of human.

III. LAW-ROOMS

Legal discourse should be recognized as a site of political struggle over sex differences.[18]

- Mary Joe Frug

The political struggle over sex and race differences takes place in the void outside the two law-rooms; "Agitation Against an Ethnic Group"[19] and "Prohibiting Purchases of Sexual Services."[20] The Swedish literal translation of "lagrum," which defines a specific paragraph within a code, is "law-room." The term "law-room" refers to the closed space of a particular legal paragraph. The closed space is only open to the legal subject, provided that the legal subject qualifies for entry according to the specific requisites, stated in each law-room or in the preparatory work in connection to the law-room.

Law-rooms have many similarities with rooms of a house. To be in a house is also to be in a room. There is no place that is both outside of a room and inside of the house. It is the same with the system of law. To be in a system of law is to be in a law-room. There is no place that is both outside of a law-room and inside of the system of law. The only way to be protected from the cold wind blowing outside is to be surrounded by the walls of a room in a house. Not everyone is allowed into the warmth of the house, the entrance might be both small and gated, and some people are not wanted. Furthermore, it is not enough to pass through the entrance; there are requirements of specific behavior for all who enter. The rules also differ: the master of the house has different rules from the servant of the house. The task of the servant is to please the master. The servant is there because of the master, but a master without a servant is no longer a master. Different rules come with different identities.

Not unlike a room in a house, one law-room leads to another law-room. There is a belief that the ways in which the different law-rooms are connected follow a specific pre-fixed system that can be discovered through

---

18. MARY JOE FRUG, *A Postmodern Feminist Legal Manifesto*, *in* POSTMODERN LEGAL FEMINISM, *supra* note 1, at 126.

19. BrB 16:8.

20. Law (1998:408).

the use of legal method as a mental map. The ability to find the way around the different law-rooms is based on the ability to "think like a lawyer." Different rooms in the same house serve different functions depending on their locations and relationships to the other rooms.

Sweden has a civil law system with a static view of the hierarchy of legal sources. The legal sources within the legal method are, therefore, organized according to a fixed hierarchy. A large part of Swedish legal method is devoted to setting forth the correct "jumps" between the different levels of the hierarchy. To "jump" in the right way is to follow the legal method correctly. The hierarchy of Swedish legal sources is as follows: constitution, code, preparatory work, case law, and doctrine. The method requires the following "jumps": when the code is unclear, the lawyer's next step is to look at the preparatory work. The lawyer consults the case law after the preparatory work. The preparatory works are very important tools in Sweden's democracy. Before a major law or a law of public interest is enacted it is sent out to all interested parties for comments. Usually the interested parties are non-governmental organizations (NGOs) and state administrations with special knowledge or interest in the field that the law covers. The comments that are sent back to the legislators become part of the preparatory work, which, after the code itself, has the highest practical impact on legal implementation. This is why it is important to closely examine the perceptions operating in the responses from the NGOs and the state administrations. The purpose of the preparatory work is to allow the people living in Sweden to be involved in the legislative process and have an impact. Most of all, the preparatory work insures that the people who will be most affected by a law also will have an opportunity to affect the legislation and implementation of that law and its effects on their lives. All the empirical facts used in this Article are directly taken from the preparatory work preceding the legislation and the creation of the law-rooms.

IV. ENTERING THE LAW-ROOM PENAL CODE 16:8 AGITATION AGAINST AN ETHNIC GROUP

In the beginning, the Swedish anti-racist criminal law was intended as a protection against anti-Semitism. The law-room has its origin in the aftermath of the Second World War and the experience of the Holocaust. The law entered into force in 1948.[21] In 1970, in connection with the Swedish ratification of the International Convention on the Elimination of all Forms of Racial Discrimination, the subject of protection against ra-

21. Prop. 1986/87:151 page 104.

cism was widened to include non-white people. This was done in order to fulfill the requirements in Article 4 of the Convention.[22] The earlier anti-discrimination provision had only covered people based on heritage.[23] The law-room now looks like this:

BrB 16:8 Hets mot folkgrupp.

Den som i uttalande eller i annat meddelande som sprids hotar eller uttrycker missaktning för folkgrupp eller annan sådan grupp av personer med anspelning på ras, hudfärg, nationellt eller etniskt ursprung eller trosbekännelse, döms för hets mot folkgrupp till fängelse i högst två år eller, om brottet är ringa, till böter. Lag (1988:835).

BrB 16:8 Agitation Against an Ethnic Group.

The one who through an expression, or through other means, communicates, threatens, or expresses disdain for an Ethnic Group or any like group of peo-

---

22. INTERNATIONAL CONVENTION ON THE ELIMINATION OF ALL FORMS OF RACIAL DISCRIMINATION, adopted and opened for signature and ratification by G.A. Res. 2106 (XX) (1965), *entry into force* 4 January 1969.

Article 4.

States Parties condemn all propaganda and all organizations which are based on ideas or theories of superiority of one race or group of persons of one colour or ethnic origin, or which attempt to justify or promote racial hatred and discrimination in any form, and undertake to adopt immediate and positive measures designed to eradicate all incitement to, or acts of, such discrimination and, to this end, with due regard to the principles embodied in the Universal Declaration of Human Rights and the rights expressly set forth in article 5 of this Convention, inter alia:

(a) Shall declare an offence punishable by law all dissemination of ideas based on racial superiority or hatred, incitement to racial discrimination, as well as all acts of violence or incitement to such acts against any race or group of persons of another colour or ethnic origin, and also the provision of any assistance to racist activities, including the financing thereof;

(b) Shall declare illegal and prohibit organizations, and also organized and all other propaganda activities, which promote and incite racial discrimination, and shall recognize participation in such organizations or activities as an offence punishable by law;

(c) Shall not permit public authorities or public institutions, national or local, to promote or incite racial discrimination.

G.A. Res 2106 (XX) (1965).

23. Prop. 1970:87 page 27 paragraph 3.

ple with allusions to race, skin color, nationality or Ethnic heritage or creed, shall be sentenced as having committed *Agitation Against an Ethnic Group* to prison for no more than two years or, if the crime is insignificant, to a fine. [Trans. Author]

Four requisites constitute the architecture of this law-room: (1) a threat or an expression of disdain has to be (2) communicated with (3) an allusion to (4) an Ethnic Group. The meaning and interpretation of these requisites are perceived as being neutral within the legal method. The notion is that by using the legal method in a correct way, a correct answer will appear and contribute to a proper resolution of the legal question.

It is enough that the message is communicated to a few people, the communication does not have to be made public[24] or have been intentional.[25] An earlier requisite was that the message had to be made publicly known or communicated to the general public but this was changed in Law (1988:835).[26] The legal interpretation of the word "threatens"[27] in this law-room is wider than the criminal meaning of the word "threat" described in the law-rooms "unlawful force" and "unlawful threat" in the Penal Code.[28] "Disdain"[29] has to be carefully interpreted, it has to be clear

---

24. Karnov 2000.

25. NJA 1996 s. 577, RH 1997:28, RH 1997:53, RH 1998:77 (all the cases refer to wearing Nazi-uniforms or similar attire).

26. Prop. 1986/87:151 page 110, SOU 1981:38 page 62.

27. Karnov 2000.

28. BrB 4:4 (1) *Olaga tvång*
Den som genom misshandel eller eljest med våld eller genom hot om brottslig gärning tvingar annan att göra, tåla eller underlåta något, dömes för olaga tvång till böter eller fängelse i högst två år.

BrB 4:4 (1) Unlawful Force
The one, who through abuse or through other violence or through threat of criminal act forces another person to commit, stand, or neglect to do something, will be sentenced for unlawful force to a fine or not more than two years in prison. [Trans. Author]

BrB 4:5 (1) *Olaga hot*
Om någon lyfter vapen mot annan eller eljest hotar med brottslig gärning på sätt som är ägnat att hos den hotade framkalla allvarlig fruktan för egen eller annans säkerhet till person eller egendom, döms för olaga hot till böter eller fängelse i högst ett år. Lag (1993:207).

BrB 4:5 (1) Unlawful Threat
The one who raises arms against someone or through other ways threatens the person with an act that is criminal in a way that is intended to create a serious fear for personal or other persons safety or property within the person being threatened, will be sentenced for unlawful threat to a fine or not more than one year in prison.
Law (1993:207). [Trans. Author]

that what is communicated about the Ethnic Group clearly surpasses what could be labeled as a factual and objective debate about the Ethnic Group as such.[30] The threat or the disdain does not have to be a direct "allusion"[31] to the race, skin color, nationality or ethnic heritage or creed. A derogatory expression that has understandable or indirect connection to the race, skin color, nationality or ethnic heritage or creed falls within this provision.[32] It is not clear exactly what kind of expression falls within the meaning of allusion. The term "Ethnic Group"[33] has two meanings: First, it covers all individuals of a *homogeneous* group representing any of the listed groupings: race, skin color, nationality or ethnic heritage or creed. Second, the phrase "any like group of people" is used to clarify the fact that Ethnic Group also covers a *heterogeneous* group of people constituted by different individual races, skin-colors, nationalities or ethnic heritages or creeds.[34] Individual identifiable persons are not protected by this provision.[35] When it comes to this provision there cannot be a plaintiff.

---

29. Karnov 2000.

30. NJA 1982 s. 128 the case is about a sign in the entry of a camping location that did not allow the Romani people to enter. It was found that the sign reached the limit that defines the word disdain in this context. RH 1991:20 The case is about whether a local radio program had reached the level of disdain in its communication about the Jewish group.

31. Karnov 2000.

32. NJA 1989 s. 374 "*Jävla svartskalle.*" *Jävla svartskalle* in English means "Damn Black Head," a reference to black hair. This case was not tested as an "Agitation Against an Ethnic Group" crime, probably because it was directed towards an individual. This case, labeled a case of "Insult," states that it is an insult to be called "*Jävla svartskalle.*" The expression is a generic derogatory term used towards immigrants in Sweden. The court states that the term has an insulting meaning.

BrB 5:3 *Förolämpning*

Den som smädar annan genom kränkande tillmäle eller beskyllning eller genom annat skymfligt beteende mot honom, dömes, om gärningen ej är belagd med straff enligt 1 eller 2 §, för förolämpning till böter. Är brottet grovt, dömes till böter eller fängelse i högst sex månader.

BrB 5:3 Insult

The one who defames someone with humiliating words of abuse or through other derogatory behavior against him, shall be sentenced, if the act is not covered with punishment according to 1 or 2 §, to a fine for the insult. If the crime is severe, the sentence shall be a fine or – no more than six months in prison. [Trans. Author]

33. Karnov 2000

34. Prop. 1970:87 page 27 paragraph 3.

35. Prop. 1970:87 page 25 paragraph 2. Paragraph two states that the Swedish Penal Code does not have any protection from racism that is directed towards the individual.

V. ENTERING THE LAW-ROOM 1998:408 ABOUT PROHIBITING PURCHASES OF SEXUAL SERVICES

Prostitution was legal in Sweden until Law-Room 1998:408 about Prohibiting Purchases of Sexual Services entered into force in 1999. The question of whether to criminalize prostitution was investigated numerous times before 1998[36] but never led to full criminalization.[37]

Approximately 2,500 *Women* sell sexual services in Sweden. Approximately 650 *Women* work as street prostitutes. Each *Woman* in prostitution has, on average, about fifty customers and about 125,000 men purchase sexual services each year. During the 1980's there was a decrease of *Women* in prostitution but in 1990 the number of *Women* in prostitution began to increase. The number of foreign *Women* and the number of *Women* with mental disorders in prostitution also increased.[38] About half of the *Women* in prostitution are abusers of drugs or alcohol.[39] Ten percent of all Swedish men have purchased sexual services. These men do not, according to the committee preparing the SOU 1995:15, deviate from the "normal" man with regard to profession, financial situation, or marital status. These men, according to the committee, are deviants with regard to their views on *Women* and sexuality.[40]

The law-room About Prohibiting Purchases of Sexual Services looks like this:

> Lag (1998:408) om förbud mot köp av sexuella tjänster
>
> Den som mot ersättning skaffar sig en tillfällig sexuell förbindelse, döms-om inte gärningen är belagd med straff enligt brottsbalken – för köp av sexuella tjänster till böter eller fängelse I högst sex månader. För försök dömes till ansvar enligt 23 kap. Brottsbalken.
>
> Law (1998:408) about prohibiting purchases of sexual services
>
> The one, who against compensation acquires a temporary sexual connection, will be sentenced for the crime – if the act is not criminalized according to

---

36. SOU 1981:71, SOU 1982:61 and SOU 1995:15.

37. It has never been illegal to offer sexual services. Before 1999, however, on some occasions it was illegal to purchase sexual services. BrB 16:10 Prohibiting Purchase or Offer to Purchase Sexual Services of Person Under Age Eighteen. There were also some limitations in the law regulating social services and the the law regulating illnesses: Socialtjänstlagen (1980:620), LVU (1990:52); the Law Regulating Substance Abuse, Socialtjänstlagen (1988:870); and the Law Regulating the Risks of Transfers of Sexually Transmitted Diseases, Socialtjänstlagen (1988:1472). It was, and is, illegal to earn money from other people's prostitution. BrB 6:18.

38. One argument for criminalizing of purchasing sexual services was that the international sex industry also affects Sweden. Prop. 1997/98:55 page 23.

39. Prop. 1997/98:55 page 102.

40. Prop. 1997/98:55 page 103.

> the penal code – of purchasing sexual services to a fine or prison not more than six months. For attempts sentenced according to [law-rooms] 23 kap. Brottsbalken. [Trans. Author]

The only requisite elaborated in Proposition 1997/98:55 is what it means to be "the one" in the above law-room. This singular requisite is partly connected to the fact that in Sweden it is the act of purchase, not the sale, of sexual services that is criminal. The decision to criminalize only the purchaser, and not both the purchaser and the seller, was the object of major public debate in Sweden. The national debate is captured in the preparatory work, Prop. 1997/98:55.[41] The committee preparing the SOU 1995:15 suggested that both purchaser and seller should be criminalized. This suggestion was supported by only two remissinstanser:[42] Stiftelsen Kvinnoforum (The Foundation Women's Forum) and Polisstyrelsen i Stockholm (The Police Board of Stockholm). Some of the remissisntanser that suggested criminalization of only the purchasers were: Folkhälsoinstitutet (The People's Health Institute), Jämstalldshetsombudmannen (The Equal Rights Ombudsman[43]), Riksorganisationen för kvinnojourer i Sverige (The National Organization for Women's Hostels in Sweden[44]), and the Fredrika Bremerförbundet (The Fredrika Bremer Association). Some of the remissinstanser that completely rejected criminalization of either the purchaser or the seller were: Rikspolisstyrelsen (The National Police Board), Domstolsverket (The National Courts[45]), and Justitsiekanslern (The Attorney General). The Riksåklagaren (The National Prosecutor) showed sympathy with a criminalization of the purchaser, but found that the circumstances in favor of a criminalization were not strong enough.

In the end, only the purchaser was criminalized, even though that was

---

41. Prop. 1997/98:55 pages 18, 22, 100-101, 104.

42. Before a major law becomes a law in Sweden, the law, in the form of a proposal, often as a SOU (Statligoffentligutredning), is sent out to the leading institutions, organizations and NGOs (remissisnstanser) in the field that the law concerns. The Institutions, Organizations, and NGOs comment on the proposal and their comments become a legal source in the Proposition. This makes it possible for many people to have input on the legal system and the production of laws in Sweden.

43. An Ombudsman is an official person appointed by the Government with residence in the Parliament. The task of the Ombudsman is to give individual people, in this case people that are victims of gender discrimination, a voice. The Ombudsman is a serious appointment, and the Ombudsman works closely with the NGOs in the field. The Ombudsman is supposed to critically survey the activities of the government within the field of the appointment. The Ombudsman serves as an internal control of the government making sure that the government and the state administration follow the anti-discrimination principles and laws.

44. This is a NGO that is nationwide. It works on local levels by protecting women from abuse and helping to give them a voice.

45. This is the body that controls all the courts in Sweden.

considered to be a break of the symmetry that derived from the belief that two parties entering into a contractual situation with each other are equal. There was also a fear that the efficiency of the law-room would be low, because it only criminalized the buyer. The women in the Social Democratic party were very strong advocates for the criminalization of only the purchaser and the government was Social Democratic at that time. The Swedish government also argued for criminalization of only the purchaser and not of the seller. The government was motivated by the notion that it is unreasonable to criminalize the party that, in most cases, is the weaker party and, further, is exploited for the sexual satisfaction of the stronger party. The government also argued that it is important to enable persons involved in prostitution to seek help without risk of punishment.[46]

The law-rooms also criminalize *attempt* by referring to the law-rooms in 23 Kap. Brottsbalken that regulate *attempts* to commit acts criminalized within the Penal Code. The Government realized that the actual implementation of the law would be difficult. By criminalizing attempts to purchase sexual services, they hoped that the law would be more effective. If they could intervene as early as when the attempt to purchase sexual services was made, they hoped it would make it easier for the police to disrupt the trade on the streets.[47] The law-room BrB 23:1 (1) looks like this:

BrB 23:1 (1)

Har någon påbörjat utförandet av visst brott utan att detta kommit till fullbordan, skall han i de fall särskilt stadgande givits därom dömas för försök till brottet, såframt fara förelegat att handlingen skulle leda till brottets fullbordan eller sådan fara endast på grund av tillfälliga omständigheter varit utesluten.

BrB 23:1 (1)

The one who begins the execution of a crime without its completion shall, in cases where a specific provision so provides, be sentenced for attempt to commit the crime if there was a danger that the crime would have been completed and that such danger was averted only due to temporary circumstances. [Trans. Author]

46. Prop. 1997/98:55 page 22.
47. Prop. 1997/98:55 page 104.

VI. THE SYSTEM OF LAW AS THE SYSTEM OF MARKS

Law requires all legal claimants to assume a particular posture -a particular identity- in seeking judicial assistance; we must leave aside much of the multiplicity and complexity of our lives in order to engage in legal discourse.[48]

- Mary Joe Frug

It is difficult for law to deal with anti-essentialist concepts.[49] Indeed, the essential is a necessary condition of law's existence. The individual and the local have no place in the law-room. In the moment the law-rooms above deal with a human being, with the anti-essential, the law-rooms are no longer part of a system of law. And without law-rooms there is no system of law. Law, in that moment, leaves the general, the abstract, the totalizing that constitutes it as a system of law. In the moment the law-room falls into the individual, into the localized, it is no longer law. To be saved from the disruption of the localized, to be spared the fall from the abstract into the individual, the system of law depends on essentialist beliefs.

The servant sleeps in the servant's quarters because he is a servant: the quarters are called the servant's quarters because the servant sleeps there. A servant is a servant because the master is served by the servant. The master bedroom is the master bedroom because the master sleeps there. A master is a master because he is served by the servant. Without the servant there is no master bedroom and without the master there is no servant's quarter. There is no house if there are no rooms.

---

48. MARY JOE FRUG, *Feminist Doctrine, in* POSTMODERN LEGAL FEMINISM 12, 22 (Routledge 1992).

49. *See e.g.*, GLOBAL CRITICAL RACE FEMINISM: AN INTERNATIONAL READER (Adrien Katherine Wing ed., New York University Press 2000); CRITICAL RACE FEMINISM: A READER (Adrien Katherine Wing ed., New York University Press 1997); BLACK MEN ON RACE, GENDER, AND SEXUALITY: A CRITICAL READER (Devon Carbado ed., New York University Press 1999).

VII. THE HYPERREAL

*The law is the expression of the ideological/*practical *techniques of the system of* domination. [50]

- Colette Guillaumin

Power relationships between people are expressed in who and what are essentialized. Law's essentialism has a naturalizing effect on the power relationship between social groups. It is only in serving the master that the servant is a servant and it is only as a servant that a person is let into the servant's quarters, because only servants sleep in the servant's quarters. It is in this moment, the essentialist moment, that the system of law becomes a force for the oppression of the people marked as *Women* and the people marked as *Colored.* There is no place for shelter against racism and sexism in the system of law if one does not become what the law-room demands. It is in this essentialist moment that the system of law changes from a passive object, a vessel, a mere tool of oppression, into an active subject, an agent of oppression of people marked as *Women* and people marked as *Colored.*[51]

The system of law does not simply oppress, it actively contributes to the creation of the *Woman* and the *Colored.* It is through a liaison between the system of law and the system of marks that the *Woman* and the *Colored* become hyperreal. Guillaumin argues that the scientific community could not elevate its idea of natural groups to the level where it could receive support from the state without the system of law. It was through law that the idea of natural groups was made available for implementation through the machinery of the state. Guillaumin gives the examples of the Jim Crow laws, the Nuremberg laws, and the Apartheid laws. These discriminatory laws reached into every level of life, including marriage, housing, em-

---

50. COLETTE GUILLAUMIN, *Race and Nature: The System of Marks, in* RACISM, SEXISM, POWER AND IDEOLOGY, *supra* note 5, at 133, 140.

51. Barbara Johnson writes of Mary Joe Frug's Manifesto in a way that is instructive:

> Sentence after sentence takes the form: "Legal rules" + active verb (permit, construct, mandate, facilitate, deny, encourage, prohibit, restrict, allow, encript, compel, assign, promote, favor, deter, influence, designate, create...) + consequences for women. The cumulative rhetorical effect of these sentences is to reinforce the claim they contain: by making "legal rules" the subject of at least eighty sentences in the essay, Mary Joe Frug underlines the *agency* of the law in determining the shape of women's lives.

BARBARA JOHNSON, *The Postmodern in Feminism: A Response to Mary Joe Frug, in* THE FEMINIST DIFFERENCE 183, 188 (Harvard University Press 1998) (citing MARY JOE FRUG POSTMODERN LEGAL FEMINISM 125-153 (1992)).

ployment, and education.[52]

The system of law replaces itself through the system of marks and the system of marks replaces itself through the system of law. The system of law is a copy of something that does not exist: it is a simulation of the "natural" groups designated by the system of marks. The system of law is a simulacrum.[53] Law functions as the alibi of a naturalized social hierarchy. A fundament of law is the abstraction of the essential in the system of marks, an abstraction that takes its form in the passive legal subject. The passive legal subject is what it is based on its essence and not on its actions. The politics of essential subjects is the politics of social hierarchy. The passive legal subject is a simulation of what is believed to truly exist in the society: the essential subject.[54] The essential subject *Woman* or *Colored* is one whom others consider a *Woman* or a *Colored*. That is the simple truth from which a postmodern politics of identity must start.[55]

It is through belief in the existence of natural subjects that we create groups of sub-humans. The *Servant*, the *Woman*, and the *Colored* are what they are and the process through which they are made into *Servant*, *Woman* and *Colored* is hidden. The process is hidden by the fact that the identities: *Servant*, *Woman* and *Colored* appear to be natural. The essential classification of the group justifies its placement in the social hierarchy.[56] In this way law becomes not only a reflection of social hierarchy but also an "active agent" in the creation and maintenance of social hierarchy.[57] The system of law is the system of order upon which the system of

---

52. COLETTE GUILLAUMIN, *Race and Nature: The System of Marks, in* RACISM, SEXISM, POWER AND IDEOLOGY, *supra* note 5, at 133, 147.

53. "[N]ot unreal, but a simulacrum, that is to say never exchanged for the real, but exchanged for itself, in an uninterrupted circuit without reference or circumference." JEAN BAUDRILLARD, *The Precession of the Simulacra, in* SIMULACRA AND SIMULATION 1, 6 (Sheila Faria Glaser trans., University of Michigan Press 1997) (1981).

54. As Baudrillard writes:

> The territory no longer precedes the map, nor does it survive it. It is nevertheless the map that precedes the territory-the precession of simulacra-that engenders the territory, and if one must return to the fable, today it is the territory whose shreds slowly rot across the extent of the map. It is the real, and not the map, whose vestiges persist here and there in the desert that are no longer those of the Empire, but ours. *The desert of the real itself.*

Id. at 6.

55. I am, of course, paraphrasing Sartre's observation that: "The Jew is one whom other men consider a Jew: That is the simple truth from which we must start." JEAN-PAUL SARTRE, ANTI-SEMITE AND JEW, *supra* note 10, at 69.

56. *See* Andrea Dworkin, *Against the Male Flood: Censorship, Pornography, and Equality*, 8 HARV. WOMEN'S L.J. 1, 15-16 (1985) (writing about four elements of social subordination; hierarchy, objectification, submission, and violence).

57. COLETTE GUILLAMIN, *The Idea of Race and Its Elevation to an Autonomous*

marks depends. The system of marks is the system of order upon which the system of law depends. The system of law regulates the system of marks; it does so by relying upon the abstractions of natural groups believed-into-being through the system of marks. These products of belief upon which the system of laws is based are the so-called legal subjects. These legal subjects, the abstractions of the natural subjects in the system of marks, create a new subject: the hyperreal subject.[58] This is accomplished through a simulation in the name of law[59] and order.[60] The hyperreal is not about good or evil; it is about function, how well it works.[61] It is measured according to how much law and order it produces. The system of marks claims to be an abstraction of the "natural" social order. The system of law claims to be an abstraction of the system of marks. Together, the system of marks and the system of law create something that goes beyond the original power relationship between the people who have and the people who have not. Together, the system of marks and the system of law move beyond the distinction between the original and the copy. Together, the system of marks and the system of law bring about the realm of the hyperreal.

The system of marks claims to represent the real, which is described in the system of law. The system of law claims to represent the real, which is described in the system of marks. Together, these two systems create something that goes beyond either: the hyperreal. The hyperreal goes beyond mere distortion or disguise of the "real" relationship between those who have power to oppress and those who lack the power to effectively resist oppression. Together, these two systems, the system of law and the system of marks, each claiming the representation of the other, create the hyperreal: the copy for which there is no original, the supercession of the distinction between the real and the false. The subjects that emerge from this supercession, the *Woman* and the *Colored*, are subjects of the hyperreal, there is no distinction between a real and a false *Woman*: there is no original *Woman*. There is no distinction between a real and a false *Col-*

Scientific and Legal Status, in RACISM, SEXISM, POWER AND IDEOLOGY 61, 65-66 (Routledge 1995).

58. There are three types of simulation: the first is a copy of the real, the second is when the copy and the real take the same dimensions, and the line between the real and the copy becomes blurred, the third is beyond the distinction real and copy. It is when the image claims to be a copy, but there is no original to copy. This is the "hyperreal." RICHARD J. LANE, JEAN BAUDRILLARD 86 (Routledge 2000); *see also* Anthony Paul Farley, *Lilies of the Field: A Critique of Adjucation*, 22 CARDOZO L. REV. 1013 (2001).

59. The system of law.

60. The system of marks.

61. RICHARD J. LANE, JEAN BAUDRILLARD, *supra* note 58, at 86.

ored: there is no original *Colored.*

The law-rooms above reproduce the system of marks. The passive legal subjects, *Colored* and *Woman,* begin with their natural identities, their natural identities begin with the passive legal subjects.[62] They come into being as hyperreal subjects when, as the condition for entry into law-rooms, they become their marks. After entering the law-room it is not possible to separate true and false. There is no true and there is no false. A simulacrum is different from a copy. A copy is the non-real. A simulacrum is something different; it behaves and reacts as the real. It becomes impossible to separate a simulacrum from the real, in fact the simulacrum replaces the real and becomes the hyperreal. The passive legal subject is the simulacrum of an identity that is believed to exist naturally. The passive legal subject is a simulacrum of the simulation of existence simulated through a natural identity. In other words, the imagined legal subject precedes the natural subject, both subjects hide the fact that each is the result of the other. Both subjects are constructed through "a circumstantial association between economic relations and physical traits."[63] The system of law in its simulation of the system of marks hides the fact that there is no natural race[64] or natural sex[65] and that the hyperreal subjects, the *Colored* and the *Woman*, are created. The *Colored* and the *Woman* become representations without there ever having been an original to re-present.[66]

Catharine MacKinnon argues that *Women* own neither their identities nor the lives they lead.[67] Law, she argues, creates *Women*.[68] Further, she argues that there is no knowing what a *Woman* is, because there has never been a *Woman* not formed through sexism channeled through the system of law, based on the system of marks, specifically the bodily mark of sex.[69] So far, the only *Women* that have ever been are people who have been forced to meet one of two standards: either the standard of being

---

62. In the language of Baudrillard, an implosion is "the place where simulation begins." JEAN BAUDRILLARD, *The Precession of the Simulacra, in* SIMULACRA AND SIMULATION, *supra* note 53, at 86.

63. COLETTE GUILLAUMIN, *Race and Nature: The System of Marks, in* RACISM, SEXISM, POWER AND IDEOLOGY, *supra* note 5, at 133, 140.

64. The assumption of a natural race is that *Coloreds* give birth to *Coloreds* and non-*Coloreds* give birth to non-*Coloreds*. *Id.* at 136.

65. The assumption of a natural sex is to assume that *Women* give birth to *Women*, and men give birth to men. *Id.* at 136.

66. *See generally* Maria Grahn-Farley, *Not for Sale! Race & Gender Identity in Post-colonial Europe* 17 N. Y. L. SCH. J. OF HUM. RTS. 271 (2000).

67. CATHARINE A. MACKINNON, *On Exceptionality: Women as Women in Law, in* FEMINISM UNMODIFIED: DISCOURSES ON LIFE AND LAW 70, 71 (Harvard University Press 1987).

68. *Id.*

69. *Id.*

equal to men, or the standard that men have created for *Women* to meet.[70] There is no state of being before the state of being *Colored.* Coloring erases all previous pasts. The violence of objectification finds its hiding place in the concept of difference.[71] Left visible are the myriad submissions that take the shape of the so-called "voluntary" and which, in turn, make the social hierarchy look natural.[72]

VIII. THE PASSIVE AND THE ACTIVE LEGAL SUBJECT

When examining the law-room "Agitation Against an Ethnic Group" the contours of the active legal subject come into light. The active legal subject is everything that is not described by the law-room as the Ethnic.[73] He is not of a specific race, he does not have a skin color, he does not have a national or Ethnic heritage or a creed. The reason why he is a negation is because his body has no importance: his essence is irrelevant. It is through his actions that the active legal subject becomes a legal subject. It is action that constitutes him as a legal subject. It is by threatening, and communicating the threat to other people, that a social subject starts down the road to becoming an active legal subject. It is the action that is evaluated, not the active legal subject.[74] The action of the active legal subject has to meet a qualitative standard: the action has to correspond to the essence of the passive legal subject. It is not enough that the action expresses disdain along with an allusion to the Ethnic, - a specific race, skin color, nationality, heritage, or creed, if the people referred to would not essentially be an Ethnic - a specific race, skin color, nationality, heritage, or creed. The perceived essence of these people precedes the action itself.

The Ethnic, however, does not become a legal subject through its own

---

70. *Id.*

71. Nadine Taube and Elizabeth M. Schneider look at the ways Anglo-American law has contributed to keeping *Women* as subordinates by explicitly excluding *Women* from the public space and by refusing to regulate the private space, where *Women,* as a result of their exclusion from the public space, are confined. Taube and Schneider also describe the way that the law helps to legitimize sex-discrimination through the ideology of difference between men and *Women. See* Nadine Taube & Elizabeth M. Schneider, *Women's Subordination and the Role of Law, in* THE POLITICS OF LAW: A PROGRESSIVE CRITIQUE 328 (David Kairys ed., Basic Books 1998).

72. COLETTE GUILLAUMIN, *The Idea of Race and Its Elevation to Autonomous Scientific and Legal Status, in* RACISM, SEXISM, POWER AND IDEOLOGY 29, 64 (Routledge 1995); Dworkin, *supra* note 56, at 15-16.

73. "Identities privileged by whiteness and masculinity remain non controversial, invisible, undiscussed, and therefore cloaked in privacy." Lisa C. Ikemoto, *In Sisterhood Mari Matsuda, Where Is Your Body?* 2 MICH. J. OF RACE & L. 513, 517-518 (1997).

74. The active legal subject is its actions. The passive legal subject, on the other hand, is its mark.

actions. There is no action or qualitative standard of action that constitutes the Ethnic. It is the bodily mark of Ethnicity itself that constitutes the qualifying mark of the Ethnic.

The construction of the passive legal subject starts with recognition by other individuals who have received the threat that the active legal subject communicated to them. It is irrelevant whether the threat was ever communicated to the Ethnic. It is irrelevant for the constitution of the passive legal subject whether the Ethnic understands or in any way is aware of what was communicated. One Ethnic is not equal to one without Ethnicity, a non-Ethnic. The Ethnic has to be a collective, an Ethnic Group, to become a legal subject in the law-room: "Agitation Against an Ethnic Group." There is no law-room offering protection within the Swedish criminal law for the individual Ethnic.[75] It is only when falling into the anonymous sameness of the Group that the door to the law-room might open. To qualify as a Group, the individuals have to be considered by the system of law to belong to a specific group within the system of marks: marks of races, skin colors, nationalities or heritages or creeds must therefore be placed on their bodies. Without a bodily-marked collective, the Group is not recognized as a legal subject and is not let into the protection of the law-room "Agitation Against an Ethnic Group." This law-room takes the meaning of the Other to its ultimate extreme. It is literally every social subject that might at any moment be possible to classify as the Other, and all the Others are in the law the same. The meaning of Ethnic Group goes beyond the internal interpretation of the system of marks; it puts all the sub-groups into one group. Ethnic Group means a group constituted by social subjects that can carry any of the bodily marks that label and place them below the group of white-Swedish-Protestant-men on the top. Every separate Other is the same as all the other Others.

"The Law-room 1998:408 about Prohibiting Purchases of Sexual Services:" makes it clear that to purchase sexual services is the same as being a man.[76] According to this law-room, to be purchased for the purpose of sexual services is the same as being a *Woman*.[77] Men purchase the bodies of *Women*. *Women's* bodies are purchased by men. To be a man is to be a purchaser. To be a *Woman* is to be a prostitute.[78] The active legal subject

---

75. Prop. 1970:87 page 25.

76. Prop. 1997/98:55 page 103.

77. Prop. 1997/98:55 page 23.

78. The idea of natural groups from the eighteenth century up to the present day-the chief factor among those affecting social relationships consist in ownership of oneself .... Specific cases apart, one could be excluded for life from ownership of oneself by slavery, serfdom or sex. [P]artial and temporary exclusion do not carry with them a total naturalistic sentence against the group in question.

COLETTE GUILLAMIN, *The Idea of Race and Its Elevation to an Autonomous Sci-*

comes into being through his actions. It is through the act of purchasing the sexual services of the Prostitute that he comes into existence as a legal subject.

The word Prostitute is frequently used[79] throughout the proposition as a substitute for the word seller, but the person that purchases sexual services is never described by anything other than his actions.[80] "So long as the prostitute is denied the rights of a person, she sums up all the forms of feminine slavery at once."[81] His actions define him as the buyer, he is his actions. It is also his actions (the buyer's) that define her as a prostitute, she is *his* actions. This mirrors Guillaumin's argument that to be a noun is a sign of your placement on the bottom in the social hierarchy.[82]

The Prostitute is the essence of a prostitute; she exists through time and space as a receiving sign. The law-room presumes that a prostitute is a Prostitute even without acting as a prostitute. The attempt to purchase a sexual service from a prostitute is criminal, "For attempts sentenced according to 23 kap. Brottsbalken."[83] There is no room for mistake; a *Woman* is or is not a prostitute. It is not the actual act of selling sexual services that makes a *Woman* into a prostitute: her essence fixes her identity before any action. For the purchaser to become a legal subject in the law-room as early as in the moment of attempt he has to attempt to purchase sexual services from a real Prostitute. It is not criminal to attempt to purchase sexual services from a non-Prostitute. The door to the law-room "Prohibiting purchases of sexual services" opens up to the next law-room, BrB 23:1, that regulates the attempt to commit the crime when there is "a danger that the crime would have been completed and that such danger was averted only due to temporary circumstances."[84]

If the person approached was not a Prostitute then there was no risk that attempt might have become a completed crime. The only way to know if there was a risk of the crime's completion would be to know whether the person was a Prostitute. Without an essentialist approach it is not possible to know whether or not there was a risk of completing the purchase and

---

entific and Legal Status, in RACISM, SEXISM, POWER AND IDEOLOGY, *supra* note 67, at 61, 73.

79. The word prostitute was used either on its own or in combination with "street-prostitute" thirty-nine times. Prop. 1997/98:55.

80. He is never called "John." Prop. 1997/98:55.

81. SIMONE DE BEAUVOIR, THE SECOND SEX 556 (H.M. Parshley ed. trans., Vintage Books 1989) (1953).

82. COLETTE GUILLAMIN, *The Specific Characteristics of Racist Ideology, in* RACISM, SEXISM, POWER AND IDEOLOGY 29, 50-51 (Andrew Rothwell and Max Silverman trans., Routledge 1995).

83. Law (1998:408) last paragraph.

84. BrB 23:1

sale of sexual services. The constituting moment for the active legal subject "purchaser" is the attempt to purchase sexual services from a Prostitute. Thus, the law-room relies on the essence of being a Prostitute. The law-room relies on the possibility that one can be a Prostitute without an act of prostitution: one is or is not a Prostitute based on whether one is, in essence, a Prostitute. The law-room assumes that it is possible to know who is and who is not a Prostitute without reference to an act of prostitution. The Prostitute reveals herself through her essence and not her actions: her essence precedes her existence.

The law-room "Agitation Against an Ethnic Group" and the law-room "Prohibiting Purchases of Sexual Services" are built in similar ways. Both law-rooms have one active legal subject, a subject that comes into existence as a legal subject when its actions are perceived to treat the *Woman* and the *Colored* in a manner that corresponds to their essence as W*oman* and as *Colored.* A *Prostitute* is a *Woman*, a *Woman* is treated as a *Prostitute*, an *Ethnic* is a *Colored*, and a *Colored* is treated as an *Ethnic*. The passive legal subject comes into being when the active legal subject materially treats it according to the bodily marks it bears. The passive legal subject comes into being when it is treated according to its "nature."

IX. THE HIERARCHY OF MOVEMENT

The social hierarchy is reproduced through law by its enforcement of external limitations on bodies according to different essential classifications of groups. Duncan Kennedy writes about the conflict of interests between men and women over anti-discrimination law.[85] Kennedy argues that there is a zero sum game between the sexes. He argues that if the anti-discrimination law actually functioned to prevent sex discrimination against women, then a lot of men who think of themselves as normal men, and non-sex-harassers, would find themselves affected by the anti-discrimination law, a law that they perceive to be for other types of men. There is a zero sum game between the sexes because effective anti-discrimination law would, necessarily, be regarded as unjustly sanctioning "normal" men.

There is a correlation between externally forced limitations and one's place in the hierarchy.[86] The same law-rooms that are supposed to prevent the most abusive and exploitative actions in a hierarchical society also have the exclusionary requirement of a coherent subject. One must present

---

85. DUNCAN KENNEDY, *Sexual Abuse, Sexy Dressing, and the Eroticization of Domination, in* SEXY DRESSING ETC. 126, 147 (Harvard University Press 1993).

86. *See* Zanita E. Fenton, *Mirrored Silence: Reflections on Judicial Complicity in Private Violence*, 78 OR. L. REV. 995 (1999).

oneself as a coherent subject in order to qualify for the role of the passive legal subject. The passive legal subject, which one must become in order to enter the law-rooms, reproduces the very abuses and exploitations from which it is intended to protect its occupants. The passive legal subject that seeks protection in the law-room functions as the very limitation on the protection that will be provided by the law-room. The law-room protects its occupants only from that which is not already present in the mark *Woman* or *Colored.*

The law-rooms "Agitation Against an Ethnic Group" and "Prohibition Against Purchases of Sexual Services" came into being because people are both racist and sexist, and those racist, sexist people have the power to act on those values. Racism and sexism function as external limitations upon the people marked as *Colored* and marked as *Women.* The people marked as *Colored* and marked as *Women* are forced to adjust their behavior and movements according to the racism and sexism they encounter. To have to submit to external limitations also means to be the victim of subordination and to be a subordinate. The race and sex struggle is not only over the span and the effectiveness of these law-rooms. The core of the struggle is about subordination: Who has to follow whose house rules.

To be forced to accept external limitations, through the system of marks or the system of law, is to be subordinated. The external limitation on the behavior and movement of the self is also a symbol of subordination. External limitations are for the people on the bottom of the social hierarchy, not for the people on the top. The fact of external limitation is also the sign of your place in the hierarchy: to be the subordinated is to be on the bottom. To be forced to follow someone else's house rules is to not be your own, to not own yourself.

Effective laws against sexist and racist behavior would mean a redefinition of what it means to be externally limited. In such law-rooms the "normal" white-Swedish-Protestant-man would also be externally limited, which is the sign of subordination. And because no "normal" white-Swedish-Protestant-man is ever to be subordinated by anyone other than his equal, that is, another "normal" white-Swedish-Protestant-man, such law-rooms have not been built.

The demand for coherence[87] is a way to limit the efficiency of the law-

87. For critical approaches to the law's demand for coherence, see generally Kimberlé Crenshaw, *Demarginalizing the Intersection of Race and Sex: A Black Feminist Critique of Antidiscrimination Doctrine, Feminist Theory and Antiracist Politics*, 1989 U. CHI. LEGAL F. 139 (1989); Kimberlé Crenshaw, *Mapping the Margins: Intersectionality, Identity Politics, and Violence Against Women of Color*, 43 STAN. L. REV. 1241 (1991); Angela P. Harris, *Race and Essentialism in Feminist Legal Theory*, 42 STAN. L. REV. 581 (1990).

room and thus prevent it from limiting the "normal" white-Swedish-Protestant-man's behavior towards people marked as *Colored* and people marked as *Women*. This is because external limitations are only for "abnormal" white-Swedish-Protestant-men. External limitation is a sign of subordination. Racism and sexism, for example, are signs of subordination: they mark the external limit on what people marked as *Colored* and people marked as *Women* can do. Indeed, one is not even allowed to enter the protective law-rooms without first presenting oneself as a *Colored* or a *Woman*, that is, without presenting oneself as the unambiguous sign of subordination: the mark of race and the mark of sex.

X. EXITING THE LAW-ROOM

There is no existence as a legal subject in the void outside the law-room. Just as one cannot exist outside the rooms of a house and still remain inside the house, it is impossible to exist outside the law-rooms and still remain inside the system of law. To belong to the group marked as *Women* or the group marked as *Colored* and to reject the bodily mark is also to reject being a subject in the system of law. To exist outside the system of law is also to reject the protection the system of law gives *Women* and *Colored*. In the Void, the brute force and violence of the system of law reach the surface. In the Void, the violence against people marked as *Women* and marked as *Colored* appears unveiled.[88]

First, the system of law, through its need of a passive legal subject based on the bodily marks, that is, the essential, creates the *Colored* and the *Woman*. Once created as the *Colored* and as the *Woman*, the subject then needs protection from the oppression exercised through the system of marks by the non-essentialized group. To reject essentialization is to exist in the Void outside the law-room. It is the same as not existing as a legal subject. There is no place outside of the law-room that remains inside of the system of law. To not be a legal subject is to not be a social subject. It is to not be.

The system of law replaces itself through the system of marks and the system of marks replaces itself through the system of law. The system of

---

88. Today we know perfectly well (as we probably always did, but what we know and what we are prepared to acknowledge are not always the same thing...) that any physical characteristic whatsoever can be made into a 'discriminator' in some socially or politically motivated system of classification (by opposition to a disinterested scientific one). The choice of somatic criteria is symbolic of the intentions of the classifiers, and nothing more.

COLETTE GUILLAUMIN, *I Know Its Not Nice, But... ' The Changing Face of 'Race', in* RACISM, SEXISM, POWER AND IDEOLOGY 99, 100 (Andrew Rothwell and Max Silverman trans., Routledge 1995).

law begins and ends with the system of marks. The system of marks begins and ends with the system of law. There is no separation between the system of law and the system of marks. The one is the other. To not be a social subject is to not be a legal subject. To be outside the law-room is to stand naked, without the protection and recognition that the essentialized subject needs in order to be. The deal is clear: either one accepts the essentialization of one's body and receives the partial protection of the law-room or one rejects the essentialization of one's body and loses the whole protection. Accept the naturalness of one's placement on the bottom of the social hierarchy or stand naked in the cold, outside of the protective warmth of the law-room.

XI. LOCALIZED DISRUPTION: A STRATEGY OF LIBERATION

To want existence, to want to disclose the world, and to want men to be free are one and the same will.[89]

- Simone de Beauvoir

To embrace ambiguity is not only a question of ethics. It is a strategy intended to free us from the blinding system of marks and to enable us to "imagine a world of less inequality and anguish."[90] It is a strategy about how to move from the totalizing center to the localized individual. It is about developing a practice wherein the fear of existing in ambiguity becomes a joyful way of being. It is a practice of existence in the external unlimitedness of one's strengths. It is about being aware of existence and, at the same moment, existing.

The continuation of the world of the oppressor, the marker, is a linear development. It is a world that not only today places the people marked as *Colored* and the people marked as *Women* on the bottom, it is a world that will continue to keep the people called *Colored* and the people called *Women* on the bottom of the social hierarchy. The world will continue to be carried forward by the marker. The only way to change this is a revolution that breaks the harmonious line of development that the marker has in mind and in which he projects himself into the future as the active, the one on the top, and *Colored* and *Women* as the passive, as the ones on the bottom. To be able to break the line and transform the revolution from

89. SIMONE DE BEAUVOIR, THE ETHICS OF AMBIGUITY, *supra* note 4, at 86-87.

90. Phyllis Goldfarb, *From the Worlds of "Others": Minority and Feminist Responses to Critical Legal Studies*, 26 NEW ENG. L. REV. 683, 689 (1991). Goldfarb writes, "White feminists and people of color,... would clearly benefit from a project that theoretically deconstructs the underpinnings of social arrangements characterized by gender, race, and class-subordination and that frees thinkers to imagine a world of less inequality and anguish." *Id.*

ideology into concrete actions the marked has to think about the marker as an "obstacle and a thing."[91] It is only by thinking of the marker as an obstacle and a thing that the marked can relate to the relationship between the marker and the marked with indifference and without passion. If the marked start thinking of the marker as anything other than a thing and an obstacle, they will become passionate about the marker and dependent on the very being of the marker. The marker then becomes crucial to the marked for his own being. This means that the marked will need the marker to be, in order to be able to be himself. Therefore revolution can only take place by getting rid of both the marker and the marked.

The system of law begins in the system of marks. The system of marks begins in the system of law. It is impossible to separate one system from the other system. Each system is the other system. This is why the strategy of liberation from oppression has to take place in both the legal as well as the social. It is by transforming both systems that each system can be transformed. The legal is the social and the social is the legal.

XII. LEGAL STRATEGY

The problem is not so much about squeezing incoherent subjects into law-rooms. Law-rooms often deal with incoherent legal subjects, such as the people marked as white-Swedish-Protestant-men. It is not a problem to fit the incoherent subject white-Swedish-Protestant-man into the legal definitions. The reason why it is not difficult to fit the incoherent social subject, white-Swedish-Protestant-man, into the law-room that can make him into a legal subject is because he does not need to be a white-Swedish-Protestant-man to fit into any law-room. He is what his action is. He is action.

The *Woman* and the *Colored* are not only bodily marks of sex and race; they are also signs that define the meaning of subordination. To be socially subordinated is also to be essentially subordinated. The same law-rooms that are supposed to combat the subordination of the people marked as *Women*, and the people marked as *Colored*, are also the law-rooms that help subordinate them. The problem of the incoherence of the subject is only an excuse to keep the existing social hierarchy. The problem of the incoherence of the subject is the excuse for not changing the meaning of subordination. The fact that the white-Swedish-Protestant-man, which is also an incoherent identity, qualifies as a legal subject shows that the problem of the law-rooms is not the problem of incoherence. Rather, the problem of the law-room is the order of things in the social hierarchy. The problem is connected to the power relationship between the people on the

91. SIMONE DE BEAUVOIR, THE ETHICS OF AMBIGUITY, *supra* note 4, at 83-84.

top and the people on the bottom. The problem is about the power to exist above the force of essentialism.

The system of law participates in un-doing and preventing the existence of human beings. The need of the system of law for the abstraction of the legal subject creates and is a creation of the essential social subject. This is also what creates the *Woman* and the *Colored* in the above law-rooms discussed herein as the Prostitute and as the Ethnic. It also determines the set roles of the people marked as *Women* and the people marked as *Colored.*

The legal cannot exist without the illegal. It is through legal reasoning that we find the illegal.[92] The line between legal and illegal is not a line given by "nature." It is made by man. The line is where it is through the will of man and the power of man to implement his will.[93] Likewise, the architecture of the law-room is not given by nature. It is made by man. The architecture of the legal system is made by man and manifests the will of man, represents the will of man, and represents the power of man to implement his will.

To be the subordinated is in itself a sign of one's place in the social hierarchy. The demand for a coherent identity to fit the role of the passive legal subject is itself a sign of subordination. There is no essentialist demand for a coherent identity to be an active legal subject. There is no requirement of essence, only action, to become an active legal subject.

The passive legal subject is not needed to make the law-room functional. The only function of the passive legal subject is to limit the scope of the law-room, and its effects on the active legal subject. The passive legal subject is not built into the law-room as a protection of the people that are made to fill the passive role. The passive legal subject is built into the law-room to limit the impact of the law-room on the active legal subjects. In other words, the passive legal subject serves as an external limitation on the limitation of the freedom to move and act of white-Swedish-Protestant-men, who fill the role of the active legal subject. The two law-rooms I have examined do not represent efforts to protect against or reverse the social hierarchy. They serve as the guarantors of the persistent social order.

---

92. JEAN BAUDRILLARD, *The Precession of the Simulacra, in* SIMULACRA AND SIMULATION, *supra* note 63, at 1, 14-15.

93. *See* Grahn-Farley, *supra* note 66, at 296.

XIII. SOCIAL STRATEGY

The system of marks prevents localized disruptions from taking place. It substitutes the moment of ambiguity with the continuing security of totalizing theory. The system of marks is a form of totalizing theory that prevents the moment in which one might develop one's own inner voice, one's own ethic of existence. Instead of the moment of ambiguity, there is the continuation of security in the unawareness of existence.

To achieve freedom, the people marked as *Colored* and the people marked as *Women* have to break free from the linear thinking set out by their master and oppressor. Freedom is about disrupting the harmony. Harmony that maintains the law and order that reproduces the present social order must be disrupted. Freedom is about refusing to be the sign of subordination. Freedom is about getting rid of both the marker and the marked.

To be able to break free one must change one's way of looking at the oppressor. One must begin to look at the oppressor as the *thing* and the *obstacle*. Freedom is about becoming indifferent to the intent and being of the master and oppressor. Freedom is about having the courage to question the fixed positions in the power relationship between the people who have and the people who have not. Freedom is about never feeling done. Freedom is about becoming aware of existence in order to escape categorizing people in sub-human groups. Jean Baudrillard writes that it is traumatic to lose one's reference points.[94] To exist takes courage and can be painful. It is in the questioning, in experiencing the ambiguity between the fixed positions in the power relationship between the marker and the marked, that consciousness of one's own existence emerges from within. The ethics of ambiguity is a way to see the human being in others as well as in oneself.

It is liberation for all if the coherent view of the subordinate becomes fragmented. Fragmentation is the liberation or the setting free of both sides of the now apparently fixed paradox. Without a coherent image of subordination, the image of supremacy will also fragment. Since the image of supremacy is in many ways fragmented, it owes its existence only to the belief in its own coherency.

---

94. JEAN BAUDRILLARD, *History: A Retro Scenario, in* SIMULACRA AND SIMULATION 43, 44 (Sheila Faria Glaser trans., University of Michigan Press 1997) (1994).

XIV. CONCLUSION

The system of law replaces itself through the system of marks. The system of marks replaces itself through the system of law. There is no separation between the system of law and the system of marks. In the implosion, the beginning is the end and the end is the beginning. The system of law and the system of marks become impossible to separate.[95] To deny the system of marks is to deny the system of law and to deny the system of law is to deny the system of marks. There is not one without the other. The path to a society in which people regard others and themselves as human beings is a question of ethics. It is a question of choice. It is a question of rejecting the totalizing system of marks and the totalizing system of law and the way that our understandings of others and ourselves are fixed. It is a question of exiting the hyperreal. It is an ethics of ambiguity.

The ethics of ambiguity is an invitation to always question and to always remain in the position between the poles. It is about always questioning one's way of being and others'. It is about breaking the harmonious line projected by the masters and instead, through the choice of constant questioning, creating a localized disruption from within. To get rid of the marker is the same as getting rid of the marked. It is about liberating the being of the marked from the dependence on the marker.

A legal change can only come through a social change and a social change can only come through a legal change. A legal strategy is also a social strategy in the same way that a social strategy is also a legal strategy. To focus only on the system of law as the site of change will fail because the simulacrum of the natural identity contributes to the creation of a hyperreal identity. To focus only on the social, the system of marks, as the site of change will fail because it is through the legal that the marks placed on the body become essential and natural. Hyperreal identities cannot be called true or false. There is no way to separate the true from the false in the simulacra.

It is by making a change from within, in the personal, through localized disruptions, that we can reject the totalizing theory of marks and law. It is by having the courage to exist in ambiguity, to embrace an ethics of ambiguity, that one can create a personal voice and reject the totalizing theory of marks and law. The ethics of ambiguity is a personal localized ethics, a disruptive ethics, and a repudiation of totalizing theory. This is a psychological strategy of being: a practice of existing as a human being.

---

95. JEAN BAUDRILLARD, *The Precession of the Simulacra, in* SIMULACRA AND SIMULATION, *supra* note 63, at 31.

It is the foundation of the system of law, not its failing, to create and demand essential coherent identities from the oppressed. Rejection of one is also the rejection of the other. There is no law to be saved in the struggle against the oppression of people marked as *Colored* and the people marked as *Women*. The system of law is the system of marks.

The law does not have a problem in dealing with incoherent social subjects. In the law-rooms "Agitation Against an Ethnic Group" and "Prohibiting Purchases of Sexual Services" the law has no problem in dealing with the incoherent subject, white-Swedish-Protestant-man, a subject as incoherent as any. But in the law-rooms this subject becomes the active legal subject. The qualification for becoming an active legal subject does not reside within one's identity; the qualification to become an active legal subject is to act.

The problem is not a problem solved by reference to essence. The problem is the social hierarchy that has been justified through the system of marks, a system that naturalizes the existing social hierarchy. The problem is about the politics of essentialism, operationalized through the system of law. The purpose and meaning of both systems, the system of marks and the system of law, is to mask, justify, and violently uphold a hierarchy. Since the hierarchy is hyperreal it is well able to predetermine the subject positions of its occupants, rendering those it disfavors "incoherent" and those it favors "coherent." But the logic of hyperreality is politics: a politics that places the subject white-Swedish-Protestant-men on the top. All the rest fall to the bottom, into the sameness of the *Woman* and the *Colored*. That which we label "incoherent" is merely that which has been imprisoned by the very law-room to which it has turned for sanctuary.

Foster Care: The Border of Family Identity Maintaining, (Re)creating, Destroying

Zanita E. Fenton*

INTRODUCTION

This session on foster care, in title, attempts to reflect the postmodern approach to life and legal theory of Mary Joe Frug. Where this title attempts to engage us in a postmodern approach to the subject, the underlying reality of foster care is one that could only be understood as postmodern. Defining family and, in many instances care, is elusive for those who participate in the system of foster care and for the children who are subjected to it. These participants include: the legal parent(s) (natural or adoptive), the foster parents or legal guardians, judges and court administrators, legislators, lawyers, social workers, and the children. In this list, I include children last because that is often the way the system effectively considers them; though the system was created *for* children, it is too often the case it is only *about* them. I also placed children last so that they will stay at the forefront of our conversation. I have not directly identified mothers, either biological or foster, as they are not explicitly considered in the policy objectives of the policy approaches. Though unstated, it is the unfulfilled maternal stereotypes and the vilification of impoverished and minority women that fuels the underlying purpose of the approaches.[1]

The purpose of our session was to have a conversation about future directions for policy. There have been roughly two approaches to foster care policy in American history: the "child rescue orientation" and the "family support approach."[2] I asked conference attendees, and now I ask the

---

* Associate Professor of Law, Wayne State University Law School; A. B., Princeton University, 1990; J.D., Harvard Law School, 1993.

1. Professor Frug argued that the legal rules are structured to permit and sometimes mandate the "*terrorization* of the female body," the "*maternalization* of the female body," and the "*sexualization* of the female body." Mary Joe Frug, *A Postmodern Legal Manifesto (An Unfinished Draft)*, 105 HARV. L. REV. 1045, 1049-50 (1992). In the specific context of social services, specifically foster care, these three observations most assuredly combine and intensify.

2. Patricia Schene, *Past, Present, and Future Roles of Child Protective Services in Protecting Children from Abuse and Neglect*, *in* 8 THE FUTURE OF CHILDREN 23, 24

reader, to imagine themselves as members of an advisory board for a foster care facility and to consider the discussions of the speakers/authors as experts on their topics. To ground our policy considerations in some sense of reality, I asked the participants to consider the sample case studies,[3] included at the end of this introduction, in their proposals for policy objectives.

In addition to those cases, we also considered the broad topics that affect foster care, such as abuse and neglect, drug use, unemployment, health care needs, and child care issues. Among the solutions explored, in addition to traditional placements, were kinship placements and open adoptions.

BRIEF HISTORY

State intervention for the protection of children is relatively recent in historical terms. During most of the nineteenth century, primarily destitute white children were sent to institutions operated by private charitable institutions or almshouses.[4] A home specifically for the care of children, the Children's Aid Society, was founded by Charles Loring Brace in New York in 1853. Later, a model of care which is considered to be the precursor to the modern-day child protective services, spread throughout the Northeast and the Midwest.[5] Between 1920 and 1950, the modern system of child welfare emerged.[6]

Of course, the history of foster care is not unitary because American history includes divergent racial histories. For example, the institution of slavery not only served to maintain black children, but also, in most instances, created the need for their care by dismantling existing families.[7] Ratification of the 13th Amendment not only ended formal slavery, but

---

(Spring 1998). Both Susan Mangold and Richard Wexler discuss these approaches in their symposium pieces in this section. *See* Susan Vivian Mangold, *Transgressing the Border Between Protection and Empowerment for Domestic Violence Victims and Older Children: Empowerment as Protection in the Foster Care System*, 36 NEW ENG. L. REV. 69 (2001) (discussing the approaches as a "tension between parental empowerment/autonomy and child protection"); Richard Wexler, *Take the Child and Run: Tales From Age of AFSA*, 36 NEW ENG. L. REV. 129 (2001) (taking specific issue with the modern manifestation of the "child rescue orientation" exhibited by the Adoption and Safe Families Act).

3. I note that these are only samples and cannot reflect the vast and diverse sets of circumstances of each child, or even the majority of children.

4. *See* Schene, *supra* note 2, at 25.

5. *See id.*

6. *See id.* at 26-27.

7. This statement easily brings to mind the current welfare system. Perhaps this criticism is harsh, but it should not go unheeded.

also ended indentured servitude as a means of child care for white children. After emancipation, ratification of the 13th Amendment, and Reconstruction, the advent of "Jim Crow" laws prevented the care of black children by formal institutions in white society.[8] Even upon the dismantling of legal segregation, the children of black families were turned away from services because they were considered "unadoptable,"[9] and later identified within the "special needs" category of formal institutions.[10] This is not to say that black children did not benefit from foster care, it was just most often informal through extended family networks.[11] Race must be included race in the history of foster care because it continues to have an impact on the system and necessarily must be part of policy formation.

SELECTED CURRENT STATISTICS

There are approximately 600,000 children in foster care today.[12] Of these children, approximately 110,000 are waiting to be adopted, with the majority (57%) waiting in foster homes:[13]

By Gender: Male (52%); Female (48%)

---

8. *See* JOYCE A. LADNER, MIXED FAMILIES: ADOPTING ACROSS RACIAL BOUNDARIES 67 (1977); *see also* ANDREW BILLINGSLEY, BLACK FAMILIES IN WHITE AMERICA 69-71 (1968).

9. *See generally* DAWN DAY, THE ADOPTION OF BLACK CHILDREN: COUNTERACTING INSTITUTIONAL DISCRIMINATION (1979).

10. Modern definitions of special needs generally include age of the child, race, and mental or physical disability, but differ from state to state. Special needs is broadly defined under the Adoption and Safe Families Act. *See* 42 U.S.C. §673(c) (2001).

11. *See* Zanita E. Fenton, *In a World Not Their Own: The Adoption of Black Children*, 10 HARV. BLACKLETTER J. 39, 42-43 (1993).

12. As of the end of 1998, there were 560,000 children in foster care. *See* U.S. Dept. of Health and Human Services, Administration for Children and Families, *Children in Foster Care, 1983-1998*, *available at* http://www.acf.dhhs.gov/news/stats/fc.htm (last visited June 11, 2001). The number of children in foster care has consistently gone up since 1983, so 600,000 is an extrapolation from this trend. *See id.*

13. *See* U.S. Dept. of Health and Human Services, Administration for Children and Families, *Children's Bureau AFCARS REPORT*, *available at* http://www.acf.dhhs.gov/programs/cbpublications/afcars/rpt0199/ar0199e.htm (last visited June 11, 2001) [hereinafter *AFCARS Report*].

14. *See id.*

15. *See id.*

By Race: White (28%); Black (56%); Hispanic (9%); American Indian/Alaskan Native (1%); Asian/Pacific Islander (1%); Unknown/Unable to Determine (5%).[14]

The mean age of children removed from their parents was 4.4 years with the median age at 3.3 years. The mean length of time in foster care was thirty-eight months, and the median was thirty months.[15] Of the 124,500 children exiting foster care, 60% were reunited with their families, 13% were placed with other relatives, 12% were adopted and 5% were emancipated.[16] Approximately 37,000 of children in foster care had parents whose parental rights were terminated.[17]

CASE STUDIES[18]

SAM

When Sam was almost two, his single mother married. A year later she gave birth to Jack. Sam's mother was cool and aloof with him, viewing Sam as a difficult child to parent. She frequently used harsh methods of discipline, such as tying his hands to a chair and leaving him there for several hours. She also belittled him and compared him unfavorably with Jack. Sam's stepfather was more nurturing and developed a relationship with Sam that was satisfying to them both. When Sam was almost five, and Jack two, his stepfather left the family because he could not tolerate his wife's parenting style. He was awarded custody of his son Jack. Sam remained with his mother and did not have any contact with either his stepfather or his brother. Within months, a substantiated report of physical abuse was made against Sam's mother, and Sam was placed in a non-relative foster home. After experiencing three disruptive foster home placements and one disrupted adoption, Sam is currently living with a foster family while he awaits adoption by a new, recently identified family. He is ten years old.

---

16. *See AFCARS Report*, *supra* note 13. Of the remaining children exiting foster care 2% were the subject of a guardianship, 3% were transferred to another agency, 2% were runaways, less than 1% was due to the death of the child, and 1% was other. *See id.*

17. *See id.*

18. The case studies are based on real foster cases with minor alterations for simplicity. The session used these cases as a point of departure in discussing broad policy objectives. We wanted the participants to be as engaged in the discussions as possible.

CARA

Cara is eleven years old. Cara first came into care three and a half years ago following a fire that left both she and her mother badly burned. Cara's mother suffered a miscarriage, her left arm and breast were amputated, her face was badly disfigured, and she lost her hearing. Over 60% of Cara's body was burned, and both of her legs were amputated below the knee. Following Cara's discharge from the hospital, she was placed in Children's Services foster home. She required the use of a jobst pressure garment, consisting of a face mask, body vest, and sleeves, which she needed to wear twenty-three hours a day. She also required a gel pad under the mask and had splints and prosthetic devices on both legs.

Cara was made a temporary court ward and placed in the home of her paternal grandmother. Her mother remained in the hospital. Her father was arrested following an investigation of the fire and was incarcerated. Cara was then placed in the home of her maternal great-aunt and uncle, where she is currently residing.

Since Cara's placement in foster care, she has had surgery to release the index and small finger on her right hand. Although she continues to wear the jobst garment, she now wears her gloves and mask primarily at night. Her face has healed well. She will continue to require occupational and physical therapy, ongoing medical care and additional surgery and skin grafting. Cara is self-conscious about the loss of her legs, and she has had problems adjusting to her mother's disfiguration.

PETE

Pete is six years old. For the first four years of his life, Pete lived with his mother and two younger brothers. He had no contact with his father. While living with his mother, Pete was not provided with proper clothing or helped to follow through with hygiene. The home was consistently unclean and lacked adequate furnishings or food. His mother had a history of physical and sexual abuse, used crack cocaine and other drugs, and was part of an escort service. The family moved frequently, often with little preparation.

When Pete was four, he and his brothers were removed from his mother's home and placed in separate foster care homes. Permanent custody was recently taken on the two younger children. After one previous placement, Pete was placed in the care of his aunt, who has been very nurturing and supportive of Pete. Pete's mother does not participate in assessments or visitations. Pete's aunt reports that she must constantly monitor him because he does not understand boundaries. He also has few self-care skills. Pete's aunt has been attempting to obtain respite care but has been unable to because of limited funds and resources.

ELLIE

Ellie is two years old. Ellie was removed from her mother's care at birth and placed in foster care. When she was eighteen months old, Ellie received a liver transplant. Two months ago she was placed with a second foster family, consisting of a mother, father, five biological children aged twelve to twenty-two and one adopted child, age ten. All of the biological children are currently living at home. The two eldest of these, aged twenty and twenty-two, neither work nor attend school. The adopted child was sexually abused prior to her arrival in the family two years ago.

The mother has assigned primary care of Ellie to the two oldest daughters. Recently, the oldest daughter reported to her mother that she had stopped giving Ellie her medication "because she seemed to do better without the medication." The foster mother concurred with her daughter, and has reported to the worker that she will not comply with giving Ellie her medication because Ellie's doctor has not made adjustments in the medication dosage as requested by the foster mother.

ROSA

Rosa's mother was the victim of domestic abuse. Rosa's mother was responsible for the childcare and housework. She and her husband argued a lot, and he belittled her mother. He also insisted on sex with her when she had health problems, isolated her from friends, controlled the finances, and would not let her get a job. When the verbal abuse began escalating into physical abuse, she went to a shelter.

Ten months after Rosa's mother left her father, Administrative Services (ADS) charged her with neglect for allowing Rosa to witness domestic abuse: "Respondents constantly argue in the presence of the child, with the child intervening and the child . . . states that she wished her parents would not argue as much." ADS placed Rosa in foster care while the neglect proceedings were pending. Charges were added that the mother "admits to being present when the father would beat the child" and that she "failed to protect the child from being beaten." Her mother says that she only told the social worker that she argued with her husband about spanking Rosa and that he refused to stop hitting the child.

JAMIE AND LINDA

Three year-old Jamie and his two year-old sister, Linda were placed in foster care because their mother's home was not a suitable environment. Social workers said that her home was unhealthy, dirty, and unsafe. Criticism ranged from having three pennies on the floor that the children could choke on to allowing dirt on the rubber seal of the refrigerator door. The report included that the children's bedroom was "cluttered with toys, clothing and other objects strewn about."

Social services first began monitoring the home of the twenty-year-old single mother when Linda's pediatrician reported that Linda was not gaining enough weight. The mother claimed that her daughter was like herself when she was young, smaller than most children. The social worker also stated in the report that there was not enough food in the home. The mother claimed that there was because four nights a week they ate dinner from Carl Junior's restaurant where the children's father worked. Eventually doctors diagnosed Linda's case as "failure to thrive."

The children's mother attempted to follow the social worker's instructions, but continued to clash with him. Their twenty-one-year-old father also requested custody but did not comply with the requests from social workers to become involved with his children. Thus, the children were removed.

Two weeks after the removal, their mother visited the children and noticed that Linda had a large bruise on the side of her face. A social worker told her that the foster mother said Linda fell on a concrete park bench when her brother tried to lift her from a grocery cart. Linda's mother thought the story was unbelievable because three-year-old Jamie is too short to lift Linda out of a cart. She also noted that Linda had no scratches from hitting the concrete.

Linda was taken to the hospital but details about the visit were unclear because the caregivers in the foster home speak mostly Spanish. Linda's mother requested a copy of the medical report, but never got one. Each week Linda's mother complained about a new bruise. She complained for months, but nothing happened. Social workers said that Linda had fallen or hit herself. Her pale skin seemed to bruise easily. Then during one visit with her mother, Linda's eye was nearly swollen shut and her bottom lip was cut. She had several bruises on her face and a red mark on the back of her head. When social workers removed Linda's clothes they found bruises all over her legs. Based on these findings, the children were removed to another foster home.

In the second home, the foster mother said Jamie was too hard to control. He pulled the sheets off the bed and did not follow directions. She wanted him out of her home immediately, but would keep Linda. Jamie was placed in another home a significant distance from where his mother and sister were. The siblings lived apart for two months, during which time they did not see their mother or each other. The children were reunited in a third foster home, and visits with their mother resumed.

During the shuffling, the children's mother worked to regain custody or to place the children with relatives. The sister of the children's father offered to take them, but plans fell through because she lived too far away. Now, the great-grandparents want them and are going through background checks.

JORDAN

It was 2:00 A.M. in a hospital emergency room. To the doctor on duty it was unimportant that the woman needing treatment for sickle cell anemia was an undocumented alien, or that her American-born child was healthy. The doctor admitted them both because the mother urgently needed hospital treatment, and the child had no one else who could take him.

The mother's poverty and recurrent hospitalizations had forced her several times to turn her son over to the foster care system. Getting Jordan back and providing adequately for his care at home was always difficult for her. This was because, as an illegal immigrant, she did not qualify for most government financial and medical assistance. In addition, the assistance her son was entitled to was often not enough, even with Supplemental Security Income for Jordan's learning disabilities and food stamps. Social workers said that the hospital was not a place for Jordan, and within days they charged that the mother's failure to make more suitable arrangements amounted to child neglect.

RANDY

At the age of two days, Randy, the baby of a sixteen-year-old unwed mother, was declared a dependent child and placed in the custody of the state. The child's mother was herself an incorrigible child who had been sexually promiscuous and who currently has no means of supporting herself or her child. The agency administering TANF refuses to approve aid for the child in the mother's home.

The mother wishes to care for her child in the home of her parents. Her father is eighty-years-old and disabled, her mother is seventy-two years old and has previously asked Children's Services to take the baby since she and her husband cannot help take care of it, physically or financially. Randy never lived with his mother prior to his removal.

DISCUSSION

As you read the points of the discussants that follow, please keep the facts of these case studies in mind and also consider the following questions:

1. What services would you recommend for this family?

2. What should be the ultimate case plan?

3. Should the state initiate a proceeding to terminate parental rights?

4. Is this case a good one for kinship care, open adoption, or some other form of care?

5. What additional information do you need to answer these questions?

In addition, the session participants tried to be realistic and mindful of the availability of information and the costs in time, money, and other resources in obtaining information desired for the best assessment of a child's needs.

PARTICIPANTS

The Participants in this session, acting as invited experts for our "mock board meeting," included Elizabeth Bartholet, Tonya Brito, Lynn Girton, Susan Vivian Mangold, and Richard Wexler. Elizabeth Bartholet's[19] discussion of the nature of the current manifestation of foster care, problems faced within the system, and the need to make more foster children available for adoption earlier is described in a comprehensive fashion in her book, *Nobody's Children: Abuse and Neglect, Foster Drift, and the Adoption Alternative*.[20] Tonya Brito[21] focused primarily on progressive alternatives to the system of foster care, such as open adoptions. Lynn Girton[22] focused on her experience representing foster parents, the reasons for the foster placements and the difficulties encountered by the foster parents, both in handling the children's needs and in attempting to adopt these children. Susan Mangold[23] discussed two separate issues: the particular problems associated with the use of foster care in cases of

19. Professor of Law, Harvard Law School.

20. ELIZABETH BARTHOLET, NOBODY'S CHILDREN: ABUSE AND NEGLECT, FOSTER DRIFT, AND THE ADOPTION ALTERNATIVE (1999).

21. Assistant Professor of Law, University of Wisconsin-Madison Law School; A.B. Barnard College, 1986; J.D. Harvard Law School, 1989.

22. Lynn Girton is Chief Counsel of the Volunteer Lawyers Project of the Boston Bar Association. She was formerly the Senior Managing Attorney of the Employment and Welfare unit of Greater Boston Legal Services. Her current legal work has focused primarily on probate and juvenile issues, with an emphasis on the legal needs of low income children.

23. Associate Professor of Law, University at Buffalo Law School; A.B. Harvard University, 1982; J.D. Harvard Law School, 1987.

spousal abuse and the specific needs of children who "age out" of the foster care system without permanent placement.[24] Finally, Richard Wexler[25] addressed the problems associated with the overuse of foster care. For example, when existing families are unnecessarily torn apart, usually state costs are greater than that for the services necessary to assist those families.[26]

FINAL REFLECTIONS

The discussions of the session participants in many ways captured the history of foster care at one moment in time and the respective policy approaches. We discussed the issues confronting foster care at many points on the spectrum, from problems originating with the biological family to those originating with the foster families. We also discussed administrative and resource problems of the agencies and administrators to the emotional/psychological/physical needs of the children subject to the foster system. The range of issues that challenge the foster care system obviates why there have been, and continue to be, differing approaches in policy choices. The most amazing thing about this discussion was that it evolved from the two stated policy approaches clashing, to a recognition that each position may be, both together and independently, recreating and perpetuating the problems they purport to address.[27] Once this was acknowledged, it was possible to examine other approaches that either incorporated both positions or ignored both of them in seeking solutions that were workable for the individual child involved.

The information and ideas discussed are only the beginning of the task. Hopefully we can now move forward to find broad solutions, an appropriate policy to assist the child welfare system, and individualized solutions for each child who becomes subject to the system.

---

24. *See* Susan Mangold, *Transgressing Borders Between Protection and Empowerment for Domestic Violence Victims and Older Children: Empowerment as Protection in the Foster Care System*, 36 NEW ENG. L. REV. 69 (2001).

25. Executive Director, National Coalition for Child Protection Reform; author of *Wounded Innocents: The Real Victims of the War Against Child Abuse* (Prometheus Books: 1990, 1995).

26. *See* Richard Wexler, *Take the Child and Run: Tales From the Age of ASFA*, 36 NEW ENG. L. REV. 129 (2001).

27. Deliberately, I state this situation to parallel Professor Frug's description of various feminist scholars in their approaches to difference and equality, culminating in a description of postmodernists who challenge such approaches, "which themselves are likely to recreate and perpetuate the problem of the relationship between sexual difference and equality." Mary Joe Frug, *Sexual Equality and Sexual Difference in American Law*, 26 NEW ENG. L. REV. 665, 665 (1992).

Transgressing the Border Between Protection and Empowerment for Domestic Violence Victims and Older Children: Empowerment as Protection in the Foster Care System

Susan Vivian Mangold*

INTRODUCTION

The conference title, Transgressing Borders,[1] poses interesting interpretations when applied to the field of foster care. I frame my comments in the debate over what have been identified as the competing rights of protection and empowerment. I would like to present empowerment of caretakers and protection and empowerment of children as complimentary goals in the foster care system. This means family preservation services sensitive to the complicated dynamics in homes with domestic violence. For older teens, empowerment by protection requires permanent placement options that allow children and their birth families more latitude to plan in cooperation with child welfare agencies. Both require new and real investment in families and caretakers. While this may not be fiscally realistic, it is legally possible and morally necessary.

From colonial times through twentieth-century Supreme Court action and more recent federal legislation, the tension between parental empowerment/autonomy and child protection has been in a shifting balance. In earlier times, parental authority could not be questioned for the purpose of protecting children from abuse. However, by the late nineteenth century, children could be protected by removal via criminal proceedings or philanthropic agency actions. More recently, in the 1980's, family empowerment and protection of children were seen as complimentary, not competing--the best way to protect children was not by removal but through services which

* Associate Professor of Law, University of Buffalo Law School; A.B. Harvard University, 1982; J.D. Harvard Law School, 1987.

1. Symposium, *Transgressing Borders: Women's Bodies, Identities and Families*, Mar. 31, 2001, New England School of Law, Boston, Massachusetts.

empowered parents to safely care for their children. In the last few years, federal law has again shifted the balance toward protection via removal from the home and swifter termination of parental rights. Amidst this shifting debate, new factors constantly change understandings of the complexities of family privacy and parental empowerment and the role of the state in protecting children.

Foster care is defined for this article as out-of-home care agreed to by parents or ordered by a court following a report of abuse or neglect filed with a state child welfare agency.[2] It is provided by state or local public agencies directly or by contract with private providers.[3] Federal funds reimburse the public agencies creating a federal-state matching system, which carries federal mandates in exchange for the matching funds.[4] Foster care is therefore a locally provided placement service with some federal regulation as a result of federal mandates. Foster care can be used in other public systems such as the mental health system or the juvenile justice system to house children who cannot be cared for in their own homes.

Foster care is sometimes understood as placement in home-like settings as opposed to group homes or institutions. I am using the term more broadly to refer to out-of-home placement of any type for abused and neglected children. This could include institutional settings, group homes, foster family care or supervised independent living arrangements.

Borders separate many classifications with legal significance in the area of foster care. Most obviously, biological parents as a category have different rights and responsibilities expected of them and owed to them than foster parents or adoptive parents; children in care similarly have specified rights which are not provided to children cared for in their own homes. Because of the distinguished mix of presenters on the conference foster care panel who debate the categories of biological parent, foster parent and

---

2. For a more complete discussion of the foster care system and the child protection system which triggers out-of-home care based on abuse or neglect, see Susan Vivian Mangold, *Challenging the Parent-Child-State Triangle in Public Family Law: The Importance of Private Providers in the Dependency System*, 47 BUFF. L. REV. 1397, 1438-42 (1999).

3 These are usually non-profit private providers, but foster care by for-profit providers is eligible for federal matching funds. *See* 42 U.S.C. § 672(c) (Supp. V 1999). States may contract with over 200 private provider agencies to provide foster care to abused and neglected children. *See* Susan Vivian Mangold, *Protection, Privatization, and Profit in the Foster Care System*, 60 OHIO ST. L.J. 1295, 1313 n.81 (1999).

4. *See, e.g.*, Child Abuse Prevention and Treatment Act, 42 U.S.C. §§ 5101-5106a (Supp. V 1999); Adoption Assistance and Child Welfare Act of 1980, Pub. L. No. 96-272, 94 Stat. 500 (codified as amended in scattered sections of 42 U.S.C.); Adoption and Safe Families Act of 1997, 42 U.S.C. § 671 (1994 & Supp. 1999).

adoptive parent, I focus on the eligibility borders into and out of the foster care system and the implications for protection and empowerment of transgressing those borders. Foster care remains the only open-ended entitlement program,[5] and as such, it is crucial to carefully consider how the law defines who may come in, who may go out, who must come in and who must go out.

In Part I, I draw upon some of my previous writings to summarize the history, legislation, current operation and representational standards of the legal system that responds to child abuse and neglect.[6] I include this background to depict the tensions between protection and empowerment in history, policy and practice. These tensions are balanced within a context of gender subordination, racism and classism. While the specifics of the context have changed over time, these three types of discrimination are still omnipresent and operative in the foster care system.

I discuss the right to protection as developed by constitutional case law and federal legislation. I use the historical background to highlight the conflict which developed between protection of children and empowerment of parents and also to distinguish the origins of the child protection system from those of the domestic violence system which has its own historical development. I draw upon the scholarly and practical debate regarding the proper model of representation of children to illustrate the tension between empowerment and protection, especially for older children.

In Part II, I turn to the extension of the definitions of child abuse or neglect to include domestic violence. "Child abuse and neglect" is a phrase which encompasses a spectrum of child maltreatment from actual serious injury to possible risk of harm perpetrated by caretakers upon their children. As discussed infra Part I-C and I-D, each state has its own laws defining child abuse and neglect within the parameters of federal mandates.[7] Domestic violence similarly encompasses a spectrum of abuse. It is abuse between adult intimate partners perpetrated to maintain power and control.[8]

As the comorbidity of domestic violence and child abuse and neglect is

---

5. For a discussion of the connection between public assistance and foster care entitlements, see Mangold, *supra* note 3, at 1310.

6. *See* Susan Vivian Mangold, *Extending Non-Exclusive Parenting and the Right to Protection for Older Foster Children: Creating Third Options in Permanency Planning*, 48 BUFF. L. REV. 835 (2000); Mangold, *supra* note 3, at 1295; Mangold, *supra* note 2, at 1397.

7. *See, e.g.*, 42 U.S.C. § 5106a(b)(2)(A) (1994 & Supp. 1998).

8. The Power and Control Wheel, which depicts the cycle of power and control in abusive adult relationships, was developed by the Domestic Abuse Intervention Project, 206 W. Fourth St., Duluth, MN, 55806. *See* Domestic Abuse Intervention Project's Power and Control Wheel, *reprinted in* NANCY K. D. LEMON, DOMESTIC VIOLENCE L. 43 (2001).

documented and its effects understood, more attention is being placed on coordinating the provision of services dealing with domestic violence and child protection. This coordination has resulted in identification of new forms of child abuse and neglect based upon the existence of domestic violence in the home. Domestic violence is a newly recognized form of child abuse and neglect that creates eligibility for foster care. The issues posed in recognizing domestic violence as a form of abuse are captured in part, in the case study of "Rosa" by Zanita Fenton, Chair of the conference's foster care panel.[9] That case study will be referred to in this article when illustrative and reads as follows:

Rosa

> Rosa's mother was the victim of domestic abuse. [Her husband left all the child care and housework to her, they argued a lot, he belittled her mother, he insisted on sex with her even when she had health problems, he isolated her from friends, he controlled the finances and would not let her get a job. When she feared that the verbal abuse was escalating into physical abuse, she decided to leave, going to a shelter.]
>
> Ten months after Rosa's mother left her father, Administrative Services (ADS) charged her with neglect for allowing Rosa to witness domestic abuse: "Respondents constantly argue in the presence of the child, with the child intervening and the child . . . states that she wished her parents would not argue as much." ADS placed Rosa in foster care while the neglect proceedings were pending.
>
> Charges were added that the mother "admits to being present when the father would beat the child" and that she "failed to protect the child from being beaten." Her mother says that she only told the social worker that she argued with her husband about spanking Rosa and that she refused to do so.

In the Rosa case study, ADS charges Rosa's mother with neglect for allowing Rosa to witness verbal domestic violence of her mother by her father. "Witnessing domestic violence" is a newly recognized form of child abuse. The charges against Rosa's mother later include "failure to protect" for abuse allegedly perpetrated by the father against the child and known to

---

9. *See* Zanita E. Fenton, *Foster Care: The Border of Family Identity-Maintaining, (Re)creating, Destroying*, 36 NEW ENG. L. REV. 59 (2001).

the mother. "Failure to protect" is likewise a new form of child abuse or neglect based upon the presence of domestic violence in the home.

The presence of domestic violence in the home can now create eligibility for foster care. Violence between adults in the home is now a type of child abuse or neglect, which can trigger the legal responses leading to placement outside the home. The fact that the presence of domestic violence may allow children to come into temporary foster care for protection has advantages and disadvantages which are discussed; the fact that exposure to domestic violence may force some children into foster care against their wishes and those of their mothers, who may be the ones revealing the violence, has created a more complicated discussion implicating the advantages and disadvantages of the core idea of coordination of the systems. To what extent should protection of children trump empowerment of mothers? Can we shift the debate to consider empowerment of mothers and children as the best permanent and temporary way to protect children?

The history of the legal responses to domestic violence is explained as necessary background to the convergence and divergence of the domestic violence and child protective systems in theory and practice. The goal of the domestic violence system is to empower women; the child protective system too often pits mother caretakers against the agency in a purported effort to protect children. Current statistics and socio-legal responses to domestic violence are described to lay the groundwork to consider the efforts at coordination of the two systems and the consequences of that coordination.

I conclude this section with suggested reforms to avoid the unintended results of collaboration and to begin to imagine a system that truly invests in the empowerment of caretakers in an effort to protect children.

Part III focuses on age as a criteria of eligibility in the foster care system. While domestic violence can bring a child in, age can force a child out, objectively ending a child's right to protection. I discuss age as an absolute limit to the continuation of foster care services, creating ineligibility when a birthday is crossed. In 1999, Congress estimated that approximately 20,000 teens exited foster care because they had aged-out of the system, reaching the birthday when they were no longer eligible for foster care benefits.[10] Significantly, none of the nine children in the eight case studies used by the panel at the conference were about teens in foster care. Older foster children are never the paradigm and are often forgotten in the system. The eldest child discussed in the case studies was eleven-year-old Cara, and her case

10. *See* H.R. 3443, 106th Cong. (1999). As a result of the Foster Care Independence Act, discussed *infra* Part III, states may now continue some benefits to age twenty-one and receive federal matching funds for them. In some cases, the age of ineligibility is still eighteen, in others it is twenty-one. *See* Mangold, *supra* note 6, at 835.

raised issues of disabilities more than age. We are not told the age of Rosa. Yet, for children approaching ineligibility for foster care, age can be the most relevant factor. Statistics on older children in foster care are presented in conjunction with recent federal legislation raising the age for discretionary eligibility. As in Part II, I conclude this section with suggestions for reform to mitigate the harshness of age ineligibility and to imagine a more flexible spectrum from protection to empowerment.

PART I. THE RIGHT TO EMPOWERMENT AND THE RIGHT TO PROTECTION IN THE FOSTER CARE SYSTEM

A. Theoretical Framework of the Foster Care System

While parents have a right to raise their children free from state intervention, children have a countervailing right to protection from abuse and neglect. This tension between parental rights and child protection is the key conflict in the child protection system, but I am interested here in exploring the unique tensions raised by domestic violence and age ineligibility for older children in foster care. If the allegations of abuse and neglect are severe enough, federal and state laws aimed at protecting children require that they be removed from their parents and placed in foster care.[11] Foster care is an entitlement, and every state is required by state and federal law to provide foster care for eligible children.[12] Foster care provides out-of-home care for over 500,000 children in the U.S. every day.[13] Most of these children are in foster care as a result of allegations of abuse and neglect against their parents.[14] A sense of the range of allegations is presented by Professor Fenton's case studies.

In her now famous and often misunderstood article on children's rights, Hillary Rodham Clinton wrote in 1973 that "children's rights" was a "slogan in need of a definition."[15] Twenty-seven years later, the notion of chil-

---

11. *See, e.g.*, Child Abuse Prevention and Treatment Act, 42 U.S.C. §§ 5101-5106 (Supp. 2000); Adoption Assistance and Child Welfare Act of 1980, Pub. L. No. 96-272, 94 Stat. 500 (1980) (codified as amended in scattered sections of 42 U.S.C. (Supp. 1997)); Adoption and Safe Families Act of 1997, 42 U.S.C. § 671 (1994 & Supp. 1998).

12. *See* 42 U.S.C. § 622 (1994 & Supp. V 2000).

13. *See* CHILD. DEF. FUND, STATE OF AMERICA'S CHILDREN: 2000 YEARBOOK 84, 86 (2000). *See also* STAFF OF H.R. COMMITTEE ON WAYS AND MEANS, 105TH CONG., 1998 GREEN BOOK: BACKGROUND MATERIAL AND DATA ON PROGRAMS WITHIN THE JURISDICTION OF THE COMMITTEE ON WAYS AND MEANS 783-90 (Comm. Print 1998) (noting statistics for 1995 and earlier) [hereinafter 1998 GREEN BOOK].

14. *See* 1998 GREEN BOOK, *supra* note 13, at 790 (noting statistics for 1995 and earlier).

15. Hillary Rodham Clinton, *Children Under the Law*, 43 HARV. EDUC. REV.

dren's rights is still not well defined. Perhaps it is more accurate to say that there is no singularly accepted definition or theory of the rights held by children. The rights claimed tend to be of two general types: those advocating for children as autonomous persons under the law[16] and those placing a claim on society for protection from harms perpetrated on children because of their dependency.[17] I label the first type as the right of empowerment and the second as the right to protection.

The rights of empowerment for children were advocated in the manner in which many rights connected with oppressed minorities were pressed in the 1960's. Some children's advocates explicitly cited children as the next group after blacks and women who were entitled to a revolutionary expansion of rights.[18] Empowerment rights were necessary to recognize children's equality with adults. They would give children such privileges of citizenry as the right to vote, work and contract.[19] While children may always hold empowerment rights, those rights cannot be exercised until the child reaches a certain level of maturity.[20] While the age at which such an exercise occurs may vary, the key is that the right of empowerment evolves for each child from inability to exercise to ability to exercise.

Independent living initiatives[21] for teens in foster care seek to enhance the empowerment rights of children. Unfortunately, the goal of empowerment is unconnected to any objective assessment, which could continue protective services. Regardless of a child's maturity and resources, and regardless of their success in independent living programs, all children are liberated from

---

487, 487 (1973).

16. *See, e.g.*, JOHN HOLT, ESCAPE FROM CHILDHOOD (1974), HENRY H. FOSTER, JR., A "BILL OF RIGHTS" FOR CHILDREN (1977).

17. *See, e.g.*, Bruce Hafen, *Children's Liberation and the New Egalitarianism: Some Reservations About Abandoning Children to Their 'Rights'*, 1976 BYU L. REV. 605 (1976).

18. *See* Martha Minow, *What Ever Happened to Children's Rights?*, 80 MINN. L. REV. 267, 270 (1995). Minow discusses works such as John Holt's ESCAPE FROM CHILDHOOD (1974) and Richard Farson's BIRTHRIGHTS (1974) to explain this movement in the 1960's and 70's, but whose origins were derived much earlier in the writings of Jean-Jacques Rousseau and John Dewey.

19. *See, e.g.*, Minow, *supra* note 18, at 271; *see also, e.g.*, RICHARD FARSON, BIRTHRIGHTS (1974).

20. *See* Emily Buss, *Confronting Developmental Barriers to the Empowerment of Child Clients*, 84 CORNELL L. REV. 895 (1999) (arguing that developmental issues are important in considering empowerment rights of children). *But see* Katherine Hunt Federle, *The Ethics of Empowerment: Rethinking the Role of Lawyers in Interviewing and Counseling the Child Client*, 64 FORDHAM L. REV. 1655 (1996) (arguing that empowerment is an important consideration for all children in determining the model of attorney representation of children).

21. *See* Independent Living Initiative, Pub. L. No. 99-272, 100 Stat. 294 (codified as amended at 42 U.S.C. § 677).

foster care when they reach the age of majority.

Unlike the rights connected with empowerment, protection based rights evolve in the opposite pattern. The right is strongest at the front end of the foster care system after cases are investigated and initial placements are developed. The very young or incompetent have the strongest claim on these rights, but the ability to exercise them diminishes with maturity. Our foster care system has been developed to give children a right to protection until the age of eighteen or, in some cases, twenty-one. At the age of eighteen or twenty-one, the right to protection is superceded by a right of empowerment as children mature out of the system whether they are ready for independence or not.[22] Foster care ends when children reach the age of majority, eighteen or twenty-one, depending on state law and regulations.[23] In 1999, Congress estimated that 20,000 teens exited foster care because they "aged-out" of the system, reaching the age at which eligibility for foster care benefits is terminated.[24]

There is tension between the right to protection of abused and neglected children and the empowerment rights of older foster children. How much authority should be ceded to older children to determine their services and placement arrangements? How do we justify terminating needy children from foster care solely because of their age? Similarly, there is tension within the domestic violence system and between the domestic violence and child protection systems over the indeterminant border between protection and empowerment. Is the goal of the domestic violence system to protect or empower in the short and long term? How does and should the children's protective services system interpret its role to protect children while empowering caretakers? These core questions regarding protection and empowerment frame my discussion of age and domestic violence as determinants of eligibility for foster care. Transgressing the border between empowerment and protection to understand empowerment of older children and of caretakers as the heart of protection frame my proposed reforms.

---

22. *See* Gerald P. Mallon, *After Care, Then Where? Outcomes of an Independent Living Program*, 77 CHILD WELFARE 61, 62 (1998) (stating at the outset that "[P]reparing young people in out-of-home care for independent living and for successful adulthood has not been one of the child welfare system's primary goals.").

23. *See* WASH. REV. CODE § 13.34.030 (1) (2000) (stating that a child is "any individual under the age of eighteen years"); HAW. REV. STAT. § 346-17.4 (a) (1) (1999) (stating that a child is eligible for foster care if "[t]he person is twenty-one years old or younger").

24. *See* H.R. 3443, 106th Cong., (1999). As a result of the Foster Care Independence Act, discussed *infra*, states may now continue some benefits to age twenty-one and receive federal reimbursement for those foster care benefits.

B. Supreme Court Cases Developing the Boundaries of the Right to Protection[25]

The right to protection frames the entire child welfare system and much of the state law governing the state systems, but is not clearly recognized under constitutional law.[26] In 1944, the Supreme Court in *Prince v. Massachusetts*[27] recognized the authority of the state to protect children even when this infringed upon the legal caretaker's authority. This case was preceded by *Meyer v. Nebraska* and *Pierce v. Society of Sisters*[28] which voided state laws limiting parental authority in the choice of schools.

In *Meyer v. Nebraska* and *Pierce v. Society of Sisters*, state laws infringing upon the authority of parents were deemed unconstitutional. The *Pierce* court recognized the rights of parents, their duty to their children and the right to exercise these rights and duties free of state intervention.

> Under the doctrine of *Meyer v. Nebraska*, 262 U.S. 390, we think it entirely plain that the Act of 1922 [mandating that children attend public schools] unreasonably interferes with the liberty of parents and guardians to direct the upbringing and education of children under their control . . . . The child is not the mere creature of the State; those who nurture him and direct his destiny have the right, coupled with the high duty, to recognize and prepare him for additional obligations. [29]

This oft-quoted phrase recognizes the rights and duties of parents and the possible role of the state on behalf of children. The parent-child-state triangular balance is introduced constitutionally.[30] The court in *Prince v. Massachusetts* built upon the foundation laid by *Meyer* and *Pierce* but decided that the circumstances of the case warranted state intervention to protect the child. In *Prince*, *parens patriae* power of the state to protect children was upheld. The *Prince* court held valid a state child labor law against both the legal guardian's assertion that the law violated her right to raise the child as she saw fit and the child's right to practice Jehovah Witness beliefs by selling religious magazines. In discussing the applicable precedents, the court

---

25. Discussion of these cases summarizes previous writings. *See* Mangold, *supra* note 9, at 844-49; Mangold, *supra* note 2, at 1402-10.

26. *See* DeShaney v. Winnebago County Dept. of Social Services, 489 U.S. 189 (1989) (deciding that the state does not owe a duty to protect a child from violence at least until the child is in state custody); *infra* text accompanying notes 35, 39-42.

27. 321 U.S. 158 (1944).

28. Meyer v. Nebraska, 262 U.S. 390 (1923); Pierce v. Society of Sisters, 268 U.S. 510 (1925).

29. *Pierce*, 268 U.S. at 534-35.

30. For a discussion of the balance between parental rights and the state's authority to intervene into the parent-child relationship, see Mangold, *supra* note 2, at 1438-42.

limited the holdings of *Meyer* and *Pierce* and recognized the state's ability to exercise its authority to protect a child's welfare:

> Previously in *Pierce v. Society of Sisters*, 268 U.S. 510, this Court had sustained the parent's authority to provide religious with secular schooling, and the child's right to receive it, as against the state's requirement of attendance at public schools. And in *Meyer v. Nebraska*, 262 U.S. 390, children's rights to receive teaching in languages other than the nation's common tongue were guarded against the state's encroachment. It is cardinal with us that the custody, care and nurture of the child reside first in the parents, whose primary function and freedom include preparation for obligations the state can neither supply nor hinder. *Pierce v. Society of Sisters*, *supra*. And it is in recognition of this that these decisions have respected the private realm of family life that the state cannot enter . . . . It is sufficient to show what indeed appellant hardly disputes, that the state has a wide range of power for limiting parental freedom and authority in things affecting the child's welfare . . . .[31]

This framework was not constitutionally developed by the Supreme Court until the 1972 case of *Wisconsin v. Yoder*[32] reinforced the authority of parents. That case resulted in a successful challenge to compulsory education laws imposed on the Amish. In *Yoder*, parents were convicted under a Wisconsin law requiring a child's attendance at school until the age of sixteen. The parents argued that sending their teens to school past the eighth grade violated their Amish beliefs and lifestyle. The court agreed with the parents against the state's authority to protect the children in this case.[33]

Meyer, *Pierce*, *Prince* and *Yoder* all dealt with state intervention to protect children but were not public family law[34] cases dealing with children in state custody as a result of allegations of child abuse or neglect. In *Santosky v. Kramer*,[35] and then in *DeShaney v. Winnebago County Department of Social Services*,[36] parental rights and duties and state rights and duties to-

---

31. Prince v. Massachusetts, 321 U.S. 158, 166-67 (1943).

32. 406 U.S. 205 (1972).

33. The rights of children were not explicitly considered by the majority, but they were raised by Justice Douglas in his dissent. Douglas argued for a remand to consider the wishes of the children whose parents were convicted under the compulsory education law. The weight to be given to their wishes was not stated. Seeking such input from children could be viewed as granting them empowerment rights over their parents' control, but Douglas did not detail the authority to be given to their wishes, nor did he speak in terms of children's rights in his dissent. *See Yoder*, 406 U.S. at 241 (Douglas, J., dissenting).

34. *See* Mangold, *supra* note 2, at 1397. "I coin the phrase *public family law* to depict cases where the state has intervened into the 'private' family to assume some custodial interest from the parents." *Id.*

35. 455 U.S. 745 (1982).

36. 489 U.S. 189 (1989).

ward abused and neglected children were addressed by the Supreme Court. In *Santosky*, the court held that the standard necessary to involuntarily terminate parental rights was "clear and convincing evidence."[37] Even when children were in the dependency system and their care was subject to procedural safeguards at each juncture, the court found that the importance of the parental right to the care and control of their child could not be severed absent a showing by the state of clear and convincing evidence of unfitness.[38] The *Santosky* court relied on a line of cases, beginning with *Meyer*, *Pierce*, and *Prince*, to demonstrate historical recognition of parental rights.

> [F]reedom of personal choice in matters of family life is a fundamental liberty interest protected by the Fourteenth Amendment. The fundamental liberty interest of natural parents in the care, custody, and management of their child does not evaporate simply because they have not been model parents or have lost temporary custody of their child to the State. Even when blood relations are strained, parents retain a vital interest in preventing the irretrievable destruction of their family life.[39]

In *DeShaney*,[40] the rights of parents to raise their children free of state intervention was upheld over a child's right to protection in a case in which the facts were deeply challenging to the weight given to parental rights. The court declined to find a state duty to protect a child who was in the custody of his father, not in state custody, when the child suffered permanent serious injury at the hands of his father.[41] Winnebago County Department of Social Services was repeatedly informed of incidents of abuse and the risk of further abuse, but the agency did not remove the young child from his father's care.[42] The court reasoned that the state right to intervene, investigate, and monitor the situation did not implicate a duty to protect the child who remained in his father's care.[43] In accordance with the parent-child-state framework developed in the *Meyer-Pierce-Prince* line, the state had not taken on the custodial right and therefore did not hold the accompanying duty to protect the child. The right of control had been left exclusively to the father, and the child could not make out a liberty claim for denial of a duty to protect based on the father's acts of "private violence."

These court cases discuss the *parens patriae* power of the state to protect

---

37. *Santosky*, 455 U.S. at 769.

38. *See id.*

39. *Id.* at 753.

40. 489 U.S. 189 (1989).

41. For a provocative discussion of the Thirteenth Amendment as the more appropriate cause of action in this case, see Akhil Reed Amar & Daniel Widawsky, *Child Abuse as Slavery: A Thirteenth Amendment Response to DeShaney*, 105 HARV. L. REV. 1359 (1992).

42. *See DeShaney*, 489 U.S. at 191.

43. *See id.* at 197.

children but leave blurred where the border of the right to protection exists amidst the parents' fundamental right to raise their children. None of these cases discuss children's empowerment rights or the importance of the age and maturity of the children. Legal doctrine mediating the tension between protection and empowerment is undeveloped. The boundary for eligibility for protection is therefore left to myriad state laws within the framework of federal mandates.

C. History of the Balance Between Protection and Empowerment in the Foster Care System[44]

1. Early Protection of Children

Since the earliest colonial days, some children have needed protection beyond that which families could provide. Orphans are the obvious children in need of protection.[45] In the seventeenth and eighteenth centuries, children whose parents did not provide adequate religious or other instruction could also be removed to insure that the children would grow to be productive adults.[46] Children had no recognized "right" to protection, and physical punishment was permitted and even required,[47] but religion and orderly

---

44. For a fuller history of child protection before federal legislation to address child abuse and neglect, see Mangold, *supra* note 2, at 1397 (summarized in part in this section).

45. Mary Ann Mason examined the record of two Virginia parishes to provide demographic information on involuntary apprentices. "Orphans constituted 38.1% of all child apprentices; 39.3% were classified poor children; 11% were described as illegitimate; and 12.6% were termed mulatto." Mary Ann Mason, *Masters and Servants: The American Colonial Model of Child Custody and Control*, 2 INT'L J. CHILDREN'S RTS. 317, 326 (1994).

46. *See* CHILDREN AND YOUTH IN AMERICA: A DOCUMENTARY HISTORY, Vol. I, 1600-1865, 27-29 (Robert H. Bremner ed., 1970) [hereinafter CHILDREN AND YOUTH, Vol. I].

47. The earliest colonial laws included provisions prohibiting excessive corporal punishment of children. *See* The Body of Liberties of 1641: The Liberties of the Massachusetts Collonie in New England, 1641, *reprinted in* EDWIN POWERS, CRIME AND PUNISHMENT IN EARLY MASSACHUSETTS: 1620-1692, 543 app. (1966). The Body of Liberties addressed the issue of the physical punishment of children by their parents in Chapter 83 stating: "If any parents shall wilfullie and unreasonably deny any childe timely or convenient marriage, or shall exercise any unnaturall severitie towards them, such childeren shall have free libertie to complaine to Authoritie for redresse." WILLIAM H. WHITMORE, THE COLONIAL LAWS OF MASSACHUSETTS: REPRINTED FROM THE EDITION OF 1672 (1890) (from The Body of Liberties of 1641) (Boston, Rockwell and Churchill 1890). For a discussion of the English stockholder interests prompting codification of colonial law see ELIZABETH PLECK, DOMESTIC TYRANNY: THE MAKING OF AMERICAN SOCIAL POLICY AGAINST FAMILY VIOLENCE FROM COLONIAL TIMES TO

societal norms prompted early settlers to protect the well-being of children even against the prerogatives of their fathers.

Colonial fathers were charged with the proper upbringing of their children, responsible for educating and training them to be productive citizens of the community. Fathers who failed to properly instruct their children could lose custody of the children.[48] As early as the 1640's, the colonial laws authorized public authorities to remove children from their families and place them with other families who could raise them in a manner deemed appropriate.[49] Colonial laws allowed private intervention into the parent-child relationship to assure that child rearing was appropriate for raising employable and moral children. Tightly woven religious communities provided moral guidance and often acted with public authorities to provide supervision of family life.[50]

> Forasmuch as the good education of children is of singular behoof and benefit to any Common-wealth; and whereas many parents & masters are too indulgent and negligent of their duty in that kinde. It is therefore ordered that the Select men of everie town, in the severall precincts and quarters where they dwell, shall have a vigilant eye over their brethren & neighbours, to fee, first that none of them shall suffer so much barbarism in any of their families as not to indeavour to teach by themselves or others, their children & apprentices, so much learning as may inable them perfectly to read the english tongue, & knowledge of the Capital lawes . . . . Also that all masters of families doe once a week (at the least) catechize their children and servants in the grounds & principles of Religion . . . . And further that all parents and masters do breed & bring up their children & apprentices in some honest lawful calling, labour or imployment, either in husbandry, or some other trade profitable for themselves, and the Common-wealth if they will not or cannot train them up in learning to fit them for higher imployments.[51]

THE PRESENT 21-22 (1987). There is, however, no evidence of colonial fathers being punished for excessive corporal punishment alone. *See* Mason, *supra* note 45; PLECK, *supra*. In fact, corporal punishment of children was accepted and encouraged in colonial times. The 1674 Records of the Suffolk County Court record two instances, one involving Governor Leveret's grandson, where parents were ordered to whip their children at home in the presence of the constable as punishment for misbehavior by the children. *See* POWERS, *supra* at 178. Even capital punishment for incorrigibility was codified. The General Laws of Massachusetts Colony, 1658, state that a son who is "stubborn and rebellious and will not obey [his parents'] voice and chastisement, but lives in sundry and notorious crimes, such a son shall be put to death." THE GENERAL LAWS OF THE MASSACHUSETTS COLONY 15 (1658) (Capital Laws, ch.13).

48. *See* PLECK, *supra* note 47, at 27-29.

49. *See id.*

50. *See generally* PLECK, *supra* note 47.

51. THE BOOK OF THE GENERAL LAUUES AND LIBERTYES CONCERNING THE INHABITANTS OF THE MASSACHUSETTS (1648) *reprinted in*, THE LAWS AND LIBERTIES OF MASSACHUSETTS 11 (1929).

In 1642, Massachusetts Bay enacted a law, to be enforced through the courts, that children could be removed from their parents' home involuntarily, based upon the manner in which parents were raising them. The goal was not to protect children from abuse but rather to protect society from poorly raised youth.

> And if any of the Select men after admonition by them given to such masters of families shal finde them still negligent of their dutie in the particulars aforementioned, whereby children and servants become rude, stubborn & unruly; the said Select men with the help of two Magistrates, or the next County court for that Shire, shall take such children or apprentices from them & place them with some masters for years (boyes till they come to twenty one, and girls eighteen years of age compleat) which will more strictly look unto, and force them to submit unto government according to the rules of this order, if by fair means and former instructions they will not be drawn unto it.[52]

Such children were removed by the town authorities and placed in an apprenticeship or indenture, called "binding out."[53] These indentures, or contracts to bind out children, could be arranged voluntarily by parents seeking training for their children or involuntarily by authorities who removed children from parents whose child rearing was seen as inadequate.[54] Involuntary indentures required local authorities to contract with families to care for the removed children.[55] These contracts were the first contracting to care for children whose families were considered unable to properly raise them to protect the community from the negative impact of these improperly raised children. Throughout the eighteenth century, public authorities acted to protect vulnerable children, such as orphans, emancipated slaves, unaccompanied laborers and contract laborers by binding them out to masters who would provide acceptable supervision and training.

Because the family unit was considered a form of governance and social control at the time, such involuntary indentures were not really public-private contracting as we understand it today. These arrangements predated a uniform child protection system, but they introduced interventions by the state as *parens patriae* that would later be assumed by nineteenth century criminal law prosecutions and anti-cruelty agency interventions. Historical beliefs of appropriate child rearing triggered these early interventions. Such early colonial arrangements are evidence of substitute care facilitated by local authorities; the same is true with modern foster care.

From the earliest colonial days through the antebellum years, binding out

---

52. *Id.*
53. *See id.*
54. *See* CHILDREN AND YOUTH, Vol. I, *supra* note 46, at 64.
55. *See id.*

remained the preferred way to deal with children of the poor. Although the first orphanage was established in 1728 in New Orleans, specialized institutional care for children was scarce until the nineteenth century.[56] Until that time, children of the poor who could not be bound out to private families were placed in almshouses, publicly operated warehouses for the poor not segregated by age.[57]

2. Protection of Children from Abuse and Neglect

The middle of the nineteenth century brought cases of criminal prosecutions against parents for beating their children. These prosecutions introduced an era when the legal system began to intervene in family life and compromise parental autonomy to protect children from physical assaults at the hands of their parents. At this nascent stage of development of a legal response to child abuse, many children considered "poor" or "neglected" were already under public supervision. Now, the states began to prosecute cases of physical assaults by parents on their children and to refer some of those children to the community resources available for neglected children.

The early cases prosecuting parents for physical assaults illustrate the historical hesitancy of the court to infringe on parental authority to protect children on the basis of allegations of physical beatings by fathers. Such

---

56. *See id.* at 60-63.

57. In February 1775, of the 622 paupers on the books of the New York City Almshouse, 259 were children, mostly under nine years of age. The authorities made every effort to bind out even these young children. In 1788, laws were passed allowing children in New York City, Albany and Hudson to be bound out without parental consent. The mayor, recorder, or aldermen could approve such an arrangement for any child found begging in the streets. By 1795, forty percent of the "inmates" at the almshouse in New York City were children under nine years of age. In 1797, following a yellow fever epidemic that filled the New York City almshouse with widows and their children, a group of women founded the Ladies Society for the Relief of Poor Widows with Small Children. The group was incorporated by law in 1802. In-kind help was given to assist widows in making a livelihood, but no relief was granted to women who refused to place out their children who were able to work. By 1800, the Society helped 152 widows with 420 children and in 1803 received state funding, raised by lottery, to continue its work. In 1806, an orphanage was founded by the Society. At the time, it was only the second such institution in the young country, the other being in Charleston, South Carolina. *See* DAVID M. SCHNEIDER, THE HISTORY OF PUBLIC WELFARE IN NEW YORK STATE 1609-1866, at 179-89 (1938). Amidst the development of institutional care for children in the mid-19th century, the population of children in almshouses continued to grow. "The census of 1880 showed that 7,770 children between the ages of two and sixteen were in almshouses in the United States . . . ." *See* Mason P. Thomas Jr., *Child Abuse and Neglect Part I: Historical Overview, Legal Matrix, and Social Perspectives*, 50 N.C. L. REV. 293, 302 n.37 (1972) (*citing* HOMER FOLKS, THE CARE OF DESTITUTE, NEGLECTED AND DEPENDENT CHILDREN 80 (1902)).

intrusions would challenge the accepted paternal role and parenting prerogatives. The texts reveal a range of legal questions posed by the courts in attempting to redefine the limits of state intervention. Courts debated whether it was the state of mind of the parent when perpetrating the beating,[58] the instrument used in the beating,[59] or the injury caused by the beating which should constitute evidence of abuse sufficient to sustain a prosecution.[60] Courts searched for objective measures of actionable abuse to avoid over-intrusion into family government. In one of a few early reported cases, a permanent serious injury was necessary to prosecute to protect children.

> It will be observed that the test of the defendant's criminal liability is the infliction of a punishment "cruel and excessive," and thus it is left to the jury without the aid of any rule of law for their guidance to determine. It is quite obvious that this would subject every exercise of parental authority in the

---

58. In *Stanfield v. State*, the court rejected the lower court charge, which focused on the instrument used in the beating, instead focusing on the manner of the Defendant.

> The charge asked and given does not mend the matter, which was that the jury could not convict the defendant unless the chastisement was done in a cruel or vindictive manner. . . . Was the correction moderate? . . . If it was, defendant was not guilty of an assault and battery at all. . . . If it was not moderate, but excessive, he was guilty as charged of an aggravated assault and battery by having exceeded the boundary of his legal right as guardian under the law, and placed himself in the attitude of a stranger and not a parent to the child . . . . Whether it is moderate or excessive must necessarily depend upon the age, sex, condition and disposition of the child, with all the attending and surrounding circumstances, to be judged of by the jury . . . .

43 Tex. 167, 168 (1875) (citations omitted). In *Johnson v. State*, similar reasoning was used.

> The right of parents to chastise their refractory and disobedient children is so necessary to the government of families, to the good of society, that no moralist or lawgiver has ever thought of interfering with its existence, or of calling upon them to account for the manner of its exercise, upon light or frivolous pretenses. But, at the same time, that the law has created and preserved this right, in its regard for the safety of the child it has prescribed bounds beyond which it shall not be carried. In chastising a child, the parent must be careful that he does not exceed the bounds of moderation and inflict cruel and merciless punishment; if he do, he is a trespasser, and liable to be punished by indictment. It is not, then, the infliction of punishment, but the excess, which constitutes the offence, and what this excess shall be is not a conclusion of law, but a question of fact for the determination of the jury.

21 Tenn. 283, 283 (1840).

59. In *Neal v. Georgia*, 54 Ga. 281, 282 (1875), the court affirmed that one "lick" with an old saw was "cruel and outrageous abuse of the parental authority, and made the perpetrator of it guilty," while also noting that a "very large margin must be left to the judgment of the parent." *Id.* at 283.

60. *See* State v. Jones, 95 N.C. 588, 589 (1886).

> correction and discipline of children--in other words, domestic government--to the supervision and control of jurors, who might, in a given case, deem the punishment disproportionate to the offense, and unreasonable and excessive. It seems to us, that such a rule would tend, if not to subvert family government, greatly to impair its efficiency, and remove restraints upon the conduct of children . . . . The test, then, of criminal responsibility is the infliction of permanent injury by means of the administered punishment, or that it proceeded from malice, and was not in the exercise of a corrective authority . . . . We do not propose to palliate or excuse the conduct of the defendant in the present case. The punishment seems to have been needlessly severe, but we refuse to take cognizance of it as a criminal act, because it belongs to the domestic rather than legal power, to a domain into which the penal law is reluctant to enter, unless induced by an imperious necessity.[61]

These criminal prosecutions were contemporaneous with the more famous 1874 case of Mary Ellen.[62] That case was championed in the front pages of the New York Times. It was brought by leaders of the New York Society for the Protection of Cruelty to Animals, heralding an era of child rescue and protection by private philanthropic agencies.[63] The growing work by such agencies developed the social and legal response to protect children from the abusive actions of their parents by "rescuing" children from them.

The Society leaders argued to the court on behalf of Mary Ellen, a young girl whose care was at issue, that children, as members of the animal kingdom, were entitled to protections at least equal to those provided animals.[64] The arguments were fashioned by Henry Bergh, founder and president of the Society, and Elbridge T. Gerry, counsel for the Society. The New York Times article of April 10, 1874 opened:

> It appears from proceedings had in Supreme Court yesterday, in the case of a child named Mary Ellen, that Mr. Bergh does not confine the humane impulses of his heart to smoothing the pathway of the brute creation toward the grave or elsewhere, but that he embraces within the sphere of his kindly efforts the human species also.[65]

---

61. *Id.* at 590-93.

62. *See* Lela B. Costin, *Unraveling the* Mary Ellen *Legend: Origins of the "Cruelty" Movement*, 1992 SOC. SERVICES REV. 203 (1991).

63. For a summary discussion of earlier foundations of child protection in response to child labor and orphans in colonial times and through the civil war, see NEIL A. COHEN, *Child Welfare History in the United States*, *in* CHILD WELFARE: A MULTICULTURAL FOCUS 13-21 (Neil A. Cohen ed., 1992). For a discussion of the emergence of protection of children within the societies for the protection of cruelty to animals before the Mary Ellen case, see Costin, *supra* note 62.

64. *See* Costin, *supra* note 62, at 204.

65. *Mr. Bergh Enlarging His Sphere of Usefulness: Inhuman Treatment of a Little Waif-Her Treatment-A Mystery to Be Cleared Up*, N.Y. TIMES, Apr. 10, 1874, at

The news articles explained that the child had been discovered when a woman, Etta Angell Wheeler, was on an "errand of mercy" to a dying woman. She was told by the woman of the desperate cries of a child in the next tenement building. Wheeler had tried repeatedly to gain entrance to the apartment to see the child. She was eventually let into the flat when Mr. Connolly, the man of the house, was not present, and she was able to observe and have a short visit with Mary Connolly, his wife, and Mary Ellen. While the news accounts of the time do not include commentary on domestic violence, the circumstances of Wheeler's failed attempts and subsequent entry, only when Mr. Connolly was absent, portray a scenario recognizable today as comorbid domestic violence and child abuse. Perhaps Mr. Connolly perpetrated some form of domestic violence against Mary Connolly and was also abusive toward Mary Ellen; Mary Connolly was therefore fearful of allowing Wheeler's entry in his presence. In another possible scenario, perhaps Mr. Connolly was abusive toward his wife but Mary Connolly was perpetrating the abuse against Mary Ellen either individually or along with Mr. Connolly.

Reports indicate that Wheeler went to several institutions to seek help for the child, before she found Bergh and pleaded for his assistance.[66] It was known at the first hearing that Mary Ellen was living with Mary and Francis Connolly and they were charged with cruel abuse against her, but that they were not her natural parents.[67] It is not clear how this casual custodianship affected the willingness of the agency, court and public to champion prosecution of the Connollys.

The case was originally prosecuted against both Mr. and Mrs. Connolly. Mary Ellen's ill health, lack of proper clothing and frequent abuse with whips, scissors, and slaps must have been known, if not perpetrated, by both adults in the home. Even if this could not be proven, it was only Mrs. Connolly, on a day when Mr. Connolly was not present, who allowed Mrs. Wheeler into the apartment to discuss Mary Ellen. Only Mrs. Connolly ever appeared in court. Only she was ultimately tried and sentenced for the

---

8.

66. Perhaps because Mrs. Wheeler's husband was a newspaper man, the case is graphically and fully reported in the paper. *See* Costin, *supra* note 62, at 210; N.Y. TIMES, Apr. 10, 1874, at 8; N.Y. TIMES, Apr. 11, 1874, at 2; N.Y. TIMES, Apr. 14, 1874, at 2; N.Y. TIMES, Apr. 28, 1874, at 8; N.Y. TIMES, Dec. 27, 1874, at 12. For a compilation of related articles and papers of the New York Society for the Prevention of Cruelty to Children, see CHILDREN AND YOUTH IN AMERICA: A DOCUMENTARY HISTORY, Vol. II, 1866-1932 Parts 1-6, 185-97 (Robert H. Bremner ed., 1971) [hereinafter CHILDREN AND YOUTH, Vol. II, Parts 1-6].

67. *See The Mission of Humanity*, N.Y. TIMES, Apr. 11, 1874, at 2; Costin, *supra* note 62, at 207; CHILDREN AND YOUTH, Vol. II, Parts 1-6, *supra* note 66, at 187.

abuse.[68] This celebrated abuse case targeted the "mother" caretaker. No male was held accountable. The case signals the entry of private philanthropic agencies into the legal system on behalf of abused and neglected children. It also foreshadows the treatment of mothers and lack of attention paid to holding abusive fathers and husbands accountable before dependency courts.[69]

This may be an infamous case of "failure to protect." Because he was deceased, this case did not impinge on the rights of Mary Ellen's father. Mary Ellen's mother's rights were not at issue in the prosecution because her mother had abandoned her or had been separated from her much earlier. Mary Connolly, taking on the custodianship of Mary Ellen as her stepmother, may have been prosecuted because of this legal relationship. It is also possible that she was prosecuted, despite her ultimate willingness to allow Mrs. Wheeler into the flat against the wishes of her husband, because she was the "mother" of the child and was therefore responsible for her care under nineteenth century notions of parenting. Mr. Connolly's drunkenness, violence against Mary Ellen and possible violence against Mrs. Connolly were ignored, perhaps because he had no legal relationship to the child, or perhaps because his behavior, as an unrelated "man in the house" was not as shocking to the norms of child care at the time.

The publicity surrounding this case led to important results for the future of child protection. The activities of private provider agencies acting on behalf of abused and neglected children increased significantly. In the same year, a private provider agency, the New York Society for the Prevention of Cruelty to Children, was formed with Elbridge Gerry as its counsel.[70] By 1880, thirty-three such societies existed in the United States,[71] most of them in the business of rescuing both animals and children. As Bergh explained:

> The protection of children and the protection of animals are combined be-

---

68. *See Mrs. Connolly, the Guardian, Found Guilty, and Sentenced to One Year's Imprisonment at Hard Labor*, N.Y. TIMES, Apr. 28, 1874, at 8.

69. *See generally* Marie Ashe, *Postmodernism, Legal Ethics, and Representation of "Bad Mothers," in* MOTHERS IN LAW: FEMINIST THEORY AND THE LEGAL REGULATION OF MOTHERHOOD 142-66 (Martha Albertson Fineman & Isabel Karpin eds., 1995); Marie Ashe & Naomi R. Cahn, *Child Abuse: A Problem for Feminist Theory*, 2 TEX. J. WOMEN & L. 75 (1993); V. Pualani Enos, *Prosecuting Battered Mothers: State Laws' Failure to Protect Battered Women and Abused Children*, 19 HARV. WOMEN'S L.J. 229 (1996); NAT'L CENTER ON WOMEN AND FAMILY LAW, FAILURE TO PROTECT: A REFERENCE MANUAL FOR NEW YORK ATTORNEYS REPRESENTING BATTERED WOMEN AT RISK OF LOSING THEIR PARENTAL RIGHTS FOR FAILURE TO PROTECT THEIR CHILDREN FROM THE ABUSER (1993) [unpublished monograph].

70. *See* CHILDREN AND YOUTH, Vol. II Parts 1-6, *supra* note 66, at 189-97.

71. *See* Linda Gordon, *Child Abuse, Gender and the Myth of Family Independence: Thoughts on the History of Family Violence and Its Social Control 1880-1920*, 12 REV. L. & SOC. CHANGE 523 (1983-84).

> cause the principle involved, i.e., their helplessness, is the same; because all life is the same, differing only in degree of development and expression; and because each profits by association with the other.[72]

These early efforts were aimed at rescuing children and, sometimes, prosecuting the adults who brutalized them. The societies did not see as their mission any education or support of parents or treatment of families. As Gerry explained:

> The SPCC was simply created as a hand affixed to the arm of the law, by which the body politic reaches out and enforces the law. The arm of the law seizes the child when it is in an atmosphere of impurity, or in the care of those who are not fit to be entrusted with it, wrenches the child out of these surroundings, brings it to the court, and submits it to the decision of the court--unless, on the other hand, it reaches out that arm of the law to the cruelist, seizes him within its grasp, brings him also to the criminal court and insures his prosecution and punishment. These are the functions of our societies.[73]

Until the mid-twentieth century, private philanthropic agencies intervened on behalf of abused and neglected children. States made fledgling criminal prosecution efforts on behalf of individual children but neither the civil nor the criminal response was uniform or broadly applied. State criminal prosecutions for abuse continued without any federal legislative guidance into the 1970's.[74] No federal law codified the duties of the state to protect children from abuse and neglect.

---

72. *See* Costin, *supra* note 62, at 213.

73. ELBRIDGE GERRY, THIRTY FIRST ANNUAL MEETING OF THE AMERICAN HUMANE ASSOCIATION (Albany, N.Y. 1907), *cited in* Costin, *supra* note 62, at 219. At this 1907 speech at the Annual Meeting of the Society, Gerry was also clear that the society's purpose was to rescue children and refer their parents for prosecution, not to provide treatment. The New York Society for the Prevention of Cruelty to Children (hereinafter "N.Y.S.P.C.C.") was "not created for the purpose of educating or reforming children, or seeing that they were transported into other homes." *Id.* This description was meant in part to differentiate the purposes of the Society from the work of the New York Children's Aid Society which gathered up children from the industrializing Northeast cities and sent them on "orphan trains" to the rural Midwest where they were given "proper homes" through an informal indenture. These *rescue* efforts are also distinguishable from broader turn-of-the-century *preventative* efforts. Carl Carstens led the first of these agencies, the Massachusetts Society for the Prevention of Cruelty to Children. In 1907, at that Society's Annual Meeting, Carstens stated the broader mission of child protection: "Children will still need to be rescued from degrading surroundings for many years to come . . . but the society recognizes more definitely that it is a preventive agency." MASSACHUSETTS SOC. FOR PREVENTION OF CRUELTY TO CHILDREN, 1907 ANN. REP. 27, *in* Paul Gerard Anderson, *The Origin, Emergence, and Professional Recognition of Child Protection*, 63 SOC. SERVICES REV. 222, 224 (1989).

74. For a summary of the history of child protection and of the present operation of the system, see Mangold, *supra* note 6.

A seminal event in the history of state intervention to protect children from child abuse and neglect was the 1962 publication of Battered-Child Syndrome by Dr. Henry Kempe.[75] Kempe was a pediatrician who worked with pediatricians and radiologists to identify causes of suspicious injuries to children.[76] With new knowledge about injuries that could only be caused by abusive behavior, states moved to codify responses to protect children. Between 1963 and 1967, every state passed a statute requiring some form of reporting of incidents of child abuse. Reporting triggers the state response based on a child's right to protection from abuse and neglect.

In 1973, the Senate Subcommittee on Children and Youth of the Committee on Labor and Public Welfare held hearings in Washington and at children's hospitals around the country on the needs of abused and neglected children.[77] Bills were introduced in both the House and Senate, but the Senate subcommittee chaired by Walter Mondale held the main hearings. In a letter of transmittal to the Senate Committee Chairman, Mondale explained the need for legislation:

> The Subcommittee held hearings in Washington, New York, Denver and Los Angeles. Members of the Subcommittee personally visited victims of child abuse in hospitals and observed firsthand the operations of multi disciplinary child abuse teams in several cities. We were appalled to learn how many abused and neglected children there are and how little is being done to help them and their troubled families. Statistics vary widely, but there is little question that thousands and thousands of youngsters suffer severe physical and emotional abuse every year. This is a problem that cuts across social and economic barriers. It occurs in all kinds of neighborhoods. Yet there was no focused Federal effort to deal with the problem. Nowhere in the Federal government could we find one official assigned full time to the prevention, identification and treatment of child abuse and neglect.[78]

One year later in 1974, The Child Abuse Prevention and Treatment Act (CAPTA) was passed.[79] CAPTA initiated a federal response to child abuse based on a child's right to protection by the state where there is evidence of parental harm. It formulated the mandates for the development of a bureaucracy within the Department of Health, Education, and Welfare (HEW) (now

---

75. *See* C. Henry Kempe et. al., *The Battered-Child Syndrome*, 181 J. AM. MED. ASS'N. 105 (July 7, 1962).

76. *See* C. Henry Kempe, *Paediatric Implications of the Battered Baby Syndrome*, 46 ARCHIVES OF DISEASE IN CHILDHOOD 28, 28 (1971).

77. *See Child Abuse Prevention Act, 1973: Hearings on S. 1191 Before the Subcomm. on Children and Youth of the Comm. on Labor and Public Welfare*, 93d Cong. 2 (1973) [hereinafter *Mondale Hearings*].

78. *Id.* at 2 (letter Walter Mondale sent to Hon. Harrison A. Williams).

79. One year later, Congress passed the Child Abuse Prevention and Treatment Act, 42 U.S.C. § 5101 (Supp. 1995).

called the Department of Health and Human Services) to gather information and expertise on the problem of child abuse, a largely undocumented subject at the time.[80]

Most important for the subsequent history of the federal/state relationship addressing child abuse, CAPTA contained provisions for a grant program.[81] Eligibility for grants required states to follow a series of mandates in order to receive the funds.[82] Those provisions concerned reporting, investigating, confidentiality of record keeping, and law enforcement cooperation.[83] They were the earliest versions of the more complete and complicated federal-to-state reimbursement system which funds state child protection systems today.

CAPTA established a minimum law enforcement-like state response to determine which children had a plausible right to protection.[84] By focusing on reporting and investigating without attention to supporting families or preventing abuse, parents as "alleged perpetrators" were pitted against the agencies. Parental empowerment was in conflict with child protection. The key state response to protect children from abuse became the mandatory reporting, investigating, and record keeping system that is commonly known as the child protective services system. While all states had some form of reporting law in place before CAPTA, few met the more rigorous CAPTA requirements before 1974. CAPTA, in effect, maintained continuing attention on reporting laws, confidentiality, and investigation.

CAPTA addressed the "front end" of the system, bringing attention to troubled families and children and investigating them. It soon became clear that permanency planning was crucial as children were being placed in foster care in increasing numbers. Following the passage of CAPTA, the numbers of children reported as abused and neglected exploded, and state-based foster care systems were flooded with children placed as a result of reporting and investigation through child protective services. Senator Cranston summarized the situation before the Senate in 1979:

> The number of children in foster care in 1977 was approximately 500,000, nearly three times the number of children in foster care as compared to 1961. In only one of every five cases does the services plan for these foster children recommend a specific length of placement. In other words, the so-called temporary provision of foster care has no definite target date for ending the placement and for placing the child in a permanent family setting. Over half the children in foster care have been away from their families for

80. *See Mondale Hearings*, *supra* note 77 (Letter of Transmittal).
81. *See id.*
82. *See id.*
83. *See* CAPTA, § 4(b)(3).
84. *See id.*

> more than 2 years- about 100,000 children have spent more than 6 years of their lives in foster care. Nearly one-fourth of the children have been in three or more foster family homes. Even in cases where the agency had developed a plan for returning the child to his or her home, in one-third of the cases, there was no plan for visits between the child and the parent or another person who would care for the child if returned home. There are more than 100,000 children in foster care awaiting adoption.[85]

As reports mounted as a result of CAPTA requirements, foster care became the expedient and perhaps sole resource to address the children's safety.[86] Separation became the means of protection and protection was divorced from caretaker empowerment.

3. Protection of Children by Empowerment of Parents

Concerns that children were being unnecessarily placed outside their homes and were languishing without permanency in foster care led to passage of the Adoption Assistance and Child Welfare Act of 1980 (AACWA).[87] Many of the children were neglected, not abused. For such children, it was hoped that temporary supportive services would empower parents to resume their authority and safely care for their children. For the first time, empowerment became a preferred means of protection. The AACWA imposed the mandate that states provide a plan to the federal government requiring the state-based public agency to make "reasonable efforts" to prevent placement in foster care or achieve reunification for children temporarily placed.[88] The law also provided for adoption subsidies to encourage the adoption of children out of foster care who could not be reunified.[89] States codified the reasonable efforts and adoption language in their laws.[90] If states failed to meet the mandates of the law they would risk losing eligibility for matching federal reimbursement for their foster care expenses.[91] As a consequence of the fiscal incentives offered in AACWA, family preservation efforts to reunify families flourished and the number of children in foster care began to decrease.[92] Unnecessary foster care place-

85. 125 CONG. REC. 110, 59 (1979) (statement of Sen. Cranston).

86. *See id.*

87. *See* Pub. L. No. 96-272, 94 Stat. 500 (1980) (codified as amended in 42 U.S.C. §§ 602, 608, 620-28, & 670-76 (1997)).

88. *See id.* at § 671(a)(15)(A (1994 & Supp. 1997).

89. *See* 42 U.S.C. § 671(a)(15)(B) (1994 & Supp. 1997).

90. *See, e.g.*, CAL. WELF. & INST. CODE § 366 (West 1999), MISS. CODE ANN. § 43-15-13 (1999).

91. *See* 42 U.S.C. § 671(a) (1994 & Supp. 1997).

92. *See* CHILD. DEF. FUND, STATE OF AMERICA'S CHILDREN: 1994 YEARBOOK 22

ment and foster care drift were diminished by requiring case planning, case reviews, reunification efforts and subsidies for adoptions of children leaving foster care.[93] The law required that whenever the determination to place a child in foster care was pending, the court had to make a finding as to whether "reasonable efforts" had been made to prevent the placement. It was hoped that this procedural requirement could reduce the unnecessary placement of children in foster care when services to their families could maintain them safely at home.[94] The goal was never to eliminate foster care. There was and is a point on the spectrum of cases where separation is the only means of protecting children. Instead, the law was targeted at the large number of unnecessary placements of children who could remain safely at home with supportive services.[95] The provisions of the Social Security Act codifying the AACWA have been revised over the past two decades to require a myriad of planning processes in exchange for federal reimbursement for foster care.[96] The state is required to have a written document for each child in care which includes, in part,

> [A] plan for assuring that the child receives safe and proper care and that services are provided to the parents, child, and foster parents in order to improve the conditions in the parents' home, facilitate return of the child to his own safe home or the permanent placement of the child . . . .[97]

The protection of the child includes not only state-provided services to alleviate the risks of child abuse and neglect but also to reduce the risks of unnecessary placement. In seeking to protect children from unnecessary placement, the law made the right to protection and the empowerment of families complimentary, not conflicting policies. The policy of protection by empowerment of families was short lived and perhaps never fully exercised. Preventive and supportive services were always "capped" services and funding for them was never sufficient to meet the local need. In contrast, placement services were and are uncapped and available in unlimited supply. These fiscal incentives and disincentives may have undermined the spirit of AACWA.

Following passage of the AACWA, criticism began to mount over the

---

(1994).

93. *See* Pub. L. No. 96-272, § 471(a)(15)(B), § 471(a)(16), § 473(a)(1), 94 Stat. 500, 503, 504 (1980) (codified as amended in 42 U.S.C. §§ 602, 608, 620-28 and 670-76 (1997)).

94. *See* 42 U.S.C. § 671(a)(15) (1994 & Supp. 1998).

95. Richard Wexler makes a strong argument for family preservation, especially intensive family preservation services along the Homebuilders Model. Such services are designed to prevent unnecessary placements. *See* Richard Wexler, *Take the Child and Run: Tales from the Age of ASFA*, 36 NEW ENG. L. REV. 123 (2001).

96. *See* 42 U.S.C. §§ 601-79b (1994 & Supp. IV 1998).

97. 42 U.S.C. § 675(1)(B) (1994 & Supp. IV 1998).

perceived emphasis on empowering families. While the law had also provided for adoption subsidies to encourage and support the adoption of children out of foster care, the emphasis on preservation and reunification was criticized for fostering a climate where children were left in unsafe homes and sometimes returned to unsafe homes.[98] The concern was that empowering families was being mandated at the expense of the protection of children. While the policy of the AACWA sought to empower families as the best means of protecting children, criticism of the empowerment efforts pitted family empowerment against the protection of children.

In the 1990's, the number of children in foster care began to increase once again. While the reasons for this are complex, the increase is usually attributed to the crack epidemic in the inner cities and the increasing percentage of children living in desperately poor conditions with young, unmarried mothers.[99] A series of highly publicized brutal deaths of children who were "known to the public agency" and provided with failed preventive services in their own homes instead of being placed in foster care led to a growing outcry for reform of the system.[100]

4. Retreating from Protection by Empowerment of Parents with a New Emphasis on Termination of Parental Rights

The Adoption and Safe Families Act of 1997 (ASFA)[101] was a partial response to renewed calls for protection – both for swifter removal from abusive homes and for expedited adoptions. Protection was again viewed as it had been before the AACWA – as requiring swifter removal from investigated homes instead of supportive services into the homes. ASFA provides exceptions to the reasonable efforts requirement introduced in the AACWA when "aggravated circumstances" are present.[102] The section providing for the exceptions appears uncontroversial at first glance, citing torture, death of another child, or sexual abuse as examples; but a more broad exception comes from the "aggravated circumstances" catch-all provision which is left to the states to define.[103] It moves the balance from empowering parents

---

98. *See, e.g.*, NEW YORK STATE COMMISSION ON CHILD ABUSE, FINAL REPORT (1996).

99. *See id.* at 22-25.

100. *See id.* at 3-4, 28-29. For a discussion of the perceived problems with the reasonable efforts requirement leading to the adoption of the Adoption and Safe Families Act, see Christine H. Kim, *Putting Reasonable Back into the Reasonable Efforts Requirement*, 1999 U. ILL. L. REV. 287 (1999).

101. *See* Pub. L. No. 105-89, 111 Stat. 2115 (codified as amended at 42 U.S.C. §§ 671-679b (1994 & Supp. 1998)).

102. *See* 42 U.S.C. § 671(a)(15)(D)(i) (1994 & Supp. 1998).

103. *See id.* In his article in this volume, Richard Wexler describes the impact of ASFA in very stark terms. *See* Wexler, *supra* note 95. While I see ASFA as a par-

toward a stronger child's right to protection. It does not abandon protection by empowerment, but has shifted the policy focus toward protection by removal.

ASFA was passed to clarify that the health and safety of the child should always be paramount and to detail circumstances under which reasonable efforts did not have to be pursued. ASFA retreats from the goal of reunification in certain cases where "aggravated circumstances" or felony convictions exist. In a shift away from reunification and toward adoption the law also required that child welfare agencies file a petition to terminate parental rights for children in care for fifteen of the last twenty-two months.[104] This sets an end limit on protection by empowerment. In the earlier House version, the law expedited termination proceedings only for children under the age of ten.[105] While the law ultimately passed did not include this maximum age requirement, the sentiment of the Congress throughout the debate was that the bill would mainly impact adoptions for this preteen population. The final law includes additional important provisions for older children, namely allowing states to designate them as "special needs" children due to age and making them eligible for adoption assistance by providing them with continuing health insurance.[106]

In 1997, with the passage of ASFA, Congress altered the federal mandates on states providing foster care services to abused and neglected children.[107] The new mandates shifted the focus from preservation and reunification toward a new emphasis on swifter termination. Whether this was a balanced adjustment of protection and empowerment from the empowerment-driven AACWA or a hard swing toward protection that will later be re-balanced is yet unclear.

In 1999, Congress again amended the law governing these placements by focusing on the needs of older children in out-of-home care and passing the

---

tial retreat from the philosophy of family preservation, I would not go so far as him in characterizing this as the age of "take the child and run." *See id.* If no ASFA exception applies, child welfare agencies are still required to make *reasonable efforts* to prevent placements and reunify children. *See id.*

104. For an article critical of this shift in emphasis, see Naomi R. Cahn, *Children's Interests in a Familial Context: Poverty, Foster Care, and Adoption*, 60 OHIO ST. L.J. 1189 (1999).

105. *See* Adoption Promotion Act of 1997, § 3(a)(3)(E) (reported favorably out by the Committee of Ways and Means).

106. *See* Adoption and Safe Families Act of 1997, Pub. L. No. 105-89, 111 Stat. 2115, § 306 (1997) (amending 42 U.S.C. § 671(a) (1994 & Supp. 1997)). *See* 143 CONG. REC. S12, 513-26 (daily ed. Nov. 13, 1997) (statement of Sen. Chafee emphasizing importance of the medical care adoption subsidy for teens adopted out of care).

107. *See* Adoption and Safe Families Act of 1997, Pub. L. No. 105-89, 111 Stat. 2115 (codified as amended at 42 U.S.C. §§ 671-679b (1994 & Supp. 1997)).

Foster Care Independence Act (FCIA).[108] FCIA and ASFA represent a change in philosophy from the previous law, moving from protection by supported family preservation or reunification toward protection by alternative placements – adoption or independent living arrangements. FCIA was passed to enhance services for children aging out of foster care with the goal of "independent living." In a sense, this law was necessary because of the failure of all previous efforts to find and support families for children in foster care. It made possible a continuing claim to the entitlement of foster care based upon a right to protection until age twenty-one even for older teens who had reached the age of eighteen. This law is an important step in recognizing the role of child empowerment in protecting children, but as is discussed *infra* Part III, it does not go far enough to continue protecting while empowering older foster children.

D. Protection and Empowerment in the Operation of the Current Foster Care System

The fundamental rights of parents can be disturbed and the child's right to protection triggered by allegations of abuse or neglect. As discussed *supra* Part I.C.3, all states have in place a child protection system which responds to such allegations by investigating families and temporarily removing children when it is deemed necessary for their safety.[109] Sometimes parents voluntarily place their children outside the home during the course of the investigation and sometimes a court mandates the placement.[110] When placed, these children enter a system of out-of-home care which can include non-relative foster family homes, foster homes provided by relatives called kinship homes, group homes or larger residential settings. Once placed in out-of-home care by the child protection system through the voluntary or involuntary temporary surrender of parents, these children become foster children.[111] Foster care is always meant to be temporary. A permanency plan is written for every child in care with a recognized goal (reunification, adoption, independent living) and a set of action steps to achieve that goal.

The child protective services system is triggered by a report of abuse or neglect as defined in state law under the requirements of CAPTA. Reports are made by voluntary or mandated reporters to hotlines that federal law requires every state to operate.[112] Voluntary reports can be made by neighbors, friends, family members – anyone who suspects child abuse or neglect on the part of a caretaker.[113] If the reporter provides adequate information to the hotline operator, the report triggers investigation by the

---

108. Foster Care Independence Act of 1999, Pub. L. No. 106-169, 113 Stat. 1822 (1999).

local child protection agency.[114] Private individuals are vitally involved in activating the child protective services system through reporting, but hotlines are maintained and initial investigations are done by the public agency.

Mandated reporters are crucial to the child protective services system. Through mandated reporting, the *parens patriae* power of the state is exercised to conscript professionals who work with children to become partial state agents in protecting children from harm. These reporters generally are persons who work in professions or roles that bring them into contact with children.[115] If these professionals suspect or believe that children with whom they come into contact in the course of their employment are suffering from abuse or neglect, confidentiality and privilege are forfeited and the professionals are mandated to report the abuse or neglect to the state operated child protection system. Some states require that professionals who work with parents and have reason to suspect abuse are also mandated to report.[116] These professionals are not merely invited to participate on behalf of children; they are required to do so regardless of their professional opinion as to the wisdom, value, or safety of reporting.[117]

In effect, the state forces professionals to participate and invites the non-professional community member, the modern day Etta Angel Wheeler, to make reports voluntarily. This reporting system is not a novel

---

109. The state child protection systems operate under federal mandates codified in the Social Security Act since the Child Abuse Prevention and Treatment Act of 1974, Pub. L. No. 93-247, 88 Stat. 4 (codified as amended in 42 U.S.C. §§ 5101-119c (1995 & Supp. 2000)).

110. A voluntary placement agreement with the parent or legal guardian or a court adjudication is necessary for federal reimbursement for the placement under 42 U.S.C. § 672(a)(1) (Supp. 1998).

111. This article focuses on out-of-home care provided through the child protection system and not as a result of delinquency, mental illness or status offenses which result in placement outside the child welfare system. For a discussion of foster care in the delinquency system, see also Burt Galaway et al., *Specialist Foster Family Care for Delinquent Youth*, 59 FED. PROBATION 19 (1995). For a discussion of privatization of care in these systems, see Mangold, *supra* note 3 at 1300, n.14.

112. *See* 42 U.S.C. § 5106a(b)(1)(A) (1995 & Supp. 2000).

113. Abuse by a non-caretaker can only be pursued criminally. The civil system is reserved for intra-familial violence.

114. *See* 42 U.S.C. § 5106a(b)(10) (1995 & Supp. 2000). For examples of state operational directives, see N.Y. SOC. SERV. LAW § 422(2)(a) (McKinney 1992); Boland v. State, 638 N.Y.S.2d 500 (1996).

115. *See* 42 U.S.C. § 5106a(b) (1995 & Supp. 2000).

116. N.Y. SOC. SERV. LAW § 413 (McKinney 1992).

117. *See generally* MURRAY LEVINE & HOWARD J. DOUECK, THE IMPACT OF MANDATED REPORTING ON THE THERAPEUTIC PROCESS PICKING UP THE PIECES (1995).

creation of the 1960's; rather, it was an evolving codification of the child protection system developed since colonial times and expanded by private philanthropic agencies and criminal prosecutions at the turn of the century. In 1997, approximately three million children were reported abused or neglected.[118] Depending on the severity of the allegations, child protective services workers must respond within the period of time required by state law to determine whether there is sufficient evidence to support the allegations.[119] If the workers find that there is not sufficient evidence, the reports are considered "unfounded" and the cases are closed.[120] Of the three million reports in 1997, investigation by agencies confirmed that abuse or neglect had occurred in just under one million cases.[121] In other words, each year, over two million cases are investigated but no further action is deemed necessary.[122] Among substantiated reports, data shows that African American children are over-represented; African American children comprise fifteen percent of the population, but make up twenty-eight percent of the children with substantiated reports.[123] Neglect is the allegation substantiated for most children in the dependency system. In 1995, forty-two percent of the substantiated cases were classified as neglect while only twenty-two percent were classified as abuse. Sexual abuse was the confirmed allegation in eleven percent of the cases.[124] If a public agency worker confirms a report, often referred to as "indicating" or "substantiating" the report, the agency decides what further action is necessary. In many cases, the perpetrator is removed from the home by the time the investigation is completed, so no further services or supervision is required. In other instances, the family may be given the option of "voluntarily accepting services" from the agency, services that are often delivered by a private provider agency as a subcontractor to the public agency. These services may range from parenting classes and periodic visits to the home to out-of-home placement of the child. By federal law, the agency must make reasonable efforts to keep the family together unless an exception under ASFA applies. If services are not available to keep the child safely at home despite reasonable efforts to provide such

---

118. *See* CHILD. DEF. FUND, *supra* note 13.

119. *See* N.Y. SOC. SERV. LAW § 424(6) (McKinney 1992).

120. *See id.* § 424(7).

121. *See* CHILD. DEF. FUND, *supra* note 13, at 85; MICHAEL R. PETIT ET AL., CHILD ABUSE AND NEGLECT: A LOOK AT THE STATES, 3 (1997).

122. In New York, as a result of the death of Eliza Izquierdo, the law was amended to keep records of unsubstantiated reports for use in future investigations. The law is known as "Eliza's Law." N.Y. SOC. SERV. LAW § 422(5) (McKinney 1992).

123. *See* PETIT *supra* note 121, at 19.

124. *See id.* at 30.

services, the child may be removed.[125] Data compiled from thirty-five states shows that 130,685 children were removed due to abuse and neglect in 1995.[126] As of March 1999, there were a record 547,000 children in out-of-home placements as a result of removal due to allegations of abuse and neglect.[127]

Once a child is removed to out-of-home placement, the case must be reviewed periodically by the court or by administrative review.[128] If the family refuses voluntary services in the home or refuses to voluntarily place the child, the public agency may petition the court to find that the child is abused or neglected and to mandate a disposition.[129] The disposition can include services in the child's home and/or out-of-home placement for the child.[130] A system of procedural requirements comes into play to provide periodic dependency hearings on the parents' rehabilitation, the agency's efforts, and the child's safety.[131] The judge in such proceedings rarely considers the arguably important issues of child support, joint custody, or domestic violence[132] in formulating mandatory orders, but advocates against domestic violence are increasingly encouraging such considerations.[133] The ultimate protective tool of the system is removal of the children from the home. Increasingly under the provisions of ASFA, this can lead to termination of parental rights and an order that frees a child for adoption.

E. Empowerment vs. Protection in the Debate Over the Proper Role of the Attorney for the Child

---

125. *See* 42 U.S.C. § 671(a)(15)(C) and (D) (1994 & Supp. IV 1998).

126. *See* PETIT, *supra* note 121, at 38.

127. *See* CHILD. DEF. FUND, *supra* note 13, at 84, 86.

128. *See* 42 U.S.C. § 675(5)(B) (Supp. 1998).

129. *See* 42 U.S.C. § 671(a)(15)(C) (Supp. 1998).

130. *See* 42 U.S.C. § 671(a)(15)(A)-(C) (Supp. 1998).

131. *See* 42 U.S.C. § 675(5)(B) (1994).

132. While inquiries into violence against the mother in the home may be part of a risk assessment at the outset of a case, the focus is on the children. If the violence threatens the children, the children can be removed. A referral may be given to the mother to tell her how to remove the perpetrator through the domestic violence system and her swiftness and success in doing so may determine how she is judged as a parent. She will also be given a variety of tasks to work toward reunification with her child. Often, she is required to maintain a stable home and income. This may make it difficult for her to remove a batterer who is the family's source of income. *See generally* NAT'L CENTER ON WOMEN AND FAMILY LAW, *supra* note 69.

133. Without such comprehensive orders, conflicting orders regarding custody and visitation arrangements can jeopardize the safety of mothers and children. *See* NAT'L COUNCIL OF JUVENILE AND FAMILY COURT JUDGES, MODEL CODE ON DOMESTIC AND FAMILY VIOLENCE (1994).

The shifting balance between protection and empowerment is evident in the history, laws and current operation of the child protection system and the foster care system as discussed *supra* Part I.B-D. As a theoretical debate, it is still being waged among scholars studying the child protection system when considering the proper role of the attorney for the child.The Child Abuse Prevention and Treatment Act ties delivery of federal matching funds to the requirement that a guardian ad litem be appointed for children in all dependency court proceedings.[134] Regulations promulgated to interpret this provision of CAPTA state that the guardian ad litem must "represent and protect the rights and best interests of the child."[135] The ambiguity in this terminology has left open to interpretation whether attorneys must be appointed to represent children or if others, professionals or volunteers, can take on this role.[136] Even when an attorney is appointed, as is required in many states, the role of that attorney in the proceedings is unclear. Must the lawyer operate in a traditional lawyer-client relationship to fully protect and exercise their child client's rights (empowerment model) or can the lawyer generally, or at least in some instances, represent some notion of the child's best interest (protection model)?[137]

Scholarship and commentary on the appropriate model of representation of children in dependency proceedings collectively assumes one of these two positions.[138] The empowerment model posits that children should be treated as autonomous clients and their positions should be zealously represented before the court. This is sometimes called the "autonomy" or

---

134. *See* 42 U.S.C.A. § 5106a(b)(2)(A)(ix) (1995 & Supp. 2000).

135. 45 C.F.R. § 1340.14(g) (1998).

136. For a discussion of the advantages of employing social workers and others who are specifically trained to work with children, especially young children, see Annette R. Appell, *Decontextualizing the Child Client: The Efficacy of the Attorney-Client Model for Very Young Children*, 64 FORDHAM L. REV. 1955 (1996).

137. This ambiguity exists in law and in practice. For a more detailed discussion of the openness to interpretation of the advocates' role, see Emily Buss, *"You're My What?" The Problem of Children's Misperceptions of Their Lawyers' Roles*, 64 FORDHAM L. REV. 1699, 1700 (1996); Donald N. Duquette & Sarah H. Ramsey, *Representation of Children in Child Abuse and Neglect Cases: An Empirical Look at What Constitutes Effective Representation*, 20 U. MICH. J.L. REFORM 341 (1987); Sarah Ramsey, *Representation of the Child in Protection Proceedings: The Determination of Decision-Making Capacity*, 17 FAM. L.Q. 287 (1983).

138. There is general agreement that children in delinquency proceedings should be represented as autonomous clients. This consensus flows from a reading of *In re Gault*, 387 U.S. 1 (1967), which focuses on protection of a child's procedural due process rights when a child's liberty is at stake. In dependency cases, where a child's placement is at issue, the right to an attorney should be interpreted as in *In re Gault*. In custody cases, the role of counsel is often debated as it is in dependency proceedings. *See* Buss, *supra* note 137, at 1700-01 n.3.

"expressed interest" view. The protection model advocates that children, as not fully competent clients, need to be protected and a position of their "best interest," whether or not it coincides with their expressed interest, should be advanced.[139] Proponents of the autonomy or empowerment model advocate that children should be carefully interviewed, and their attorneys should put the child's expressed interests before the court. This view, prevalent in legal literature,[140] states that a lawyer's professional role dictates such advocacy on behalf of the client and that any other model which allows for the opinion of the lawyer to be presented to the court is not legitimate.[141]

There are many commentators who state their preference for an empowering model of representation but import a variety of caveats. Some argue that a lawyer should be excused from following her client's wishes when the client is too young, the matter is too important, or the proceeding is too chaotic to assure that all proper information will be before the court.[142] Usually these concerns are most keen when the issue before the court is whether a child should be returned to her parents. The concern is that young, impressionable children will wish to return to their parents against their own best interest. Some commentators suggest that when these concerns are present in a case, the child's attorney should assert her own opinion about what is in the child's "best interest."[143] Others argue that the lawyer should act as an investigator and objectively ensure that all information is in evidence before the court so that a judge can properly decide what is in the child's best interest.[144] At least one commentator urges that lawyers for young children

---

139. *See generally, Proceedings of the Conference on Ethical Issues in the Legal Representation of Children*, 64 FORDHAM L. REV. 1301 (1996) (hereinafter "*Conference Proceedings*"). Articles in this special issue deal specifically with the different models. *See, e.g.*, Martin Guggenheim, *A Paradigm for Determining the Role of Counsel for Children*, 64 FORDHAM L. REV. 1399 (1996); Jean Koh Peters, *The Roles and Content of Best Interests in Client-Directed Lawyering for Children in Child Protective Proceedings*, 64 FORDHAM. L. REV. 1505 (1996); Katherine Hunt Federle, *The Ethics of Empowerment: Rethinking the Role of Lawyers in Interviewing and Counseling the Child Client*, 64 FORDHAM L. REV. 1655 (1996).

140. *See generally, Conference Proceedings, supra* note 140; *American Bar Association Standards of Practice For Lawyers Who Represent Children in Abuse and Neglect Cases*, 14 CHILD. RTS. CHRON. 25, 25-26 (1995-96).

141. *See* Federle, *supra* note 139, at 1655.

142. *See* Ramsey, *supra* note 137 at 287.

143. *See* Peters, *supra* note 139; Peter Margulies, *The Lawyer as Caregiver: Child Client's Competence in Context*, 64 FORDHAM L. REV. 1473, 1473-78 (1996).

144. *See* Martin Guggenheim, *The Right To Be Represented But Not Heard: Reflections on Legal Representation for Children*, 59 N.Y.U. L. REV. 76 (1984) (discussing the various attorney roles, including the role of investigator).

follow the dictates of the underlying substantive law.[145]

The first caveat, that capacity must be considered before a child can be represented as an autonomous client, is the most prevalent in scholarly legal literature. It has led to a debate over the proper age at which capacity can be presumed,[146] who should determine capacity,[147] and the attempts a lawyer can make to discern a sense of the child's wishes even when the child is considered impaired.[148]

The second caveat, that a child's lawyer can be excused from the traditional lawyer role when the stakes are too high for the child's safety, usually imagines a situation where placement or reunification is before the court and the child is urging a position which would place her with her parents.[149] This occurs at the dispositional or dispositional review stage of proceedings. In the chaotic, crisis driven dependency system, the concern is also raised that the adversarial system cannot be presumed to act properly. Critics suggest that not all of the information will be brought out by the overburdened child welfare agency, and the child's position, if prepared in a thorough and aggressive manner, will be given undue weight. Even if both the parental representatives and the public agency prepare fully, the child's attorney can be seen as a third party whose decision lends determinative weight to the position of one or the other party. This can be problematic when the child's attorney follows her client's wishes instead of her own opinion about what is in the child's best interest, and the other parties do not fully develop the facts and present dangers which lead the attorney to reach that opinion.

Other commentators see the role of a child advocate, even an appointed attorney, allowing representation of a child's best interest regardless of the child's expressed interest.[150] Under the best interest approach, possible incompetence, underlying substantive law, and the nature of the proceedings are persuasive in releasing the lawyer from her traditional role and taking a subjective "best interest" approach.

---

145. *See id.* Guggenheim advocates a traditional lawyer-client relationship for unimpaired children in any legal context. For impaired children, the underlying substantive law should inform the lawyer's representation.

146. *See* Ramsey, *supra* note 137, at 312 (suggesting a presumptive age of seven).

147. *See Proceedings of the Conference on Ethical Issues in the Legal Representation of Children: Report of the Working Group on Determining the Child's Capacity to Make Decisions*, 64 FORDHAM L. REV. 1339 (1996).

148. *See* Margulies, *supra* note 143, at 1473.

149. *See* Ramsey, *supra* note 137, at 309-20 (suggesting that a determination of capacity can be tied to the risk presented by the position: if the risk is high, a higher degree of capacity can be required before the lawyer must be bound by the client's position. Ramsey acknowledges the subjectivity inherent in this position).

150. *See* Peters, *supra* note 139 (discussing the best interest model).

The law governing child protection in abuse and neglect proceedings is unclear in establishing a preference for empowerment or protection even in the exercise of the basic right to counsel. This tension is even present for older children who have the strongest claim on empowerment rights and clearest capacity to inform their attorney of their wishes. For such children, empowering them is the key to protecting them as foster children and when they leave foster care. They should be the source of information and guidance both to their attorneys and to the agency workers planning their protective care.

PART II. DOMESTIC VIOLENCE, CHILD PROTECTION AND THE TENSION BETWEEN EMPOWERMENT AND PROTECTION

A. Balancing Empowerment and Protection of Women - History of the Legal Response to Domestic Violence

The legal response to domestic violence in the United States must begin with an understanding that beating one's wife, like beating one's child, was not only ignored but was condoned by law. The famous "rule of thumb" was part of an accepted Law of Chastisement in English Common Law which allowed wife beating by a husband to keep his wife from acting in unacceptable or illegal ways at a time when the husband and wife were considered one person (the man) and he would be liable for her offenses. In early Colonial times, the Bodie of Liberties prohibited wife assault of any form in Massachusetts in an effort to create civility and order in the New World; but, there is no evidence that this, or its accompanying proscription against unnaturally severe corporal punishment of children, were ever strictly enforced.[151] Until the mid-1800's criminal prosecutions, discussed *supra* Part I.C.2, of child abuse by fathers did not trigger state intervention. Similarly, an acceptance of at least moderate chastisement of wives was prevalent, either under the guise that beatings were within the husband's prerogative or that, even if beyond his reasonable prerogative, legal

---

151. The Bodie of Liberties of 1641 states, "Number 80. Everie marryed woeman shall be free from bodilie correction or stripes by her husband, unlesse it be in his owne defence upon her assalt." and "Number 83. If any parents shall wilfullie and unreasonably deny any childe timely or convenient mariage, or shall exercise any unnaturall severitie towards them, such childeren shall have free libertie to complaine to Authoritie for redresse." Bodie of Liberties of 1641 as contained in The Colonial Laws of Massachusetts (reprinted from the Edition of 1672 with the Supplements through 1686) at 51. For a discussion of this early effort, see PLECK, *supra* note 47 at 4. A second attempt to address domestic violence was made in the progressive era, but was not sustained. *See id.*; *see also* LINDA GORDON, HEROES OF THEIR OWN LIVES: THE POLITICS AND HISTORY OF FAMILY VIOLENCE (1988).

intervention would improperly intervene into the private domestic sphere.

In 1824, in the oft-quoted case of *Bradley v. State*, the Supreme Court of Mississippi ruled that:

> Family broils and dissentions cannot be investigated before the tribunals of the country, without casting a shade over the character of those who are unfortunately engaged in the controversy. To screen from public reproach those who may be thus unhappily situated, let the husband be permitted to exercise the right of moderate chastisement, in cases of great emergency, and use salutary restraints in every case of misbehaviour, without being subjected to vexatious prosecutions, resulting in the mutual discredit and shame of all parties concerned.[152]

It was not until the late 1960's, as part of the feminist movement sweeping the country, that efforts to change attitudes about domestic violence and to protect women from its harms began and grew.[153] Until this time, there was no legal response to the systemic problem of domestic violence in the United States.[154] Unlike the concurrent response to child abuse which was developed in each state by law in the 1960's and then by federal legislation in 1974, the initial activity to reform the response to domestic violence was a grass-roots, private effort. Perhaps due to distrust of government prevalent at the time, and to the broader understanding of domestic violence as a form of gender subordination among feminist leaders,[155] the movement was carried on outside of government with the establishment of shelters, public education and support groups for women.[156] These grass-roots interventions were premised on the idea that women could be protected by empowering them – the theories of protection and empowerment were not at odds but were instead complimentary.

One of the first battered women's shelter, Chiswick Women's Aid, was started in England in 1971. United States activists visited the shelter, and subsequently, in the 1970's, shelters began opening in the United States. State legislation providing for civil and criminal remedies followed. During the 1970's and 80's, legal advances were hindered by lingering prejudices amongst judges, prosecutors and police which often frustrated the

---

152. Bradley v. State, 1 Miss. 156 (1 Walker) (Miss. 1824).

153. There were earlier efforts to address family violence but these were not sustained and did not result in legal responses. *See* PLECK, *supra* note 47, at 4.

154. *See* CLARE DALTON & ELIZABETH M. SCHNEIDER, BATTERED WOMEN AND THE LAW 1 (2001).

155. *See generally* ELIZABETH M. SCHNEIDER, BATTERED WOMEN AND FEMINIST LAWMAKING (2000); SUSAN SCHECHTER, WOMEN AND MALE VIOLENCE: THE VISIONS AND STRUGGLES OF THE BATTERED WOMEN'S MOVEMENT (1982).

156. *See* Deborah Epstein, *Redefining the State's Response to Domestic Violence: Past Victories and Future Challenges*, 1 GEO. J. GENDER & L. 127, 128 (1999).

efforts of women to free themselves from abusive relationships.

While the federal government enacted CAPTA in 1974, it took another twenty years before major federal law addressing domestic violence was passed in 1994. The Violence Against Women Act of 1994 (VAWA), part of the Violent Crime Control and Law Enforcement Act of 1994,[157] created federal enforcement mechanisms to augment state laws. Interstate domestic violence[158] and interstate stalking[159] became federal offenses along with interstate violation of state protection orders.[160] A host of firearm offenses were also made subject to federal jurisdiction.[161] States are now required to give full faith and credit to protective orders from other states.[162]

VAWA also included a civil rights provision, declaring by federal law that crimes of violence motivated by gender violated a victim's civil rights. The provision provided relief in the form of compensatory damages, punitive damages, and injunctive relief.[163] The civil rights provision was declared unconstitutional in *United States v. Morrison*.[164] The other provisions of VAWA are still in effect and give relief for interstate as well as gun-related violence.

B. Current Domestic Violence System and Coordination with the Child Protection System

1. Statistics

Currently, the institutional responses to family violence are bifurcated between two systems: the child protective services system and the domestic violence system. Each of these systems is premised on a clearly identified victim: the innocent child in need of protection in the child protective services system and the battered wife in the domestic violence system. For those families who do not neatly fit into the paradigm of single victim image of one system, access to legal or social service assistance to escape the violence is complicated, sometimes complimentary but often contradictory.

Increasingly, there is recognition that family violence does not fit neatly into these two categories and that different forms of family violence occur

---

157. *See* Pub. L. No. 103-322, 108 Stat., 1796 (codified as amended in 1996, 18 U.S.C. § 2261).

158. *See* 18 U.S.C. § 2261 (1994 & Supp. V 2000).

159. *See* 18 U.S.C. § 2261A (Supp. V 2000).

160. *See* 18 U.S.C. § 2262 (1994 & Supp. V 2000).

161. *See* 18 U.S.C. § 922 (1994 & Supp. V 2000).

162. *See* 18 U.S.C. § 2265 (1994 & Supp. V 2000).

163. *See* 42 U.S.C. § 13981 (1994).

164. 529 U.S. 598 (2000).

concurrently and need to be addressed simultaneously. Because of the high comorbidity between domestic violence and child abuse, there is a fledgling effort to coordinate institutional responses to violence against women and children.[165] Tragically, this coordination has resulted in new system-imposed harms to the women and children it seeks to protect.[166] In the case study of Rosa, her mother sought non-legal assistance in the form of shelter to escape her abusive husband. Ten months later, the child welfare agency removed Rosa from her mother's care and placed her in foster care. The allegations prompting the removal were based on Rosa witnessing verbal abuse in her home. By her mother reaching out for support to protect herself and Rosa, she was pitted against the child protection agency. We are not told Rosa's age or wishes in this case. Assuming she is either too young to articulate a preference or wants to be with her mother, there is no evidence that empowerment of her mother and protection of Rosa could not be complimentary instead of contradictory goals.

Studies to date have begun to illuminate the interrelationships between different forms of family violence.[167] Feminists must play an integral part in

---

165. *See, e.g.*, NAT'L COUNCIL OF JUVENILE AND FAMILY COURT JUDGES, MODEL CODE ON DOMESTIC AND FAMILY VIOLENCE (1994); AMERICAN BAR ASSOCIATION, THE LEGAL IMPACT OF DOMESTIC VIOLENCE ON CHILDREN (1994).

166. For a discussion of early recognition of this problem, see NAT'L CENTER ON WOMEN AND FAMILY LAW, *supra* note 69 (including position paper, Grossier-Keller, Battered Women and Their Battered Children: Criminal and Civil Allegations of the Women's Failure to Protect).

167. *See* Lee H. Bowker et al., *On the Relationship Between Wife Beating and Child Abuse*, *in* FEMINIST PERSPECTIVES ON WIFE ABUSE 158 (Kersti Yllö & Michele Bograd ed. 1988). As late as 1988, Bowker, Arbitell and McFerron wrote that empirical data illuminating possible links between forms of family violence was sparse, but we held assumptions that the links exist. The authors review the scant literature on the correlation between child abuse and wife abuse. In their own study, the authors find that child abuse accompanied wife abuse in seventy percent of the studied cases where children were in the home. *See id.* at 162.

In *A Guide to Research on Family Violence*, the Urban Institute and National Council on Juvenile and Family Court Judges reports that, "[t]he risk of child abuse is significantly higher when partner assault is also reported." THE URBAN INSTITUTE AND THE NATIONAL CENTER ON JUVENILE AND FAMILY COURT JUDGES, A GUIDE TO RESEARCH ON FAMILY VIOLENCE 27 (1993) (citing Gerald T. Hotaling et al., *Intrafamily Violence and Crime and Violence Outside the Family*, *in* FAMILY VIOLENCE 315 (Lloyd Ohlin & Michael Tonry ed. 1989)). After this generally accepted statement, the studies vary greatly in their numbers, but all show a high correlation between partner assault and child abuse. *See, e.g.*, LENORE E. WALKER, TERRIFYING LOVE: WHY BATTERED WOMEN KILL AND HOW SOCIETY RESPONDS (Harper & Row 1989) (nearly half of men who abuse their female partners also abuse their children); MURRAY A. STRAUS ET AL., BEHIND CLOSED DOORS: VIOLENCE IN THE AMERICAN FAMILY (Anchor Books 1980) (nationally, seventy-five percent of battered women say that their children are also battered); Jean Giles-Sims, *A Longitudinal Study of Battered Wives*, 34

gathering the statistics of family violence to ensure an accurate representation of the experiences of women and children[168] in households where violence is present.[169]

Existing statistics detailing the violence within families are overwhelming. Between two million and four million women are victims of family violence annually.[170] Violence among adult intimate partners, often called, "domestic violence," is almost exclusively perpetrated by men against women. In the National Crime Survey, ninety-one percent of spousal assaults were by men against their wives or ex-wives.[171]

FAM. REL. 205 (1985) (one-third of families reporting a violent incident between the parents also reported the presence of child abuse).

168. In 1991, the National Commission on Children wrote,

> [A] recent analysis of the factors that place children at risk of maltreatment suggest that only family income is consistently related to all categories of abuse and neglect. When other factors, such as single parenthood and race, are controlled for income, there is no positive correlation with heightened risk of abuse and neglect. In fact, this analysis suggests that when the same resources are available to families headed by single mothers as to two-parent households, children are actually at lower risk of maltreatment.

NAT'L COMMISSION ON CHILDREN, BEYOND RHETORIC, A NEW AMERICAN AGENDA FOR CHILDREN AND FAMILIES, FINAL REPORT OF THE NATIONAL COMMISSION ON CHILDREN 284 (1991) (citing Andrea Sedlak, *National Prevalence of Child Abuse and Neglect*, paper presented for the Conference on Child Welfare Reform Experiments, American Expertise Institute, Washington, D.C. (1991)).

169. Katharine Bartlett describes the importance of "asking the woman question," in describing feminist legal methods. *See* Katharine T. Bartlett, *Feminist Legal Methods*, 103 HARV. L. REV. 829 (1990). Given the overwhelming overrepresentation of women as victims of domestic violence, gathering statistics with a woman's perspective is crucial. For instance, evaluation of success in the domestic violence system often focuses on whether or not a woman complainant appears for all proceedings. Prosecution achieving a guilty finding or plea is the ultimate goal. In changing the terms of the debate, feminists need to collect data on whether women who access the system believe the system has been successful for them even if they chose not to participate in the full formal legal process. *See generally*, STATE UNIVERSITY OF NEW YORK AT BUFFALO SCHOOL OF LAW FAMILY VIOLENCE CLINIC, MONROE COUNTY FAMILY COURT DOMESTIC VIOLENCE INTENSIVE INTERVENTION COURT EVALUATION (2000) (evaluation included in person interviews with 143 women seeking protective orders and follow up contacts three months later). As aptly put by Colleen McGrath of the New York State Office for the Prevention of Domestic Violence, "A DA's witness who failed to show up may in fact be a savvy consumer of the system." Interview with Colleen McGrath, New York State Office for the Prevention of Domestic Violence, in Albany, N.Y. (Nov. 21, 1994).

170. *See* Antonia C. Novello, *From the Surgeon General, U.S. Public Health Service*, 23 JAMA 267 (June 17, 1992).

171. *See* Patsy Klaus & Michael Rand, U.S. DEPARTMENT OF JUSTICE STATISTICS, FAMILY VIOLENCE: BUREAU OF JUSTICE STATISTICS SPECIAL REPORT (April 1984).

It is commonly estimated that the co-morbidity of domestic abuse and child abuse is 30% to 60%.[172] In one survey of battered women, 87% of the women who lived with children reported that the children knew about the battering.[173] Living within a dynamic of violence makes the children indirect victims. Often, the children are the direct targets of the violence. Researchers estimate that there are 810,000 families in the U.S. where there is concurrent spouse abuse and child abuse.[174] It is also estimated that 3.3 million to 10 million children witness domestic violence annually.[175]

2. Practice

Once a child is removed, or if the family refuses voluntary services, a system of procedural requirements comes into play to provide periodic hearings on the parent's rehabilitation, the agency's efforts and the child's safety. The judge in a child protective services case rarely considers issues of child support, joint custody or adult violence in formulating orders. The ultimate "protective" tool of the child protective service system is removal of the children from the home. In extreme cases, this can lead to termination of parental rights, an act which frees a child for adoption.

While inquiries into violence against the mother in the home may be part of a risk assessment at the outset of a case, it is not the job of the child protective services system to protect the adults in the home. If violence threatens children, the child protective services system has within its power to remove the children, and will not usually take steps to remove the perpetrator. At best, a referral may be given to the mother to tell her how to remove the perpetrator through the domestic violence system.[176] Her swiftness and

---

172. *See* CHILD. DEF. FUND, *supra* note 13 at 86 (citing Jeffrey L. Edelson, *The Overlap Between Child Maltreatment and Woman Battering*, 5 VIOLENCE AGAINST WOMEN 134-54 (Feb. 1999)).

173. Lenore Walker reported results from a sample of battered women asked about concurrent child abuse. Fifty-three percent (53%) of the women reported that the batterer abused the children. Twenty-eight percent (28%) reported that they themselves abused the children. *See* LENORE WALKER, THE BATTERED WOMAN SYNDROME 188 (1984).

174. *See* NAT'L COUNCIL OF JUVENILE AND FAMILY COURT JUDGES, FAMILY VIOLENCE: IMPROVING COURT PRACTICE iii (1990) (citing MARIA ROY, CHILDREN IN THE CROSSFIRE: VIOLENCE IN THE HOME—HOW DOES IT AFFECT OUR CHILDREN? (1988)).

175. *See* CHILD. DEF. FUND, *supra* note 13, at 86.

176. Removal via divorce or separation carries pitfalls. The Model Code on Family Violence recommends a rebuttable presumption that it is not in the best interest of a child to award sole, joint legal or joint physical custody to a perpetrator of domes-

success in doing so may determine how she is judged as a parent. She will probably also be given a variety of tasks to work toward family reunification. Often, she is required to maintain a stable home and income. This may make it difficult for her to remove a batterer who is the family's source of income from the home.[177] The domestic violence system is usually triggered by an adult victim who seeks the assistance of the police or the courts to stop an act or acts of violence against her. If she alleges sufficient facts, she can receive a civil or criminal protective order from the court to order him to cease the abuse and perhaps to keep him away from her for the period of time detailed in the order. A criminal order requires the arrest of the alleged perpetrator. Most states also have procedures in place to allow her to petition for attorney's fees, temporary support, and child custody on a short term emergency basis, but such orders are not always easy to receive, especially when there are existing orders from the domestic relations branch of the court.

Even in localities where police are mandated to make an arrest when there exists probable cause that an act of violence has occurred, release is almost immediate. Maintenance of a violence-free home depends on the woman's ability to enforce the protective order and on the ability and willingness of law enforcement to respond to violations and hold the perpetrator accountable to the terms of the order.

Because of statistics drawing parallels between violence against women and children, efforts to coordinate the responses between the two systems which respond to each form of violence have begun. Tragically, the results of this attempted coordination have frequently been to blame the woman victim of domestic violence for failing to protect her children from the violence. Instead of illuminating the dynamic of violence, stopping the perpe-

---

tic violence. The Model Code also recommends that visitation by a parent who committed domestic violence only be permitted if the safety of the child and parent victim can be protected with a list of suggested safeguards. *See* NAT'L COUNCIL OF JUVENILE AND FAMILY COURT JUDGES, FAMILY VIOLENCE: MODEL STATE CODE 33, 34 (1994). The "friendly parent" doctrine in child custody cases also has the unintended consequence of rewarding the abuser and frustrating the mother's efforts to separate herself from a batterer. Under that doctrine, the parent most likely to facilitate the relationship between the child and the other parent is given custody. In a domestic violence situation, the woman's arguments to decrease a batterer's access to herself and their child may work against her and give him custody. *See e.g.*, Karen Czapansky, *Domestic Violence, the Family, and the Lawyering Process: Lessons from Studies on Gender Bias in the Courts*, 27 FAM. L. Q. 247, 256 (1993).

177. Many of the interviewed women at the State University of New York at Buffalo School of Law's Family Violence Clinic discussed the financial difficulties which caused them to stay with the batterer, or suffer financial hardship after ending the relationship. *See* STATE UNIVERSITY OF NEW YORK AT BUFFALO SCHOOL OF LAW FAMILY VIOLENCE CLINIC, *supra* note 169, at 35.

trator from committing violent acts, and supporting the child's non-offending caregiver, the coordination efforts have only made law enforcement and child protective services workers more conscious of the risk of concurrent incidents of violence. Instead of stopping the violent perpetrator, the institutional response has often been to separate the mother caretakers from their children. While violence by the same perpetrator cannot occur simultaneously if the mother and children are in different places, the harm of separation and the failure to address the dynamic which causes the violence may result in more tragedy.

Without a unifying theory to conceptualize the complexity and co-morbidity of the violence, the coordination efforts have employed easy solutions with tragic consequences. I highlight two such "solutions" here. The first is prosecutions[178] of mothers for failing to protect their children from family violence when the mother has never herself abused the children and is a victim at the hands of the same abuser. The second is the loss of custody by the parents to the child protective services agency ensuing, in part, from the portrayal of the mother as a helpless victim.

Cases have begun to appear in many jurisdictions prosecuting women who are in a relationship with a violent man with civil or criminal failure to protect her children.[179] In some cases, this response has brought victims of violence before the criminal courts due to the actions of their abusers and deprived children of the vital caretaker in their lives. In others, the mother's perceived "inaction" is viewed as failure to protect a child from witnessing abuse or being concurrently abused. Her "action" in calling authorities and seeking shelter or protective orders can alert child welfare authorities to the presence of domestic violence in the home and lead to a civil child protection proceeding. Fear of losing children can prevent caretakers from seeking the help they need to protect children.

Protection via removal should only be an extreme remedy necessary when efforts at supporting the non-offending parent have failed or when the domestic violence victim is herself abusive and needs services to rid violence from the home. Even in such extreme cases, removal should be thought of as temporary, and aggressive reunification efforts need to be provided and repeated as women seek to break out of the cycle of violence and control. Children and mothers must be empowered in developing safety plans that draw on their resources with the support of domestic violence and

178. The term "prosecution" is used to describe criminal cases brought against women for failure to protect. There is an even larger body of case law in the civil child protective services system using "failure to protect" as grounds for an adjudication and possible removal of the children. *See, e.g.*, NAT'L CENTER ON WOMEN AND FAMILY LAW, *supra* note 69.

179. *See id.*

child welfare systems.

Feminist theories provide a solid framework to examine all forms of family violence. While not all family violence follows strict gender-based identification of male perpetrator and female victim, that gendered paradigm does in fact continue to describe a disproportionate amount of family violence. Gender is a central factor in understanding family violence.

The feminist theory of power and control leading to escalating violence, without its gendered underpinning, broadens the theory to encompass all victims and perpetrators. The richness of the differences among women in violent homes is also a vital component in conceptualizing and addressing family violence. Together, gender, power and control dynamics and an appreciation of non-essentialism within feminist legal theory provide a framework for understanding all forms of family violence.

From this framework, a new institutional response can be imagined. Two scenarios are illustrative:

> Police respond to a call for assistance by Rosa's mother for a "domestic dispute." When they arrive, all is calm but the mother seems very nervous and the father stands between her and the police. Rosa is in the other room but appears to be in no harm. The police ask the mother, "Why did you call?" When she tells them there is no problem and asks them to please leave, they warn her that she cannot abuse the 911 service. They further explain that if they come back again, they will think about reporting the family to child protective services to let them investigate the mess.

Imagine instead an approach to family violence which recognizes the centrality of gender in our society and assumes a dynamic of power and control which can be asserted silently, without violence, whenever the perpetrator makes his authority clear. The system would be able to deliver services responsive to the needs of family members who sought assistance. Empowerment would be a means to protection.

Well-trained police could view the role of their response as breaking the cycle of power and control, thus limiting the abuser's authority. Their job performance would not be judged based on the number of arrests they made or successful prosecutions from their arrests but rather on their ability to protect known victims and stop known abusers from becoming violent again with the same or a new victim. If a family member failed to participate in the subsequent stages of the intervention, that would reflect poorly on the system not on the family member. Why did they decide their life was better without the intervention, and how can the system be tailored to be more responsive?

Another example similarly illustrates the danger of a poorly coordinated response:

> The child protective services system is investigating a call from a school reporting bruises on Rosa which she says were caused by her mother hitting her with a wooden spoon. When the child protective services investigator speaks to the mother, she is wearing sunglasses in an attempt to hide a black eye. She appears to be on drugs. The child protective services investigator warns her that Rosa can be removed if she does not stop the violence in her home and get her act together. She denies any violence in the home and says Rosa was in a fight after school with another child. The case is unfounded by the child protective services system. Fearful that her life is out of control, the woman calls the domestic violence hotline and asks for a referral for counseling. She is told that mental health professionals are mandated reporters and she is at risk of losing her children to child protective services. She decides not to seek help.

If the child protective services investigator worked within an integrated family violence services system which responded to all forms of family violence, the investigator may have been able to offer services to the mother, not as a victim or perpetrator, but as someone in need of help. Removal of children should not be a threat or punishment but only a drastic step, sometimes necessary, to secure the child's safety while the caretakers get effective help. With an immediate goal of keeping the child with a family caretaker, assessments must be made to define the "family" to be preserved. Again with a reference to the dynamic of power and control, can removal of a violent adult allow a child to remain safely with a parent? If that is not possible in the short term, what services can make it possible?

In Rosa's case, her mother's efforts to separate from the violence and seek shelter should be supported. Working with her and Rosa, if Rosa is old enough to communicate her wishes, what services and supports do they need to remain safely together? The child protection system and domestic violence system must dovetail efforts to protect children by empowering them and their caretakers.

PART III. EMPOWERING AND PROTECTING OLDER CHILDREN

A. Necessary But Meager Protection: Background to the Foster Care Independence Act (FCIA)[180]

At age eighteen or twenty-one, children have spent their right to protection and are considered autonomous adults, although few are prepared for

---

180. For a discussion of the FCIA and the needs of older foster children, see Mangold, *supra* note 6 at 835 (summarized in part in this section).

such empowerment. For many older children, relationships with birth families are long-lasting if not strong. Adoptive families are hard to locate. Therefore, permanency, through termination of parental rights and then adoption, is not a viable alternative.[181] For those who remain in foster care because their families cannot safely care for them, the failure to successfully reunify or to identify an adoptive family may mean that the children leave foster care with no caretakers. For them, the goal of "independent living" becomes the fall-back permanency option. Faced with this, the public agency with custody of the child in foster care attempts to prepare the aging-out foster child with the skills and resources necessary to succeed on their own. Instead of retaining protection of the child while increasing empowerment, efforts collapse into preparing youth for independent living with no parental guidance and only limited state assistance available.

Revisions in planning requirements specifically addressed the needs of older children in foster care in the mid-1980's and the need to wean them from the protection of the state toward independence through empowerment. In 1986, the Independent Living Initiative (ILI) was passed.[182] For children sixteen years and over, the ILI requires specific planning to help these older children before they age out of the foster care system. The provisions of the ILI provide funding for states to assist children who were currently, or had in the past, received foster care maintenance payments or had been in foster care. The goal of the ILI is "to help the individuals participating . . . to prepare to live independently upon leaving foster care."[183] The programs suggested by the federal legislation for states to provide to older foster children included programs to "enable participants to seek a high school diploma[,] . . . vocational training[,] . . . provide training in daily living skills[and] . . . provide each participant a written transitional independent living plan which shall be based on an assessment of his needs."[184]

The Foster Care Independence Act of 1999[185] attempts to improve inde-

---

181. States have standards for termination of parental rights, which usually include abandonment, permanent incapacity to parent and failure to remedy the circumstances which placed a child into foster care. Federal law now requires that petitions be filed if a child has been in placement for fifteen of the past twenty-two months, but if contested, states will still need to meet their statutory grounds for termination. The federal law includes exclusions from the termination requirement which are enumerated in 42 U.S.C. § 675(5)(E) (1994 & Supp. IV 1998).

182. Pub. L. No. 99-272, Title XII, § 12307(a), 100 Stat. 294 (codified at 42 U.S.C. § 677).

183. 42 U.S.C. § 677(d) (1991 & 1997 Supp.).

184. *Id.*

185. The Foster Care Independence Act of 1999, Pub. L. No.106-169, 113 Stat. 1822 (codified as amended in scattered sections of 42 U.S.C.).

pendent living services. The plight of older foster children, and of those newly empowered into independence, was detailed in the legislative hearings for the Foster Care Independence Act. Unfortunately, many young adults who age out of foster care do not successfully make the transition to independent living by the time they age out of care. Too many reappear in shelters, jails or on the streets. The needs of aging-out foster children were at least partially recognized by Congress in passing the Foster Care Independence Act of 1999. This law provides funding for services to assist youths up to age twenty-one who are aging-out of the foster care system into independent living. This law provides reimbursement to the state for services for a few more years, but does not provide a creative solution to address the need for ongoing protection. It does not transgress an arbitrary birthday deadline to allow ongoing support for youths who still need state protection.

In passing the law, Congress explicitly recognized that the overhauls of the 1997 Adoption and Safe Families Act were not sufficient in ensuring permanent homes for all foster children. In debate, Rep. Pryce remarked:

> In 1997, Congress tried to help these children by passing legislation to facilitate the adoption of children in foster care. As a result, the dream of a permanent family and a loving home is becoming a reality for more and more children. Yet despite our best efforts to streamline the system and find willing families to adopt these kids, the reality is that there are thousands of children who will never leave the foster care system during their childhood.[186]

Rep. Camp expressed similar understandings when he stated:

> I was very proud to be a part of our efforts to revamp the Foster Care system when this House passed the Adoption and Safe Families Act two years ago. And our efforts are paying off — preliminary numbers show that adoption of foster children have [sic] increased 40 percent since 1995.
>
> But this bill takes the next step—it recognizes that no matter how hard we work, some kids will turn 18 in foster care. They'll "age out" of the foster care system without a network of family and loved ones to turn to.[187]

Despite widespread recognition of the need for "family and loved ones to turn to,"[188] the Foster Care Independence Act only provides for increased

---

186. 145 CONG. REC. 92, H4958 (daily ed. June 25, 1999) (statement of Rep. Pryce).

187. *Id.* at H4971 (statement of Rep. Camp).

188. Many members of Congress echoed this plea to address the need for caring parent-like adults for older foster children. "Those of us who have teenagers know that when the child becomes 18, they still need the guidance, the support, the direction of parents." *Id.* at H4968 (statement of Rep. Lofgren) (debating Foster Care Independence Act of 1999 (H.R. 1802) (106th Congress, 1st Session)). Rep. Pryce stated, "[a]s parents, we do not cut off our children once they turn 18, although I think it is safe to

funds for enhanced programming by the states and not for family supports for children who need them.

Many children still exit foster care with no home, adult assistance or real promise of a future. There are approximately 500,000 children in out-of-home care on any given day.[189] Nearly one-third of the children entering care in any year or in care during that year are teens.[190] Of the children leaving care, almost 40% are teenagers.[191] Only eleven percent of these are in care due to status offenses such as truancy or ungovernability or delinquency.[192] The remainder are in care mainly due to protective services, parental condition or absence, or relinquishment of parental rights.[193] Data is available from the federal government from twenty-four states on the outcomes for children of all ages exiting out-of-home care. In the six months from April 1, 1996 through September 30, 1996, sixty-three percent who

---

say that even if we did, our children would have a better chance at survival than the products of the foster care system." *Id.* at H4958; Rep. Rangel remarked, "[m]ost all of us know as parents that a child becoming 18 does not necessarily mean that they are ready to assume the responsibility of adulthood." *Id.* at H4960; Rep. Greenwood added, "[w]hen we think of ourselves as parents, how many of us with our children, who have the fortune to have had good, stable upbringings where they are loved, how many of us say, here is your 18th birthday card, hit the street? We do not do that." *Id.* at H4961; Rep. Cardin stated, "[h]ow many of us as parents tell our children at 18 that they are on their own? We have a responsibility." *Id.* at H4961; Rep. Foley remarked, "[w]e can all remember how hard growing up can be. Fortunately for most of us we had loving and supportive. . . family and parents to nurture, encourage, and teach us how to gradually enter adulthood. I could never imagine the feelings of fear or uncertainty that a foster care child approaching his or her 18th birthday must have." *Id.* at H4961; Rep. Eshoo said, "[f]or those of us with teenage children, we know that 18-year-olds aren't often prepared to live on their own, paying their own bills." *Id.* at H4964. In hearings before the House Committee on Ways and Means, Subcommittees on Human Resources, similar sentiments were expressed by William Pinto, Adolescent Services Coordinator, Department of Children and Families, Hartford, CT (March 9, 1999): "For most young people in America, leaving one's home to be on your own means voluntarily giving up the security of the family. You leave when emotionally and economically ready for independence. The move out coincides with a positive event, such as getting married or landing that first big job. When setting up the first apartment, Mom has saved silverware and dishes, Aunt Millie has that pull-out couch in the basement, and Dad may put a fresh coat of paint on the wall. Most importantly, underneath it all is the security of knowing that if it doesn't work out, you can always go back home. (And don't all the parents of young adults in this room know that they often end up back at our front door?)." *Id.* at H4965 (statement by William Pinto).

189. *See* 1998 GREEN BOOK, *supra* note 13, at 783 (noting that the most recent statistics in the 1998 Green Book are for 1995 and previous years).

190. *See id.* at 786.

191. *See id.*

192. *See id.* at 790.

193. *See id.*

left care were reunited with their families and eleven percent were adopted.[194]

In the Findings accompanying the Foster Care Independence Act, Congress estimated that 20,000 teens leave foster care each year "because they have reached the age of eighteen and are expected to support themselves."[195] The option of independent living is the fall-back option. For these children, an expectation of successful independence is uniquely problematic.

> Studies show that nearly 30 percent of these youths average nine years in foster care without a permanent living arrangement. As a result, within two years of leaving foster care, only half have completed high school, fewer still are employed, and nearly 60 percent of the young women have given birth, almost always outside of marriage.[196]

Studies of children who have aged out of foster care are scarce[197] but findings are consistently disturbing.[198] One study of all forty-six young men discharged from an independent living group home in New York City found that the overall mean length of time these young adults had spent in care before discharge was eight years.[199] Three studies[200] relied upon by gov-

---

194. *See id.* at 796.

195. Foster Care Independence Act of 1999, Pub. L. No. 106-169, 113 Stat. 1822 (codified at 42 U.S.C. § 677).

196. Nancy Johnson, *Foster Care Youth to Discuss Transition to Adulthood at Hearing Tomorrow*, GOVERNMENT PRESS RELEASES BY FEDERAL GOVERNMENT DOCUMENT CLEARING HOUSE, May 12, 1999, *available in* 1999 WL 2224152.

197. *See* Post-*Foster Care Independent Living Programs: Hearing Before the Subcomm. On Human Resources of the House Comm. On Ways and Means*, 106th Cong. (1999) *available in* 1999 WL 16947799 (statement of Cynthia Fagnoni, Director, Education, Workforce, and Income Security Issues, Health, Education and Human Services Division, U.S. General Accounting Office) [hereinafter Fagnoni Statement].

> Many foster youths have a difficult time making the transition from the foster care system to self-sufficiency. While there are few available studies tracking youths who have exited foster care, our review of these studies reveals some consistent findings. Research has shown that many former foster care youths have serious education deficiencies and rely on public assistance . . . . In addition, former foster care youths often find themselves lacking adequate housing.

Id.

198. For a literature summary on the topic, see Richard P. Barth & Marianne Berry, *Implications of Research on the Welfare of Children Under Permanency Planning, in* 1 CHILD WELFARE RESEARCH REV. 323, 345-48 (Richard Barth et al. eds., 1994).

199. *See* Mallon, *supra* note 22 at 65.

200. *See* Fagnoni Statement, *supra* note 197; *see also* Westat, Inc., *A National Evaluation of Title IV-E Foster Care Independent Living Programs for Youth* (Washington, D.C.: HHS, 1991), Richard P. Barth, *On Their Own: The Experiences of Youth After Foster Care*, 7 CHILD & ADOLESCENT SOC. WORK 419 (Oct. 1990), Mark E. Courtney and Irving Piliavin, *Foster Youth Transitions to Adulthood: Outcomes 12 to*

ernment sources report the tragic next chapter for children exiting foster care.[201]

As success in the current job market becomes even more tightly connected to educational attainment, children aging out of the foster care system are ill-equipped to compete. The Westat study found that forty-six percent of the 810 young adults studied who had left foster care had not completed high school.[202] That number was thirty-seven percent for Courtney and Piliavin who studied 113 former foster children and thirty-eight percent for Barth who studied fifty-five former foster care youths.[203] Employment success was equally dire. Westat, the largest study, found fifty-one percent unemployed two and a half to four years after leaving care.[204] Sixty-two percent had not maintained a job for at least one year.[205] Courtney and Piliavin found thirty-nine percent unemployed twelve to eighteen months after aging-out of the system.[206] Barth found twenty-five percent unemployed one to ten years after leaving foster care.[207]

Perhaps the two most disturbing correlations are between aging out of foster care and homelessness[208] and incarceration. Westat found that twenty-five percent of the former foster children had been homeless at least one

---

18 Months After Leaving Out-of-Home Care (Madison, Wisc.: University of Wisconsin, 1998). For an explanation of the findings of the Courtney study, see *Post-Foster Care Independent Living Programs: Hearing Before the Subcomm. On Human Resources of the House Comm. On Ways and Means*, 106th Cong. (1999) *available in* 1999 WL 16947799 (statement of Mark E. Courtney, Assistant Professor, School of Social Work and Institute for Research on Poverty, University of Wisconsin-Madison).

201. *See* Fagnoni Statement, *supra* note 197 (the government cites studies that show that foster children exiting care who have received independent living programming during their stay in foster care have better success than those who do not (citing Maria Scannapieco et al., *Independent Living Programs: Do They Make a Difference?* 12 CHILD & ADOLESCENT SOC. WORK J. 5 (Oct. 1995)).

202. *See* Fagnoni Statement, *supra* note 197.

203. *See id.*

204 *See id.*

205. *See id.*

206. *See id.*

207. *See id.*

208. The correlation between discharge from foster care and homelessness has long been recognized. *See The Adoption Assistance and Child Welfare Act and Related Proposals: Hearings on H.R. 2810 Before the Subcomm. on Public Assistance and Income Maintenance of the Comm. on Ways and Means*, 99th Cong., 1st Sess. 1-2 (1985) (statement of Leonard Bradley, Deputy Commissioner of Tennessee Dept. of Human Services on Child Welfare League of America) (state-by-state survey of state agencies asked what they believed happened to children who age out of foster care) (citing Mari Brita Maloney, *Out of the Home Onto the Street: Foster Children Discarded Into Independent Living*, 14 FORDHAM URB. L. J. 971, 988-89 (1985)).

night.[209] Courtney and Piliavin found that twelve percent were homeless at least once and that twenty percent of males and ten percent of females had been incarcerated at least once in the twelve to eighteen months since exiting care.[210] Barth recorded that thirty-five percent had been homeless or moved frequently and that thirty-five percent had spent time in jail or prison.[211]

Invisible to these counts are the children who have no stable home but live night-to-night on the generosity of friends or acquaintances. Called "couch surfers" in child welfare lingo, these teens are at risk for every frightening possibility of the street.[212] "The couch surfing phase is preliminary to being out in the street."[213] The numbers for teen pregnancy,214 inadequate health care and other indicators of maladjustment are similarly disturbing.

In introducing H.R. 3443 which later became the Foster Care Independence Act of 1999, Congress emphasized its concern for this vulnerable group of former foster children and made the following findings:

> (1) States are required to make reasonable efforts to find adoptive families for all children, including older children, for whom reunification with their biological family is not in the best interests of the child. However, some older children will continue to live in foster care. These children should be enrolled in an Independent Living program designed and conducted by State and local government to help prepare them for employment, postsecondary education, and successful management of adult responsibilities.
>
> (2) Older children who continue to be in foster care as adolescents may become eligible for Independent Living programs. These Independent Living programs are not an alternative to adoption for these children. Enrollment in

---

209. *See* Fagnoni Statement, *supra* note 197.

210. *See id.*

211. *See id.*

212. *See* Susan K. Livio, *Freedom Daunting for Ex-foster Child,* HOME NEWS TRIB., Jan. 3, 1999, at A1 (many former wards of foster care would be homeless "were it not for the generosity of friends and adults who have looked after [them] over the years. Social workers call such young people 'couch surfers,' for their habit of sleeping on a friend's couch until they wear out their welcome"); Susan K. Livio, *Death, Disappointment Stalk "Couch Surfers,"* THE COURIER-NEWS, Nov. 14, 1998, at A1.

213. Susan K. Livio, *Abusive Foster Care Left Teen Ill-Equipped for Freedom,* THE COURIER-NEWS, Dec. 13, 1998 *available in* http://www.injersey.com/news/story/0,1210,143770,00.html (last visited Dec. 14, 1998) (statement of Stephanie Schwartz of Crossroads, an outreach program for troubled youth in Burlington, County).

214. *See* Sharon G. Elstein, *Teenagers Are Adoptable: Strategies for Success,* 18 CHILD L. PRAC. 49 (June 1999). A 1991 study found that two to four years after leaving foster care, nearly 66% of teens were mothers compared to 25% in the general population. *See id.*

> Independent Living programs can occur concurrent with continued efforts to locate and achieve placement in adoptive families for older children in foster care.
>
> (3) About 20,000 adolescents leave the Nation's foster care system each year because they have reached 18 years of age and are expected to support themselves.
>
> (4) Congress has received extensive information that adolescents leaving foster care have significant difficulty making a successful transition to adulthood; this information shows that children aging out of foster care show high rates of homelessness, non-marital childbearing, poverty, and delinquent or criminal behavior; they are also frequently the target of crime and physical assaults.
>
> (5) The Nation's State and local governments, with financial support from the Federal Government, should offer an extensive program of education, training, employment, and financial support for young adults leaving foster care, with participation in such program beginning several years before high school graduation and continuing, as needed, until the young adults emancipated from foster care establish independence or reach 21 years of age.[215]

The purpose of the law was to enable states and localities to operate a spectrum of training programs to better prepare children aging out of foster care. The bill never included new or creative alternatives to create families or supportive homes for these children and thereby provide for ongoing protection. The purpose was described as follows:

> (a) PURPOSE.—The purpose of this section is to provide States with flexible funding that will enable programs to be designed and conducted -
>
> (1) to identify children who are likely to remain in foster care until 18 years of age and to help these children make the transition to self-sufficiency by providing services such as assistance in obtaining a high school diploma, career exploration, vocational training, job placement and retention, training in daily living skills, training in budgeting and financial management skills, substance abuse prevention, and preventive health activities (including smoking avoidance, nutrition education, and pregnancy prevention);
>
> (2) to help children who are likely to remain in foster care until 18 years of age receive the education, training and services necessary to obtain employment;

---

215. Foster Care Independence Act of 1999, Pub. L. No. 106-169, 113 Stat. 1822 (codified as amended in scattered sections of 42 U.S.C.). This Act was introduced January 6, 1999. S. 1327 has some additional information in the findings section (section 3) which goes on after the first sentence in H.R. 3443 to state, "In addition, approximately 5,000 adolescents (foster children over the age of 12) are adopted out of the foster care system each year, of whom approximately 620 are over the age of 16 at the time of their adoption. A large percentage of these children have not yet completed their high school education." S. 1327, 106th Cong. § 3 (1999).

> (3) to help children who are likely to remain in foster care until 18 years of age prepare for and enter postsecondary training and education institutions;
>
> (4) to provide personal and emotional support to children aging out of foster care, through mentors and the promotion of interactions with dedicated adults; and
>
> (5) to provide financial, housing, counseling, employment, education, and other appropriate support and services to former foster care recipients between 18 and 21 years of age to complement their own efforts to achieve self-sufficiency and to assure that program participants recognize and accept their personal responsibility for preparing for and then making the transition from adolescence to adulthood.[216]

Representative Johnson, co-author and lead co-sponsor of the legislation with Representative Cardin, summarized the goal of the legislation as "to prepare these young people to be able to move into the work force or to continue with their education on the very day they leave foster care. These children face very difficult problems and we must create programs to help them learn to be self-reliant."[217]

The intent of the Foster Care Independence Act was to empower teens and assist them into self-reliance. This is laudable and necessary but the evidence shows these youths need much more. They are uniquely and tragically ill-equipped for empowerment with no ongoing protection. Only up to thirty percent of any state's funding under this legislation can be used for room and board for children ages eighteen to twenty-one. These funds are valuable in helping to subsidize the former foster child's existence under a notion of protection, but the focus is on empowerment, despite the teens' documented need for protection.

Coupled with the Adoption and Safe Families Act, the Foster Care Independence Act seems to provide an increased, but still inadequate, spectrum of permanency options for older foster children. Children who are neither reunited nor adopted (but instead age out of the system with no identified

---

216. Foster Care Independence Act of 1999, Pub. L. No. 106-169, 113 Stat. 1822. This Act was passed by the House of Representatives on November 18, 1999 and by the Senate on November 19, 1999. It was signed into law by President Clinton on December 14, 1999. The earlier version debated on the floors of the House and Senate (H.R. 1802) contained substantially similar "finding" and "purpose" sections. H.R. 1802 was passed by the House in June. Both H.R. 3443 and H.R. 1802 were titled "The Foster Care Independence Act." Finding (a)(2) was added in H.R. 3443 and the final bill.

217. 145 CONG. REC. H4957-60 (daily ed. June 25, 1999) (statement of Rep. Johnson). Like many of the speakers, a litany of bad outcomes for the foster children was detailed by the speaker: "Today, two-thirds do not complete high school, 61 percent have no job experience, and 38 percent are diagnosed emotionally disturbed. Most end up jobless, addicted, pregnant, or in jail." *Id.*

caretaker) are left to independent living. If they live with, or receive support of some kind from relatives or other informal custodians, there is no public assistance for such accommodation.[218]

B. Creative Remedies to Protect Older Children by Empowering Them and Their Parents

1. Open Adoption

I use the term open adoption under the broadest definition, requiring neither a formal contract nor court order, but allowing for some ongoing contact to facilitate the surrender of rights and the finalization of the adoption.[219] Open adoption requires the termination of parental rights but allows for some continued contact between the birth family and the child. It could be as minimal as letters and pictures or could involve ongoing visitation. The term open adoption is used more generically to refer to any adoption in which confidentiality is compromised by opening records, exchanging information, visiting, etc. Here, the purpose in suggesting open adoption is to allow ongoing contact in whatever form will facilitate the finalization of the

---

218. Some localities use local money without state or federal reimbursement, and some states use exclusively state money without federal reimbursement to fund guardianships or other novel permanency plans. *See, e.g.*, Meryl Schwartz, *Reinventing Guardianship: Subsidized Guardianship, Foster Care, and Child Welfare*, 22 N.Y.U. REV. L. & SOC. CHANGE 441, 456-74 (1996) (discussing state subsidies for guardianship); Kathleen O. Byrne and Matilda T. Bellucci, *Subsidized Adoption: One County's Program*, 61 CHILD WELFARE 173, 173-74 (1982) (describing county subsidization for adoptions out of foster care before adoption subsidies were available under the Adoption Assistance and Child Welfare Act of 1980, Pub. L. No. 96-272, 94 Stat. 500).

219. *See* Carol Sanger, *Separating from Children*, 96 COLUM. L. REV. 375 (1996). Sanger uses such an inclusive working definition to describe both direct and indirect contact between the child, birth parents and adoptive parents. She makes the point that open adoption may "increase the number of children available." *See id.* at 492. Others have echoed the prospect, emphasized here for older children, that open adoption may provide stable familial relationships for some children who would not otherwise have parents who could care for them. Lawrence W. Cook, Note, *Open Adoption: Can Visitation with Natural Family Members Be in the Child's Best Interest?* 30 J. FAM. L. 471, 478 (1992). *But see* Carol Amadio and Stuart L. Deutsch, *Open Adoption: Allowing Adopted Children to "Stay in Touch" with Blood Relatives*, 22 J. FAM. L. 59, 60 (1983) (defining open adoption as an agreement in writing approved by the court). Some distinguish open adoption from cooperative adoption. *See, e.g.*, Annette Ruth Appell, *Blending Families Through Adoption: Implications for Collaborative Adoption Law and Practice*, 75 B.U. L. REV. 997, 1001-02 (1995). In cooperative adoption the ongoing contacts are arranged by mutual collaboration or agreement, a contract which is made between the birth parent(s) and the adoptive parents. *See id.*

adoption and, thus, some familial supports for older children in foster care. Open adoption can empower birth parents and older children by creating an acceptable adoptive arrangement. The adoption may be open by informal agreement, contract or court order.[220] Like traditional closed adoptions, open adoptions should be eligible for all forms of adoption assistance and subsidies, thus increasing the options available to provide familial protection for older children.

States only began to enact adoption statutes in the middle of the nineteenth century. In the twentieth century, legislatures amended the laws by providing confidentiality for the process, thereby creating the notion of closed adoptions.[221] While closed adoptions were the norm in the 1970's, some open adoptions were informally arranged and legally sanctioned at that time.[222] Only a few states legislatively endorse open adoptions.[223]

For all children in foster care with positive memories and bonds to their birth families, open adoption may provide a mechanism for permanency which does not require a total severance of ties with that family. This is especially relevant in cases of adoption by relatives where confidentiality is impossible.[224] For older children aging out of foster care, open adoption may provide a mechanism to allow the contacts that both the child and birth parent want or to allow for ongoing sibling contact.[225] Allowing these relationships to continue in some form may make the termination of parental rights more tenable for both parties. This could facilitate an adoption, which could not otherwise occur by assuring the parents that in surrendering their rights and freeing their child for adoption, they will not lose all contact with the child. This minimal empowerment may be what it takes to provide an adoptive family for a child in foster care.

---

220. *See* Appell, *supra* note 219, at 1011.

221. *See id.* at 1006.

222. Annette Baran et al., *Open Adoption*, 21 SOC. WORK 97, 98-99 (1976) (advocating that open adoption should be accepted as an alternative and discussing an "experiment" in which an author arranged an open adoption).

223. *See, e.g.*, ALASKA STAT. § 25.23.130(c) (Michie 2000); N.M. STAT. ANN. § 32A-5-35(A) (Michie 1999); N.Y. SOC. SERV. LAW § 383(c)(3)(b) (McKinney 1993); TENN. CODE ANN. § 36-1-121(f) (Supp. 2000); WASH. REV. CODE ANN. § 26.33.295(1) (West 1997). Other states have sanctioned open adoption through case law. *See, e.g.*, Adoption of Gwendolyn, 558 N.E. 2d 10 (Mass. App. Ct. 1990) (holding that post-adoption visitation permitted but at the discretion of adoptive parents); Michaud v. Wawruck, 551 A.2d 738 (Conn. 1988) (ruling that contract between birth and adoptive parents was not contrary to public policy).

224. *See* Appell, *supra* note 219, at 1011.

225. The importance of sibling bonds is one that is rarely paramount in the child welfare system. For a discussion of the importance of the bonds and their implications for advocacy, see Sharon G. Elstein, *Making Decisions About Siblings in the Child Welfare System*, 18 A.B.A. CHILD L. PRAC. 97 (Sept. 1999).

2. Guardianship

Guardianship is a judicially sanctioned arrangement; an adult can act as a guardian for a child for the duration of the court order. Guardianship was originally used to oversee a minor's property or estate when his/her parents were deceased or could not otherwise perform this function. Probate guardianships, as they are sometimes called, do not involve physical custody of the child, but are the antecedent to the current use of guardianships by the family or juvenile courts.[226] Probate guardianships, like the guardianships now ordered in the child welfare system, could be for a defined period of time or ongoing. Guardianship traditionally terminates no later than a child's eighteenth birthday. Just as foster care can continue to age twenty-one by continuation of the placement order, the court arranging the guardianship could have jurisdiction to continue the guardianship to age twenty-one or even beyond.

Today, guardianships are often provided by relative caretakers who have been kinship foster care resources. Guardianship provides for a more permanent status to the arrangement.[227] Guardianship also allows for the discharge of the child from the child welfare system in many instances, ending supervision by the agency and the court except for the negligible supervision that may be provided by the court overseeing the guardianship.[228] It is usually a single guardian or mother and father unit which are identified to provide guardianship. This does not require termination of parental rights and, in that sense, is less disempowering of parents. Like open adoption, it provides another option which older children and parents could choose and plan for with the child welfare agency. Often relatives prefer guardianship

---

226. *See, e.g.*, ROBERT D. GOLDSTEIN, CHILD ABUSE AND NEGLECT 767-77 (1999); Robert M. Gordon, *Drifting Through Byzantium: The Promise and Failure of the Adoption and Safe Families Act of 1997*, 83 MINN. L. REV. 637, 691 (1999); Bogart R. Leashore, *Demystifying Legal Guardianship: An Unexplored Option for Dependent Children*, 23 J. FAM. L. 391 (1984).

227. The Adoption and Safe Families Act, Pub. L. No. 105-89, 111 Stat. 2115 (1997) (codified as amended at 42 U.S.C. § 671 (1994 & Supp. 1998)). Section 101(b) of the Act added a definition of legal guardianship to the law. It is now codified at 42 U.S.C. § 675 (7) (1994 & Supp. 1998). While useful for clarification and uniformity of rights to be conferred onto guardians, no subsidization of guardianship is provided in federal law.

228. Professor Martha Fineman challenges us to separate state support from state supervision of families. For creative permanency planning for children who are over the age of eighteen, support and supervision must exist along a continuum to meet the individual needs of children and the adults sharing the parental responsibilities for them. *See* Martha Albertson Fineman, *What Place for Family Privacy?*, 67 GEO. WASH. L. REV. 1207 (1999).

to adoption because they do not have to endure a legal battle to terminate the parental rights of their sibling or other relative. Biological parents may be willing to consent to a guardianship by a sibling or other relative but not to adoption which severs their parental rights. Biological parents may participate in the selection of a guardian[229] and may enjoy ongoing visitation with the child. In families where property is available, guardianship also allows the child to inherit from the parent and does not release the parent from financial support.

Guardianships are frequently used by the child welfare system, but are not generally supported by federal subsidy as are adoptions.[230] Guardianships are not currently subsidized by the federal government except by waiver.[231] An end to state intervention and supervision of a well-functioning guardianship may be a permanency triumph; a custodial relationship moves from foster care to guardianship. Unfortunately, with the discharge from foster care comes an end to the subsidy which had been available for kinship foster care but is not available for guardianship. This loss of funding may prohibit some relatives from making the move to guardianship and eliminate this option for older children and their parents to exit the child welfare system.

A demonstration project in Illinois providing for subsidized guardianship estimates that 4,000 children will move out of kinship foster care and into more permanent and stable subsidized guardianships under the program.232 Other states allow for subsidization of guardians although their state funds are not reimbursed by the federal government.[233]

---

229. For terminally ill parents, stand-by guardianships or joint guardianships are used to plan for the child's care when the parent is no longer able to do so. *See, e.g.*, Sunny Rosenfeld, *Developments in Custody Options for HIV-Positive Parents*, 11 BERKELEY WOMEN'S L.J. 194 (1996). It is common practice for parents drafting a will to name guardians for their children. *See, e.g.*, Esther Appelberg, *The Significance of Personal Guardianship for Children in Casework*, 49 CHILD WELFARE 6 (1970) (advocating that social workers and caseworkers should encourage parents to draw up wills and name guardians for their children even when there is little or no property to pass on to the children).

230. The idea of subsidized guardianship has been circulated for years but never adopted by the federal government. *See, e.g.*, Marla Gottlieb Zwas, *Kinship Foster Care: A Relatively Permanent Solution*, 20 FORDHAM URB. L.J. 343 (1993); Leashore, *supra* note 224. For a discussion of subsidized guardianship as an exclusion from a declining subsidized reimbursement scale, see Gordon, *supra* note 227.

231. *See* Schwartz, *supra* note 218, at 456-74; Byrne & Bellucci, *supra* note 218.

232. *Adoption Promotion Act of 1997: Hearings on H.R. 867 Before the Subcomm. On Human Resources*, 105th Cong. 38 (1998) (statement of Jess McDonald, Director, Illinois Dep't. of Children and Family Services).

233. *See, e.g.*, ALASKA STAT. § 13.26.062 (Michie 2000); ARIZ. REV. STAT. ANN. § 8-814 (West Supp. 2000); CAL. WELF. & INST. CODE § 11405 (West 2001); MONT.

3. Mentorship

The word "mentor" has a Greek origin meaning steadfast and enduring.[234] In Homer's *Odyssey*, Odysseus leaves behind his friend, Mentor, to be the guide and educator of his son, Telemachus. Today, the term is used to describe a variety of adult-child supportive relationships[235] in which the adults offer themselves as short or long-term role models for children and offer some guidance to the child.[236] Such mentoring programs have existed in the U.S. at least since the turn of the century with the development of the Big Brother Program in 1902.[237] They are increasingly proposed and utilized [238] as a preventive measure against juvenile crime,[239] unemployment,[240] school

---

CODE ANN. § 41-3-421 (1999); N.M. STAT. ANN. § 32A-5-45 (Michie 1978 & Supp. 1999); W. VA. CODE ANN. § 49-2-17 (Michie 1999 & Supp. 2000). *See also* Timmons v. McMahon, 235 Cal. App. 3d 512, 286 Cal. Rptr. 620 (1991) (allowing payments to temporary, nonrelated guardians).

234. BREWER'S DICTIONARY OF PHRASE AND FABLE 721 (14th ed. 1989).

235. Mentoring in the one-on-one model may not be possible for all children. Alternatives are being explored with less intensive interaction. For a study of one such program in New York City, see Antronette K. Yancey, *Building Positive Self-Image in Adolescents in Foster Care: The Use of Role Models in an Interactive Group Approach*, 33 ADOLESCENCE 253 (1998).

236. Evaluations of such programs are scarce. For a review of a variety of programs, see Dionne J. Jones et al., *Reaffirming Young African American Males: Mentoring and Community Involvement by Fraternities and Other Groups*, 16 URB. LEAGUE REV. 9, 12 (1993).

237. *See* David Royse, *Mentoring High-Risk Minority Youth: Evaluation of the Brothers Project (Big Brother Agency)*, 33 ADOLESCENCE 145 (1998).

238. One positive example of a mentoring program is the NYC Independent Living Partnership where youths and their adult mentors meet semiannually for a weekend retreat, and monthly for an informal support group or network of support. *See* Mallon *supra* note 22, at 75. This model provides a network not only of adult mentors but also of former foster children who can support one another in the community. The PRIDE program (Personal and Racial/Ethnic Identity Development and Enhancement) uses successful mentors from the foster child's same cultural background to provide positive role models. This is not a one-on-one program, nor is it a one-day seminar. Instead, the program uses a series of group sessions to provide interaction between the foster child and mentor. It is not an intensive form of mentoring, but its effects have been positive. *But see* David L. DuBois & Helen A. Neville, *Youth Mentoring: Investigation of Relationship Characteristics and Perceived Benefits*, 25 J. OF COMMUNITY PSYCHOL. 227, 227-28 (defining mentoring as a "one-to-one relationship between a [youth] and a caring adult who assists the [youth] in meeting academic, social, career, or personal goals") (citing S.M. Nettles, *Community Contributions to School Outcomes of African-American Students*, 24 EDUC. & URB. SOC 'Y 139 (1991)). *See also* Yancey, *supra* note 235, at 253.

239. *See* Michelle Lea Cherne Anderson, *The High Juvenile Crime Rate: A Look at Mentoring as a Preventive Strategy*, 30 CRIM. L. BULL. 54 (1994).

240. *See* Patricia Rowe, *Volunteer Mentors Empower Inner-City Youths*, 19

dropout [241] and teen pregnancy.[242]

The term is used here to describe a supportive adult and to distinguish the mentors from those with guardianship or some other form of parental rights over the child. For children who do not want such adults in their lives or who do not have adults in their lives able or willing to enter into more formal parenting relationships of guardianship or adoption, mentors provide some adult supervision and support which could be extremely necessary at difficult times such as the transition out of foster care. The term usually refers to volunteer adults.[243] Youths receiving independent living services are often served in an aggregate setting with little one-to-one adult mentoring or care.[244] Resources for mentor programs are often scarce or non-existent despite the fact that "officials in all locations saw some type of mentoring program as one method to provide youths with a vocational role model and opportunities to practice other independent living skills they have learned."[245]

Volunteer opportunities should be encouraged; but, a more structured, federally subsidized form of mentoring where one or a variety of adults is paid to provide work, school or emotional support to a teen aging out of the foster care system is necessary. Subsidized mentors could insure that no child exits foster care without some committed adult support.

Such mentors could be identified by the public child welfare agency in the same way that foster parents are identified, either by direct contract with the public agency or through subcontracts with not-for-profit and new profit-making entities which provide a variety of services to the child welfare system.[246] Adults interested in assisting children but unable or unwilling to make the commitment as foster or adoptive parents could still play a

---

CHILD. TODAY 20 (1990) (describing an externship/mentoring program for Washington, D.C. youth to receive job training and experience along with career mentoring).

241. *See* Elaine A. Blechman, *Mentors for High-Risk Minority Youth: From Effective Communication to Bicultural Competence*, 21 J. CLINICAL CHILD PSYCHIATRY 160, 160 (1992).

242. *See* Jean E. Rhodes et al., *Natural Mentors: An Overlooked Resource in the Social Networks of Young, African American Mothers*, 20 AM. J. COMMUNITY PSYCHOL. 445 (1992).

243. *See* MARC FREEDMAN, THE KINDNESS OF STRANGERS xii (1993).

244. *See* Fagnoni statement, *supra* note 197.

245. *See id.*

246. The Guide Program for foster adolescents is operated by People Places of Charlottesville, Virginia. Teens in the Guide Program live with a responsible, trained adult for eighteen months before they age out of foster care. The adult serves as a "mentor, a sounding board, and a safety net for the teen." *Post-Foster Care Independent Living Programs: Hearing Before the Subcomm. On Human Resources of the House Comm. On Ways and Means*, 106th Cong. (1999) *available at* 1999 WL 16947802 (statement of Kelli Sutton Block, People Places of Charlottesville, Va.).

crucial role in the life of a teen. Whenever possible, the teens could identify mentors who might be willing to assist them in their transition to adulthood. Foster parents of siblings or foster parents of other teenagers are examples.

The realization that some children leave foster care without caring adults available to assist them challenges the very notion of permanency planning which, in some circles, means that all children who cannot be returned to their biological families are adoptable. [247] At least since 1986 with passage of the Independent Living Initiative, reality has overtaken theory to provide that "independent living" is a suitable goal for permanency planning. In other words, the goal for many children aging out of foster care is "independent living" because the child welfare system concludes that there is no family or permanent home for these teens. The system formulates a permanency plan with that realization in place, based on a right to empowerment triggered solely by age. Permanency planning needs to accommodate options that recognize that empowerment and protection are ongoing, complimentary goals.

By arguing that children need responsible adults supported by the state even after age eighteen, this article posits that the right to protection, so integral to the child welfare system, does not cease on a child's eighteenth birthday. Empowerment is important and teenage children should certainly be participants in their own permanency planning, but that does not require cessation of the exercise of a right to protection. As child welfare support via the Foster Care Independence Act can continue to age twenty-one, so may the suggested recommendations continue past age eighteen to at least twenty-one when the child is in agreement and the adult support is approved by the child welfare agency or court.

If properly funded, then open adoptions, guardianships and mentorships that accommodate empowerment and protection are potential and plausible arrangements for many children for whom reunification or traditional, closed adoptions are undesirable or unattainable. If the law could recognize parents as important but not exclusively so, the outcomes for older children could be healthier and more permanent. Instead of focusing on termination/adoption or independent living, older children would benefit from a system that does not disqualify any adult who wishes to be a positive presence, no matter how limitedly, in a child's life. Children aging out of foster care are a group that could significantly benefit from a cooperatively planned, shared custody relationship, developed with their input and facilitated and funded by the state.

---

247. *See* Katharine Miller et al., *Overcoming Barriers to Permanency Planning*, 63 CHILD WELFARE 45 (1984).

PART IV. CONCLUSION

The Conference, Transgressing Borders: Women's Bodies, Identities and Families, gave us all a rare opportunity to come together in memory of Mary Jo Frug and her work and speak openly about the foster care system and many other issues impacting the lives of women and families. I am hopeful that our presentations and conversations that day can continue and will influence the current debates on the foster care system. Throughout history, the foster care system in a variety of forms has been used to exercise the *parens partiae* power of the state to intervene into families and protect children. The protection of children has always been balanced against parental rights and empowerment. In this article, I try and shift the debate to consider protection and empowerment of both caretakers and older children as complimentary goals.

Dealing with domestic violence and its impact on children and the needs of older foster children are some of the many challenges facing the foster care system. Our current approaches are not working. We need renewed investment and creative permanency planning to address these complicated issues. Foster care is a vital resource but a temporary and radical one. The only way to possibly ensure lasting, safe and healthy plans to protect children is to empower caretakers and older youth.

Take the Child and Run: Tales From the Age of ASFA

Richard Wexler*

• In Orange County, California, an impoverished single mother cannot find someone to watch her children while she works at night, tending a ride at a theme park. So she leaves her eight, six, and four-year-old children alone in the motel room that is the only housing they can afford. Someone calls child protective services. Instead of helping her with babysitting or daycare, they take away the children on the spot.[1]

• In Akron, Ohio, a grandmother raises her eleven-year-old granddaughter despite being confined to a wheelchair with a lung disease. Federal budget cuts cause her to lose housekeeping help. The house becomes filthy. Instead of helping with the housekeeping, child protective services takes the granddaughter away and throws her in foster care for a month. The child still talks about how lonely and terrified she was – and about the time her foster parent took her picture and put it in a photo album under the heading: "filthy conditions."[2]

• In Paterson, New Jersey, parents lose their three children to foster care solely because they lack adequate housing. When the children are returned, one of them shows obvious signs of abuse – bruises and new and old burn marks – from foster care. The parents are suing. And so is their first caseworker. He never wanted the children taken away. The caseworker even found the family a better apartment. That is, however, not what his superiors wanted. Indeed, the caseworker says that because he insisted on trying to help the family, and refused to alter his reports to make the parents look bad, he was fired. Why were his superiors so anxious to take away the children? There was a rich, suburban couple ready and waiting to adopt them. And according to the lawsuit filed by the caseworker, a supervisor told him that "'children should be taken away

---

* Executive Director, National Coalition for Child Protection Reform (NCCPR); Author, *Wounded Innocents: The Real Victims of the War Against Child Abuse* (Prometheus Books: 1990, 1995).

1. *See* Laura Saari, *Checking Up on the Children*, ORANGE COUNTY REGISTER, Jan. 17, 1999, at E1.

2. *See* Donna J. Robb, *Child Abuse Charge Unfair, Group Says*, THE PLAIN DEALER, March 11, 1998, at 1B.

from poor parents if they can be better off elsewhere.'"[3]

• In Los Angeles, California, the pipes in a grandmother's rented house burst, flooding the basement and making the home a health hazard. Instead of helping the grandmother find another place to live, child protective workers take away the granddaughter and place her in foster care. She dies there, allegedly killed by her foster mother. The child welfare agency that would spend nothing to move the family offers $5,000 for the funeral.[4]

Welcome to child welfare in the age of ASFA, the so-called Adoption and Safe Families Act.[5] ASFA is a federal law helping to turn the nation's child welfare systems into the ultimate middle-class entitlement. Step right up and take a poor person's child for your very own.

ASFA was the culmination of an assault on safe, effective programs to keep families together that began in the 1990's.[6] The law has caused untold misery for thousands of children. While supposedly intended to solve the problems of the foster care system, it has, in fact, worsened those problems. In the name of promoting adoption, it is creating a generation of legal orphans. And worst of all, in the name of child safety, it has made children less safe.

There is nothing really new about the ASFA approach. Its guiding philosophy can be boiled down to a single sentence: "take the child and run." And that approach is, in fact, an old story. Indeed, it dates back to the inception of the system in the mid-nineteenth century when Charles Loring Brace, a Protestant clergyman in New York City with an abiding hatred of urban Catholic immigrants, preached that those immigrants were genetically inferior and that the only hope for their children was immediate removal from their parents' "evil influence."[7]

Brace practiced what he preached. He shipped tens of thousands of children around the country on "orphan trains," although many were not orphans.[8] Once taken, the children were handed out to farmers looking for cheap labor.[9] As one observer noted at a Conference of Charities in 1883, "[t]heir slaves [having been] set free, these men needing labor take these

---

3. *See* Jennifer V. Hughes, *Lawsuit Says DYFS Ordered False Reports – Caseworker Claims He Was Fired For Refusing*, THE RECORD, May 4, 2001 at L3.

4. *See* Nicholas Riccardi, *Grandmother Blames County in Latest Death of Foster Child*, LOS ANGELES TIMES, June 15, 1999, at B1.

5. *See* An Act to Promote the Adoption of Children in Foster Care, Pub. L. No. 105-89, 117 Stat. 2115 (1997).

6. *See id.*

7. *See* CHARLES LORING BRACE, THE DANGEROUS CLASSES OF NEW YORK AND TWENTY YEARS' WORK AMONG THEM 26-31 (Wynkoop and Hallenbeck, 1872).

8. *See generally* MIRIAM LANGSAM, CHILDREN WEST: A HISTORY OF THE PLACING-OUT SYSTEM OF THE NEW YORK CHILDREN'S AID SOCIETY, 1853-1890 (1964).

9. *See id.* at 21-26.

boys and treat them as slaves."[10]

Today, it is generally not considered fashionable to brand the poor genetically inferior. Instead, we label them psychologically inferior. A mother who works two jobs to make ends meet and sleeps with the lights on so the rats will not come out at night and bite her children may have good reason to believe the world is not an altogether friendly place. But instead of addressing her problems by getting her a better job and a better place to live, the child protection system is likely to take away her children, brand her as suffering from, for example, "apathy futility syndrome,"[11] and force her into counseling sessions that only add one more burden to her life.

No matter what problem may have led to removal of the children from a family, the parents almost always are required to endure "counseling" and "parent education" before they can get their children back.

CONFUSING POVERTY WITH NEGLECT

There is substantial evidence that cases like those described at the beginning of this article are far more typical than horror story cases involving the very few parents who really are brutally abusive or hopelessly addicted. For example, a study of "boarder babies," children who spend months in hospitals, found that the biggest single factor causing their forced hospital stays was lack of housing.[12] Families struggling to keep their children out of foster care are stymied by two major problems: homelessness and low public assistance grants.[13]

Similarly, a study of "lack of supervision" cases in New York City by the Child Welfare League of America found that in fifty-two percent of the cases studied, the service needed most was, as one might expect, day care or babysitting.[14] However, the "service" offered most often was foster care. Courts in New York and Illinois have found that families are repeatedly kept apart solely because they lack decent housing.[15] In Wash-

---

10. *Id.* at 59.

11. *See* NORMAN POLANSKY, ET AL., DAMAGED PARENTS: AN ANATOMY OF CHILD NEGLECT (1981).

12. CITY OF NEW YORK OFFICE OF THE COMPTROLLER, OFFICE OF POLICY MANAGEMENT, WHATEVER HAPPENED TO THE BOARDER BABIES? 11-12 (Jan. 1989).

13. KAREN BENKER & JAMES REMPEL, INEXCUSABLE HARM: THE EFFECT OF INSTITUTIONALIZATION ON YOUNG FOSTER CHILDREN IN NEW YORK CITY 25 (May 1989).

14. *See* MARY ANN JONES, PARENTAL LACK OF SUPERVISION: NATURE AND CONSEQUENCES OF A MAJOR CHILD NEGLECT PROBLEM 19-25 (1987).

15. *See* Cosentino v. Perales, 524 N.Y.S.2d 121 (S. Ct. 1987); Norman v. Johnson, 739 F. Supp. 1182 (N.D. Ill. 1990); *see also* Janita Poe & Peter Kendall, *Cases of Neglect May Be Only Poverty in Disguise*, CHI. TRIB., Dec. 24, 1995, at C1.

ington D.C., where the foster care system was taken over for several years by the federal courts, the first receiver named by the court to run the agency found that between one-third and one-half of D.C.'s foster children could be returned to their parents right now, if they just had a decent place to live.[16] In California, homeless children were given emergency shelter only on the condition that they be separated from their parents, until a successful lawsuit put an end to the practice.[17] The National Commission on Children found that children often are removed from their families "prematurely or unnecessarily" because federal aid formulas give states "a strong financial incentive" to do so rather than provide services to keep families together.[18] Indeed, state laws make the confusion of poverty with neglect almost inevitable, by "defining in" almost every poor family.

In Ohio, a child is considered neglected when his or her "condition or environment is such as to warrant the state, in the interests of the child, in assuming his guardianship."[19] In Illinois, parents neglect a child when they fail to provide for a child's well being."[20] Mississippi defines neglect as treating a child "without proper care, custody, supervision, or support."[21] In South Dakota, a child is neglected when his "environment is injurious to his welfare."[22]

Among the most tragic cases are those involving battered women. Rather than work to remove the abuser from the home, child protection agencies often behave as they did in the case of Rosa.[23] They target the battered woman, for "failure to protect" her children from witnessing her abuse. Whatever trauma the child already has endured by witnessing domestic violence then is compounded by his or her removal from the mother. According to the New York City Interagency Task Force Against Domestic Violence: "[i]ncreasingly in New York City, abuse and neglect proceedings are brought against battered mothers whose children are removed from them where the only allegation is their children's exposure to domestic violence."[24]

According to The Village Voice:

---

16. *See* Tamar Lewin, *Child Welfare Is Slow to Improve Despite Court Order*, THE N.Y. TIMES, Dec. 30, 1995, at 6.

17. Hansen v. McMahon, Cal. Sup. Ct. April 22, 1986 (No. CA000974).

18. NAT'L COMM'N ON CHILD., BEYOND RHETORIC: A NEW AMERICAN AGENDA FOR CHILDREN AND FAMILIES 289-90 (May, 1991).

19. OHIO REV. CODE ANN. § 2151.04 (West 2000).

20. 23 ILL. COMP. STAT. § 2053 (2000).

21. MISS. CODE ANN. § 43-21-105 (2000).

22. S.D. CODIFIED LAWS §§ 26-8-2, 26-8-6 (2000).

23. *See* Zanita E. Fenton, *Foster Care: The Border of Family Identity Maintaining, (Re)creating, Destroying*, 36 NEW ENG. L. REV. 59 (2001).

24. Karen Houppert, *Victimizing the Victims*, VILLAGE VOICE, June 15, 1999, at 41.

> [T]his is happening even in cases ... where the mother has already left her abuser and the kids are in no danger. Domestic violence advocates across the city report a flood of new failure-to-protect cases crossing their desks and assert that the city, in its prosecutorial zeal, is punishing victims rather than batterers, and traumatizing kids rather than rescuing them.[25]

The task force report states: "[k]nowing that they may be investigated by child protective services, battered mothers are more likely to remain in the abusive home, isolated and afraid, so that they can remain with their children."[26] The practice is now the subject of a lawsuit for which plaintiffs are seeking class action certification.[27]

New York City is not alone. According to The Ledger, a newspaper in Central Florida: "[f]earful of losing their children to state child abuse workers, many women avoid the streets by staying in abusive relationships, living with friends or leaving children with relatives."[28] "There is some validity to the fear of losing their children,' said Ruth Olinger of Lighthouse Ministries, 'and it prevents many homeless women from asking for help. It's sad. They don't know what else to do to feed their children and protect them from the (foster care) system.'"[29] Advocates who are "in denial" about the problem of wrongful removal offer only one piece of counter-evidence, a single report based on a case reading from a single state, New Jersey.[30]

Case readings, however, are inherently biased against families. The case records reflect only the caseworker's view of events, not the family's. Judging whether a family should have been destroyed based on reading a case record, is like judging the guilt or innocence of a criminal defendant based solely on reading an indictment. For example, reading a New Jersey case in this way may be particularly dubious. Children's Rights Inc. (CRI) alleged widespread falsification of case records in New Jersey – in order to make those records look better for a case reading that CRI wants to conduct in connection with a class action lawsuit.[31] A federal judge dis-

---

25. *Id.*

26. *Id.*

27. *See* Nicholson v. Williams, No. 00-CV-2229 (E.D. NY 2000) (Plaintiffs' co-counsel, Carolyn Kubitschek, is a member of the NCCPR Board of Directors).

28. Eric Pera, *Address Unknown: New Report Shows a 64 Percent Rise in Polk's Homeless; A Place to Call Home*, THE LEDGER, Dec. 16, 2000, at A1.

29. *See id.*

30. *See* CECELIA ZALKIND, STOLEN FUTURES: A REPORT ON PREVENTING FOSTER CARE PLACEMENT IN NEW JERSEY 21-33 (1994).

31. *See* Press Release, Children's Rights, Inc., *New Jersey Child Welfare Agency Accused of Corrupting Evidence in Federal Lawsuit Seeking Its Reform*, March 7, 2001; *see also* Nancy Parello *DYFS Witnesses Differ on Audits of Child Abuse*, THE RECORD, March 30, 2001, at A4.

missed the claim of the falsification.[32] As noted at the start of this article, in a separate action, a New Jersey caseworker is suing the state child welfare agency, alleging he was fired for refusing to falsify a case record to make parents look bad.[33]

Other cases fall on a broad continuum between the extremes, the parents are neither all victim nor all villain. Sometimes these cases involve substance abuse. When the issue is drugs, another question arises: Why even bother with drug treatment for such parents? Would not the children simply be "better off" if they were confiscated on the spot and never returned? The answer is that we should not provide treatment for the sake of the parent, but for the sake of the child. In a University of Florida study of infants born addicted to crack cocaine, one group was placed in foster care, and the other with birth mothers able to care for them.[34] After six months, the babies were tested using all the usual measures of infant development: rolling over, sitting up, reaching out.[35] Consistently, the children placed with their birth mothers did better.[36] For the foster children, the separation from their mothers was more toxic than the cocaine.[37]

It is extremely difficult to take a swing at "bad mothers" without the blow landing on their children. Therefore, if we really believe all the rhetoric about the needs of the children coming first, we must put those needs before anything – even our anger at their parents. But of course we do not do that. Instead, we have built a child welfare system on a foundation of myths about who is in that system and why. That is why the system is constantly on the verge of collapse, and why the latest effort to fix it – ASFA – is backfiring.

THE MYTH OF THE VAST FAMILY PRESERVATION CONSPIRACY

To hear the child welfare establishment explain it, the system is in its current condition because of a fanatical devotion to keeping families together. It is argued that children are forced to languish in foster care while agencies lavish services on their ne'er-do-well parents, or worse, return

---

32. Nancy Parello, *Group Brushes Off Losses in Suit Against DYFS*, THE RECORD, Oct. 9, 2001, at A3.

33. *See* Hughes, *supra* note 3.

34. *See Family Preservation and Substance Abuse*, Issue Papers (citing Kathleen Wobie, et al., *To Have and To Hold: A Descriptive Study of Custody Status Following Prenatal Exposure to Cocaine*, (paper presented at joint annual meeting of the American Pediatric Society and the Society for Pediatric Research) (May 3, 1998)), *at* http://www.nccpr.org/newissues/12.html (last visited Nov. 2, 2001).

35. *See id.*

36. *See id.*

37. *See id.*

those children to unsafe homes. George Orwell would have appreciated the extent to which the real history of child welfare has been sent down the "memory hole" and replaced by this fake history. Understanding the real history is crucial to understanding why ASFA is backfiring.

By the late 1970's, there were more than 500,000 children in foster care in the United States – about as many, relative to the total child population, as today.[38] Alarmed by this, Congress passed a bill in 1980 called the Adoption Assistance and Child Welfare Act (AACWA).[39] This law often is wrongly blamed for forcing children to languish in foster care. That in itself indicates a crucial flaw in the revisionist history of child welfare: If the problem is a law passed in 1980, what were all those children doing stuck in foster care in the late 1970's?

In fact, the 1980 law was the first effort to try to prevent children from languishing in foster care. It was the first federal law to set time limits on foster care. It was the first to require permanent placements for children either through return home or through adoption.[40] It was the first to provide federal aid for adoptions.[41] Nowhere in that law was there a requirement that judges or social workers focus solely on reunification or on keeping children out of foster care – indeed, it would be a violation of the law to do so. The law required only that states make "reasonable efforts" to keep families together.[42]

For a while the law worked. The foster care population fell to about 243,000 by 1982.[43] By then, however, the Reagan Administration was in office. The administration disliked the AACWA and pulled back Department of Health and Human Services regulations to enforce it. This action sent a signal to the states, which was that they could return to business as usual. By 1983, the foster care population was rising again,[44] and by 1998, it had more than doubled.[45] If family preservation advocates were as influential as alleged by critics, why were there so many children in foster

---

38. *See* LEROY PELTON, FOR REASONS OF POVERTY: A CRITICAL ANALYSIS OF THE PUBLIC CHILD WELFARE SYSTEM IN THE UNITED STATES 6 (1989).

39. 42 U.S.C. § 670 (1988).

40. *See* Adoption Assistance and Child Welfare Act, 8 U.S.C. §670 (1988) [hereinafter AACWA].

41. *See id.*

42. *See id.*

43. *See* Pelton, *supra* note 38, at 6.

44. *See id.*

45. *See* U.S. DEPARTMENT OF HEALTH AND HUMAN SERVICES, ADMINISTRATION FOR CHILDREN AND FAMILIES, ADOPTION AND FOSTER CARE REPORTING SYSTEM (AFCARS) REPORT, (Jan., 1999), *available at* http://www.acf.dhhs.gov/programs/cb/publications/afcars/rpt0199/ (last visited May 14, 2001).

care? Drugs are part of the explanation; the increase in the foster care population preceded, however, the crack epidemic, and it continues even as that epidemic has waned. Indeed, at a time when almost every other social indicator is moving in the right direction, the foster care population continues to rise, standing at 588,000 children as of March, 2000.[46] This is a stark testament to the failure of ASFA.

To attribute the fact that children are languishing in foster care to efforts to keep families together requires a belief in a kind of institutional schizophrenia. It requires us to believe that workers were perfectly willing to take huge numbers of children away from their parents and then, suddenly, a family preservation gene kicks in and these same workers move heaven and earth to get them back out. As one investigative reporter, however, said in a lecture: "[w]hen there is a complicated explanation and a simple explanation, the simple explanation is almost always correct."[47]

The simple explanation is "take the child and run," which has always been the ethos of the child welfare system. Once the states were free to go back to business as usual, too many children were taken too quickly from their homes. They were placed in a foster care system that did not focus solely on reunification. In fact, it did not focus on reunification at all. Rather the children were filed away and forgotten as overwhelmed workers rushed on to the next case. The "reasonable efforts" requirement was never enforced by the federal government, and in 1989, the U.S. Supreme Court ruled that individuals could not sue to enforce it.[48]

FOSTER CARE IS DANGEROUS

ASFA requires putting its version of "child safety" ahead of "family preservation," because Congress failed to realize that family preservation in fact makes children safer. Indeed, the family preservation movement always has made child safety its first priority. ASFA efffectively equated child safety with child removal, which is another example of dangerous historical revisionism.

In the late 1970's, a Washington State agency devised a new way of providing intensive intervention to keep families together, which they

---

46. *See* U.S. DEPARTMENT OF HEALTH AND HUMAN SERVICES, ADMINISTRATION FOR CHILDREN AND FAMILIES, ADOPTION AND FOSTER CARE REPORTING SYSTEM (AFCARS) REPORT, (Apr., 2001), *available at* http://www.acf.dhhs.gov/programs/cb/publications/index.htm (last visited May 14, 2001). [hereinafter AFCARS Report 2001].

47. Bruce Locklin, former investigative reporter for The Record, Hackensack, New Jersey, made the comment during a lecture to reporters and editors at the Albany Times Union.

48. *See* Suter v. Artist M, 503 U.S. 347 (1992).

called Homebuilders.[49] The term "family preservation" was invented specifically to apply only to programs that rigorously follow the Homebuilders model.[50]

These programs actually have a far better track record for safety than foster care – both because Homebuilders' programs are so safe, and because foster care can be very dangerous.

National data on child abuse fatalities show that a child is more than twice as likely to die of abuse in foster care than in the general population.[51]

A study of reported abuse in Baltimore, found the rate of "substantiated" cases of sexual abuse in foster care more than four times higher than the rate in the general population.[52] Using the same methodology, an Indiana study found three times more physical abuse and twice the rate of sexual abuse in foster homes than in the general population.[53] In group homes there was more than ten times the rate of physical abuse and more than twenty-eight times the rate of sexual abuse as in the general population, in part because so many children in the homes abused each other.[54]

Those studies deal only with reported maltreatment. The actual amount of abuse in foster care is likely to be far higher, since agencies have a special incentive not to investigate such reports, since they are, in effect, investigating themselves.

In New York City, for example, where Children's Rights Inc. settled a lawsuit against the child welfare system,[55] they found that, "[a]buse or neglect by foster parents is not investigated because [agencies] tolerate behavior from foster parents which would be unacceptable by birth parents."[56] Additionally, a lawyer who represents children in Broward County, Florida, says in a sworn affidavit that over a period of just eighteen months he was made personally aware of fifty instances of child-on-child sexual abuse involving more than 100 Broward County foster chil-

---

49. *See generally*, National Coalition for Child Protection Reform's Issue Paper # 9, *available at* http://www.nccpr.org (last visited Sept. 10, 2001).

50. *See id.*

51. *See* U.S. DEPT. OF HEALTH AND HUMAN SERVICES, ADMINISTRATION ON CHILDREN, YOUTH AND FAMILIES CHILD MALTREATMENT 1999 viii, 41 (2001).

52. *See* MARY I. BENEDICT & SUSAN ZURAVIN, FACTORS ASSOCIATED WITH CHILD MALTREATMENT BY FAMILY FOSTER CARE PROVIDERS 27-30 (1992).

53. *See* J. William Spencer & Dean D. Kundsen, *Out of Home Maltreatment: An Analysis of Risk in Various Settings for Children, in* CHILD. AND YOUTH SERV. REV. 485, 485-92 (1992).

54. *See id.*

55. *See* Marisol A. v. Giuliani, No. 95-Civ-10533 (S.D.N.Y., filed Dec. 13, 1995).

56. Complaint, at 75, Marisol A. v. Giuliani, No. 95-Civ-10533.

dren.[57] The official number of child-on-child sexual abuse instances during the same period was seven, because until the "epidemic of child-on-child sexual abuse" was exposed, the state child abuse hotline did not accept reports of such abuse.[58]

It is not surprising that studies not limited to official reports produce even more alarming results. Another Baltimore study, this one examining case records, found abuse in twenty-eight percent of the foster homes studied –more than one in four.[59]

Even what is said to be a model foster care program, where caseloads are kept low and workers and foster parents get special training, is not immune. When alumni of the Casey Family Program were interviewed, twenty-four percent of the girls said they were victims of actual or attempted sexual abuse in their foster homes.[60] Furthermore, this study asked only about abuse in the foster home that the children had lived in the longest. A child who had been moved from a foster home where she had been abused, after only a short stay, would not even be counted.[61] Officials at the program say they have since lowered the rate of all forms of abuse to "only" twelve percent, but this is based on an in-house survey of the program's own caseworkers, not outside interviews with the children themselves.[62]

This does not mean that all, or even many, foster parents are abusive. But the abusive minority is large enough to cause serious concern, and it is probably growing as the influx of children into the system caused by ASFA increases pressure on agencies to overcrowd foster homes and lower standards. Furthermore, abuse in foster care does not always mean abuse by foster parents. It can be caused by foster children abusing each other.

FAMILY PRESERVATION IS SAFER

Compare the record of foster care to the record of family preservation. The original Homebuilders program has served 12,000 families since 1982.[63] No child has ever died during a Homebuilders intervention and

---

57. *See* Affidavit of David S. Bazerman, Esq. at 4, Ward v. Feaver, No. 98-7137, (S.D. Fla. filed Dec. 16, 1998).

58. *See id.*

59. *See* L.J. v. Massinga, 699 F. Supp. 508 (D. Md. 1988).

60. *See* DAVID FANSHEL, ET AL., FOSTER CHILDREN IN A LIFE COURSE PERSPECTIVE 90 (1990).

61. *See id.*

62. *See* PETER J. PECORA ET AL., HOW ARE THE CHILDREN DOING? ASSESSING YOUTH OUTCOMES IN FAMILY FOSTER CARE 18-19 (1998).

63. Personal communication with Charlotte Booth, Executive Director, Homebuild-

only one child, more than a decade ago, has since died.[64] Michigan has the nation's largest family preservation program. The program rigorously follows the Homebuilders model. Since 1988, the Michigan family preservation program has served 90,000 children.[65] During the first two years, two children died during the intervention.[66] In the decade since, there has not been a single fatality.[67] In contrast, when Illinois effectively abandoned family preservation during the mid-1990s, there were five child abuse deaths in foster care in just one year.[68]

The other state in the forefront of family preservation efforts in recent years is Alabama. Alabama is implementing a consent decree resulting from a federal lawsuit requiring it to reframe its whole approach to child welfare by following family preservation principles.[69] Learning from the failures of other states that tried to change their systems overnight, the Alabama approach calls for gradual, county-by-county change. But the results already have been dramatic: in counties adopting a family preservation approach, foster care placements have declined by thirty-three percent.[70] More important, an independent, court-appointed monitor concluded that children in Alabama are safer now than before the system switched to a family preservation model. The monitor wrote that "the data strongly support the conclusion that children and families are safer in counties that have implemented the R.C. reforms[.]"[71]

One advocate for taking even more children from their parents, Elizabeth Bartholet, repeatedly argues that as bad as foster care is, children removed from their homes are better off there than in their own homes.[72] Although she makes the claim repeatedly, her evidence appears limited to

---

ers.

64. *See id.* Even in the one case in which a child died after the intervention, in 1987, Homebuilders had warned that the child was in danger, but the warning from Homebuilders was ignored. *See id.*

65. Personal Communication with Susan Kelly, former director of family preservation services, Michigan Family Independence Agency.

66. *See id.*

67. *See id.*

68. *See* Peter Kendall & Terry Wilson, *Boy's Death Casts Shadow on Relatives' Foster Care*, CHI. TRIB., Feb. 28, 1995, at 7.

69. *See* R.C. v. Hornsby, 88-D-1170-N (M.D. Ala., filed 1988).

70. *See Issue Paper # 1: Foster Care vs. Family Preservation: The Track Record on Safety* (citing IVOR D. GROVES, A SUMMARY REPORT ON IMPLEMENTATION STATUS OF THE R.C. V. PETELOS CONSENT DECREE 10 (1999)), *at* http://www.nccpr.org/newissues/1.html (last visited Nov. 2, 2001).

71. IVOR D. GROVES, SYSTEM OF CARE IMPLEMENTATION: PERFORMANCE, OUTCOMES, AND COMPLIANCE 3 (Mar. 1996).

72. *See generally* ELIZABETH BARTHOLET, NOBODY'S CHILDREN: ABUSE AND NEGLECT, FOSTER DRIFT AND THE ADOPTION ALTERNATIVE (1999).

one study and two articles that make brief references to official statistics.[73] The study compared two matched, but very small, groups of children. Because the differences in outcomes and the sample size were so small, the authors did not conclude that foster care is preferable to home placement.[74] Additionally, the lead author of the study says they compared average services for birth families with a better-than-average foster care program.[75] The relevant comparison is children who are kept at home with real family preservation services versus typical foster care.

Bartholet also cites two studies, which make brief reference to claims that children are more likely to be reabused at home than abused in foster care.[76] But these references use only officially-reported maltreatment – a major flaw that biases the results in ways discussed above. They also refer only to physical safety, ignoring the emotional trauma of foster care. And again, they do not compare foster care to leaving children at home with real family preservation services. They are also out of date. A newer study comparing foster children to children monitored by child protective services in their own homes found the children left in their own homes were safer.[77] Bartholet also cites one other study, but this one compared all violent deaths, not just child abuse fatalities.[78]

But more fundamentally, this is a case in which knowing what works best on average is particularly unhelpful. Though the evidence is strong that, on average, real family preservation programs are safer than foster care, it would be ludicrous – and dangerous – to leap to the conclusion that therefore every family should be offered such services and no child should be placed in foster care. It is equally ludicrous, therefore, even if one believes foster care is safer on average, to use that as justification to abolish family preservation and use foster care for thousands of children who would be safer in their own homes.

There are many cases in which foster care, horrible though it can be, is better than leaving a child at home. When we know a child is being severely beaten, tortured, raped, starved, abandoned or similarly harmed,

---

73. *See id.*

74. *See id.*

75. *See* MICHAEL WALD ET AL., PROTECTING ABUSED AND NEGLECTED CHILDREN 187-91 (1988).

76. *See id.* at 97.

77. *See Child Protection Oversight: Hearing on H781-20 Before the Subcomm. on Human Resources of the House Comm. on Ways and Means*, 1999 Leg., 1st Sess. 80 (Wash. 2000) (testimony of Jess McDonald, Dir., Illinois Department of Children and Family Services) [hereinafter McDonald Testimony].

78. *See generally* Richard P. Barth & Debra L. Blackwell, *Death Rates Among California's Foster Care and Former Foster Care Populations*, 20 CHILD. AND YOUTH SERV. REV. 601 (1998).

obviously the risks of foster care are worth taking. But contrary to the claims of family preservation critics, many cases do not even come close to falling within these categories. And in many of those cases, foster care is indeed far worse.

It is worse because of the emotional devastation wrought by tearing a child, who has not been seriously abused or neglected, from everything loving, friendly and familiar. But it also is worse because of the alarming possibility that the child who was not abused at home will be abused in foster care.

FAMILY PRESERVATION IS EFFECTIVE

Evidence for the effectiveness of the Homebuilders model also is strong:

- Michigan's Families First program sticks rigorously to the Homebuilders model. The Michigan program was evaluated by comparing children who received family preservation services to a "control group" that did not. After one year, among children who were referred because of abuse or neglect, the control group children were nearly twice as likely to be placed in foster care, as the Families First children. Thirty-six percent of children in the control group were placed, compared to only 19.4 % of the Families First children.[79]
- Another Michigan study went further. In this study, judges actually gave permission to researchers to "take back" some children they had just ordered into foster care and place them in Families First instead. One year later, ninety-three percent of these children still were in their own homes.[80] And Michigan's State Auditor concluded that the Families First program "has generally been effective in providing a safe alternative to the out-of-home placement of children who are at imminent risk of being removed from the home . . . . The program places a high priority on the safety of children."[81]
- An experiment in Utah and Washington State also used a comparison group. After one year, 85.2 % of the children in the comparison group were placed in foster care, compared to only 44.4 % of the children who received intensive family preservation services.[82]

---

79. *See* CAROL BERQUIST, ET AL., EVALUATION OF MICHIGAN'S FAMILIES FIRST PROGRAM (1993).

80. *See* BETTY J. BLYTHE & SRINIKA JAYARATNE, FAMILIES FIRST EFFECTIVENESS STUDY: A SUMMARY OF FINDINGS 18 (Sept. 28, 1999).

81. STATE OF MICHIGAN, OFFICE OF THE AUDITOR GENERAL, 1998 PERFORMANCE AUDIT OF THE FAMILIES FIRST OF MICHIGAN PROGRAM 2-4 (July 1998).

82. *See* MARK W. FRASER, ET AL., FAMILIES IN CRISIS: THE IMPACT OF INTENSIVE FAMILY PRESERVATION SERVICES 168 (1991).

• A study in California found that fifty-five percent of the control group children were placed, compared to only twenty-six percent of the children who received intensive family preservation services.[83]

• A North Carolina study comparing 1,254 families receiving Intensive Family Preservation Services (IFPS) to more than 100,000 families who did not find that "IFPS consistently resulted in fewer placements . . . ."[84]

• And still another study, in Minnesota, found that, in dealing with troubled adolescents, fully ninety percent of the control group children were placed, compared to only fifty-six percent of those who received intensive family preservation services.[85]

Some agencies are now using IFPS to help make sure children are safe when they are returned home after foster care. Here again, researchers are beginning to see impressive results. In a Utah study, 77.2 % of children whose families received IFPS help after reunification were still safely with their birth parents after one year, compared with 49.1 % in a control group.[86]

Critics ignore all of this evidence, preferring to cite a study done for the federal government which purports to find that IFPS is no better than conventional services. While critics of family preservation claim that this study evaluated programs that followed the Homebuilders model, this is not true. In a rigorous critique of the study, Prof. Ray Kirk of the University of North Carolina School of Social Work notes that the so-called IFPS programs in this study actually diluted the Homebuilders model, providing service that was less intensive and less timely.[87] At the same time, the "conventional" services sometimes were better than average. In at least one case, they may well have been just as intensive as the IFPS program – so it is hardly surprising that the researchers would find little difference between the two.[88]

---

83. *See* Sally Wood et al., *In-Home Treatment of Abusive Families: Cost and Placement at One Year*, 25 PSYCHOTHERAPY 409, 409-14 (1988) (*cited in* Howard Bath & David Haapala, *Family Preservation Services: What Does the Outcome Research Really Tell Us*, SOCIAL SERV. REV., Sept., 1994, at 400 app. A).

84. *See* Ray S. Kirk, *Tailoring Intensive Family Preservation Services for Family Reunification Cases: Research, Evaluation and Assessment, available at* www.nfpn.org/resources/articles/tailoring.html (last visited Nov. 2, 2001).

85. *See* I.M. Schwartz, et al., *Family Preservation Services as an Alternative to Out-of-Home Placement of Adolescents*, in FAMILY PRESERVATION SERVICES: RESEARCH AND EVALUATION 33-46 (K. Wells & D.E. Biegel, eds., 1991).

86. *See* Robert E. Lewis, et al., *Examining Family Reunification Services: A Process Analysis of a Successful Experiment*, 5 RES. ON SOC. WORK PRAC. 259, 261 (1995).

87. *See* RAY S. KIRK, A CRITIQUE OF THE "EVALUATION OF FAMILY PRESERVATION AND REUNIFICATION PROGRAMS: INTERIM REPORT" (May 2001).

88. *See id.*

Furthermore, efforts to truly assign families at random to experimental and control groups sometimes were thwarted by workers in the field who felt this was unethical. Workers resisted assigning what they considered to be "high risk" families to control groups that would not receive help from IFPS programs.[89] In addition, the study failed to target children who actually were at imminent risk of placement.[90] Given all these problems, Professor Kirk writes, "a finding of 'no difference between treatment and experimental groups' . . . is simply a non-finding from a failed study."[91]

Professor Kirk's findings mirror those of an evaluation of earlier studies purporting to show that IFPS was ineffective.[92] The evaluation found that these studies "did not adhere to rigorous methodological criteria."[93] In contrast, "there is a growing body of evidence that IFPS works, in that it is more effective than traditional services in preventing out-of-home placements of children in high-risk families."[94]

It was neither the safety nor the effectiveness of Intensive Family Preservation Services that prompted the backlash. Rather, it was the fact that the model was a threat to more than 100 years of child welfare orthodoxy, combined with the threat by Congress in the early 1990's to actually put some money behind the concept, raising the possibility that funds might be diverted from substitute care.[95]

So "family preservation" became the scapegoat whenever any child was left in any home under any circumstances and something went wrong – regardless of whether the child had been anywhere near a real family preservation program. And the horror stories eventually had an impact on Congress, prompting the passage of ASFA.

ASFA is filled with broad, vague catch-all clauses that blow huge holes in what little was left of the "reasonable efforts" clause in AACWA.[96]

---

89. *See id.*

90. *See id.*

91. *Id.*

92. *See id.*

93. Amy M. Heneghan, et al., *Evaluating Intensive Family Preservation Programs: A Methodological Review*, 97 PEDIATRICS 535-42, *cited in* RAY KIRK, FINAL REPORT: A RETROSPECTIVE EVALUATION OF NORTH CAROLINA'S INTENSIVE FAMILY PRESERVATION SERVICES PROGRAM (Chapel Hill, N.C.: Jordan Institute for Families, School of Social Work, University of North Carolina, Aug. 2000).

94. Kirk, *supra* note 87.

95. Congress ultimately passed the Family Preservation and Family Support Act of 1993, reauthorized four years later as part of ASFA, with the name changed to the Promoting Safe and Stable Families program. The program funds a huge array of programs, even including foster care and adoption. It does not actually require that any money be spent on family preservation, and certainly does not require any spending on Homebuilders-type programs.

96. *See supra*, note 5, tit. I, § 101.

ASFA states, for example, that states don't have to make "reasonable efforts" in cases of "chronic abuse," a term that is not further defined.[97] And ASFA declares that "the child's health and safety shall be the paramount concern."[98]

That was always the case – a fact made clear in the legislative history of the AACWA and in guidance from the Department of Health and Human Services.[99] By placing it in a new law as though it were some sort of change, Congress has invited agencies to take away children, even when there are alternative means to ensure safety. Worse, Congress has sent a message that boils down to: child safety equals child removal, so take away even more children.

Indeed, far more important than any specific language in ASFA was its overall message – a message that reinforced the traditional "take the child and run" approach to child welfare. And that is why, even as crime, unemployment, poverty and reports of child abuse itself have declined, the foster care population continues to go up.

ASFA was supposed to reduce the number of children in foster care. That has not happened. When ASFA was passed, there were 520,000 children in foster care.[100] As of March, 2000, the figure stands at 588,000.[101]

ADOPTION AS PANACEA

ASFA was based on the false premise that children were languishing in foster care because the Vast Family Preservation Conspiracy was holding them hostage. Get rid of family preservation, it was argued, and the foster care system would empty out as children were quickly and eagerly adopted. It has not worked out that way.

Although adoptions are up significantly in percentage terms, in raw numbers the improvement has been minimal. Between 1996 and 2000, adoptions of foster children increased from 28,000 to 50,000 per year.[102] That is an average increase of 5,500 adoptions per year.[103] However, 5,500 accounts for less than one percent of the children in foster care on

---

97. *See id.*

98. *Id.*

99. *See generally H.R. 867, The "Adoption Promotion Act of 1997" Before the House Subcomm. on Human Resources of the Comm. on Ways and Means*, 105th Cong. 44-55 (1997) (testimony of MaryLee Allen, Director, Child Welfare and Mental Health Division, Children's Defense Fund).

100. *See* AFCARS Report 2001, *supra* note 46.

101. *See id.*

102. *See* Press Release: *HHS Awards Adoption Bonuses* (Sept. 10, 2001), *available at* http://www.acf.dhhs.gov/news/press/2001/adoption.html (last visited Nov. 2, 2001).

103. *See id.*

any given day. Meanwhile, because ASFA also encourages the "take the child and run" mentality, more children are being needlessly taken from their homes. So despite the increase in adoptions, the foster care population continues to rise by about four percent per year.[104] Obviously, America is not going to adopt its way out of the foster care crisis. And it is not because states have not tried. ASFA gives states a bounty of $4,000 to $6,000 for every finalized adoption over a baseline number. As a result, states have been falling all over themselves pushing adoption at all costs. But the families simply are not out there. The rush to embrace adoption as panacea has had some serious side effects.

In the early 1990's, Professor Martin Guggenheim found that even before ASFA, state efforts to expedite termination of parental rights were creating a generation of "legal orphans" with no ties to birth parents but without adoptive homes either.[105] Thanks to ASFA, the problem is spreading. In New Jersey, between 1997 and 1999 almost four children had parental rights terminated for every one actually adopted.[106] A study of urban counties in Nebraska found that more than a year after parental rights were terminated, fewer than half the children had permanent homes.[107]

Furthermore, the adoption bounties in ASFA create a perverse incentive for quick-and-dirty, slipshod placements that are more likely to fail. That is because if an adoption of a foster child fails, states do not have to return the bounty. In fact, they can place the same child again and collect another bounty. That means the "disruption" rate for adoption of foster children, already estimated at roughly fifteen to twenty-five percent,[108] is likely to increase.

Professor Richard Gelles of the University of Pennsylvania, who was

---

104. *See Child Abuse Prevention Program, House Ways and Means Comm., Subcomm. on Human Resources*, 107th Cong. 2001 WL 21754863 (May 10, 2001) (testimony of Dr. Fred H. Wulczyn, Chapin Hall Center for Children).

105. *See* Martin Guggenheim, *The Effects of Recent Trends to Accelerate the Termination of Parental Rights of Children in Foster Care – An Empirical Analysis in Two States*, 29 FAM. L. Q. 121 (Spring 1995) (Prof. Guggenheim is president of the National Coalition for Child Protection Reform).

106. *See* Nancy Parello, *"Legal Orphans" Languish With Understaffed State Agencies, Busy Courts*, ASSOCIATED PRESS NEW JERSEY WIRE, April 12, 2000.

107. *See generally* VICTORIA WEISZ & ANGELA WILLIAMS, COURT FILE REVIEWS: ASFA REQUIREMENTS NOV. 1999 TO MARCH 2000, NEBRASKA COURT IMPROVEMENT PROJECT (2000).

108. *See* National Adoption Information Clearinghouse, *Disruption and Disolution, at* http://www.calib.com/naic/pubs/s_disrup.htm (last visited Sept. 10, 2001); *see also Adoption: What to do When Love is Not Enough*, THE DESERT NEWS, April 15, 1999, at C1; *Nobody's a Winner When Adoption Fails*, ARIZONA REPUBLIC, March 9, 1999, at B1; *Number of Children in DHS Care to Grow*, SUNDAY OKLAHOMAN, July 19, 1998, at 1.

involved in the writing of ASFA, has said: "[i]nitially, this was just supposed to be a safe families bill, not really an adoption bill at all. The adoption component was a way of sanitizing the bill, to make it more appealing to a broader group of people. Adoption is a very popular concept in the country right now."[109]

FOSTER CARE PANICS

In passing ASFA, Congress failed to learn the lessons some states have begun to learn after experiencing foster care panics – huge, sudden increases in placements that follow intensive media coverage of the death of a child who was known to the system. In April, 1993, three-year-old Joseph Wallace was killed by his mother.[110] Joseph was "known to the system."[111] "Family preservation" quickly became the scapegoat. Family preservation was attacked relentlessly by politicians and much of the media – even though most of the programs in Illinois bore little resemblance to the effective, Homebuilders-based models used in other states.

As a result, workers and judges became terrified to leave or return any child home for fear of becoming the next target of politicians and the Chicago media. Almost all efforts to keep families together were effectively abandoned amid claims that such efforts contradict "child protection." But instead of saving lives, child abuse deaths went up. They rose from seventy-eight before family preservation was abandoned to eighty-two the first year after it was abandoned, to ninety-one in 1997.[112] By 1996, a child was more likely to be placed in foster care in Illinois than in any other state.[113] Indeed, an Illinois child was more than four times as likely to be placed in foster care than a child in New Jersey.[114] Yet Illinois children were twice as likely to die of abuse as New Jersey children.[115]

The abandonment of family preservation led to a foster care panic that overwhelmed the system to the point that it created a backlog of more than

---

109. *My Law's 'Flawed' Says Richard Gelles*, CHILD WELFARE WATCH (Winter 2000) *available at* http://www.nycfuture.org/child_welfare/cww_06.htm (last visited Nov. 18, 2001).

110. *See Peace Comes to Amanda Wallace*, CHI. TRIB., Aug. 5, 1997, at 10.

111. *See id.*

112. *See* EXECUTIVE STATISTICAL SUMMARY STATE OF ILLINOIS, DEPARTMENT OF CHILDREN AND FAMILY SERVICES, OFFICE OF QUALITY ASSURANCE (Jan. 1998).

113. *See* CONNA CRAIG & DEREK HERBERT, THE STATE OF THE CHILDREN: AN EXAMINATION OF GOVERNMENT-RUN FOSTER CARE 9 (1997).

114. *See id.*

115. *See* CHING-TUNG WANG & DEBORAH DARO, CURRENT TRENDS IN CHILD ABUSE REPORTING AND FATALITIES: THE RESULTS OF THE 1996 ANNUAL FIFTY STATE SURVEY 13 (1997).

5,000 uncompleted investigations.[116] At the same time, foster care placements increased by thirty percent in just fourteen months, further straining the system.[117] Child abuse deaths in foster care in Illinois went from zero in the year before the foster care panic to five in the first year afterwards - an all-time record.[118]

The pattern showed itself in a new way in 1998 when the Illinois foster care panic finally began to abate.[119] That year, the number of child abuse deaths finally fell below the number before the panic began. And that year also was the first year since the panic in which the total number of Illinois children in foster care actually declined.[120] The decline has continued and a new study documents that as the numbers have fallen, children have gotten safer.[121] The foster care panic caused other tragedies in Illinois:

- Having supposedly "put children first," Illinois officials soon found they had no place to put children at all. So they were jammed into a hideous shelter, then overflowed into offices. Streetwise teens were thrown together with vulnerable younger children; infants were jammed into urine-soaked cribs. An eleven-year-old got hold of a gun and fired it.[122]
- Children were jammed into any foster home with a bed, with little screening of foster parents or foster children. As a result, according to Benjamin Wolf of the Illinois Affiliate of the American Civil Liberties Union, the Illinois foster care system became like "a laboratory experiment to produce the sexual abuse of children."[123]
- A study by the Child Welfare Institute found that at least one third of the children taken from their parents at the height of the foster care panic could safely have been returned to their own homes.[124] Abandoning family preservation took a bad system and made it, in Wolf's words, "unques-

---

116. *See* Sharman Stein, *DCFS Coordinator Puts Family Values to Work*, CHI. TRIB., Mar. 8, 1995, at 2.

117. *See id.*

118. *See* Kendall & Wilson, *supra* note 68.

119. *See* CRAIG & HERBERT, *supra*, note 113.

120. *See id.*

121. *See* McDonald Testimony, *supra* note 77; *see also* Illinois Department of Children and Family Services, *Signs of Progress in Child Welfare Reform, available at* www.state.il.us/dcfs/signscerap.shtml (last visited Nov. 2, 2001).

122. *See generally* Rob Karwath, *Abused Kids Sleep in DCFS Offices*, CHI. TRIB., June 29, 1993, at 1; Ellen Warren, *Toddlers, Troubled Teens All Wait Together at DCFS*, CHI. TRIB., July 21, 1993, at 1; Douglas Holt, *Boy Finds Gun at DCFS, Injured*, CHI. TRIB., Sept. 3, 1993, at 1; Rob Karwath, *Child Welfare Specialist Hired to Coordinate Overhaul of DCFS Site*, CHI. TRIB., Sept. 10, 1993, at 4; Rob Karwath, *DCFS Center Receives OK to House Kids*, CHI. TRIB., Dec. 17, 1993, at 1.

123. *See* Richard Wexler, *The Children's Crusade*, CHI. READER, Mar. 25, 1995, at 1.

124. *See* Bruce Dold, *Kids Suffer Under DCFS Reform Efforts*, CHI. TRIB., Sept. 22, 1995, at 19.

tionably worse."[125]

And what about the case that started it all? What was the role of family preservation in the case of Joseph Wallace? A family preservation worker recommended that the Wallace family not be preserved – he recommended to a judge that the child be removed. The judge agreed. The child was removed, but the records were lost when the family moved to another county. Only then was the child sent home to his death.[126] Not only was family preservation not the cause of the Wallace death -- family preservation almost saved Joseph Wallace's life.

Two years later, a child "known to the system" died in Connecticut.[127] Rather than learn from the Illinois experience, Connecticut officials repeated it. They too scapegoated family preservation. The result: the foster care population went up seventy-three percent,[128] there was a scandal over abuse in foster care,[129] and, as in Illinois, child abuse deaths increased.[130]

Then it was New York City's turn. Again, this time in late 1995, a child "known to the system" died.[131] Again officials blamed "family preservation," even though child abuse deaths had declined in New York City for four years in a row. Once again, the attacks set off a foster care panic, overwhelming the system. The result: thousands of children were forced to sleep, often on chairs and floors, in a violence-plagued, emergency makeshift shelter created from city offices.[132] A four-year-old foster child was beaten and starved to death in a foster home opened by one private agency, apparently desperate for beds after another had closed it down.[133] The decline in child abuse deaths ended. In 1996, for the first time in five

---

125. *See* Wexler, *supra* note 123.

126. *See* Joel J. Bellows, et al., THE REPORT OF THE INDEPENDENT COMMITTEE TO INQUIRE INTO THE PRACTICES, PROCESSES, AND PROCEEDINGS IN THE JUVENILE COURT AS THEY RELATE TO THE JOSEPH WALLACE CASES (Oct. 1, 1993).

127. *See* Valerie Finholm, *Report Raises Questions About Foster Care Organization*, HARTFORD COURANT, May 9, 1997, at A3.

128. *See* Susan Pearsall, *Just Who Should Care for the Children?*, N.Y. TIMES, April 20, 1997, at 13CN.

129. *See* Finholm, *supra* note 127.

130. Personal Communications with David Dearborn, and James Wiltse, former spokesmen, Connecticut Department of Children and Families. Child abuse and neglect fatalities in Connecticut rose from five in 1995, the year before the foster care panic, to nine in 1996; they fell to seven in 1997 and six in 1998. *See id.*

131. *See* Rachel L. Swarns, *For Children, Another Night on Office Cots*, N.Y. TIMES, Nov. 29, 1997, at B1.

132. *See id.*; Russ Buettner, *Foster Kids Glut System; Surge Worst Since Crack Heyday*, N.Y. DAILY NEWS, May 12, 1997, at 17; Russ Buettner, *Bid to End ACS Office Hell*, N.Y. DAILY NEWS, May 14, 1997 at 8.

133. *See* Rachel L. Swarns, *Agency Was Warned About Foster Mother Charged in Girl's Death*, N.Y. TIMES, July 2, 1997, at B3.

years, child abuse deaths increased.[134] They increased again in 1997 and 1998.[135] A study that looked only at deaths of children "known to the system" found that deaths had increased by fifty percent during that same time period.[136] Just as in Illinois, the death toll among children "known to the system" fell below the pre-panic level only in 1999 – by which time the panic had abated and the city was taking away fewer children.[137]

The fact that child abuse deaths actually increased after foster care panics is not surprising when one considers the real reasons children "known to the system" sometimes die. Those reasons have nothing to do with a supposed bias toward family preservation. Child welfare agencies are arbitrary, capricious, and cruel. They do leave some children in dangerous homes, even as they take others from homes that are safe or could be made safe with the right kinds of services. That is because overwhelmed, undertrained workers are given impossible caseloads and are sent out to make life and death decisions. The more the system is overloaded with children who do not need to be in foster care, the less time workers have to find children in real danger. As a result, they make more mistakes in both directions, and more children die.

The same mentality that led to the foster care panics in Illinois, Connecticut, and New York City, and the same extrapolation from a few widely publicized horror stories, ultimately forced Congress into passing ASFA. ASFA further overloads child welfare systems all over the country and places more and more children at risk. In contrast, the few places that have refused to embrace the ASFA mentality have shown real improvements. One, mentioned earlier, is Alabama. Another is Pittsburgh and surrounding Allegheny County, Pennsylvania.

In the mid-1990's, the child welfare system in Allegheny County was typically mediocre or worse. Foster care placements were soaring and those in charge insisted every placement was necessary.[138] New leadership changed all that. Since 1997, the foster care population has been cut by

---

134. In 1995, there were thirty-one "substantiated" child abuse deaths in New York City (down from sixty in 1991). The number rose to thirty-eight in 1996. Personal Communication, Terry McGrath, spokesman, New York State Department of Children and Families.

135. There were forty-six "substantiated" child abuse deaths in 1997 and forty-seven in 1998. Personal Communication with Terry McGrath, spokesman, New York State Department of Children and Families.

136. *See* Nina Bernstein, *Agency Head Defends Effort for Children*, N.Y. TIMES, July 23, 1999, at B1.

137. *See* Nina Bernstein, *Girl's Death Underscores Complexity of Child Welfare*, N.Y. TIMES, May 21, 2000, at 37.

138. *See* Allegheny County Department of Human Services, *Ensuring Permanency in Allegheny County*, *available at* http://trfn.clpgh.org/acdhs/CYF/permrep799.htm (last visited Sept. 10, 2001) [hereinafter Allegheny County].

twenty percent.[139] When children must be placed, half stay with relatives, and siblings are kept together eighty-two percent of the time.[140]

Allegheny County has done it by tripling the budget for primary prevention, more than doubling the budget for family preservation, embracing innovations like the Annie E. Casey Foundation's Family to Family program,[141] and adding elements of their own, such as housing counselors in every child welfare office so families are not destroyed because of housing problems.[142] And as in Alabama, children are safer. Re-abuse of children left in their own homes has declined. In addition, since January 1997, there has been only one child abuse fatality in a family previously known to the agency.[143]

Illinois also is turning itself around, undoing the damage wrought by the state's foster care panic. The main reason for the change was simple: the state finally realized that you get what you pay for. Until recently, Illinois paid private foster care agencies the way most states do – the agencies were paid for each day they kept a child in foster care. Thus, while the agencies were told their job was to keep children in their own homes or place them with adoptive parents, the agencies actually were paid for allowing children to languish in foster care. Under intense pressure from the Illinois branch of the American Civil Liberties Union, which had settled a comprehensive class action suit against the state some years before,[144] the state adopted a new system under which agencies are rewarded financially for achieving permanence for children and penalized for prolonged foster care.

Suddenly, the intractable became tractable, the dysfunctional became functional, and the Illinois foster care population plummeted from more than 51,000 at the height of the foster care panic, to 27,000 as of June 30, 2001.[145]

Although the Illinois child welfare agency likes to claim the turnaround is due solely to increased adoptions, which is where all the political points are to be scored, in fact, Illinois is returning more children to birth parents, and it has dramatically reduced the number of children it takes from their

---

139. *See id.*

140. *See id.*

141. The Annie E. Casey Foundation also helps to fund NCCPR.

142. *See* Allegheny County, *supra* note 138.

143. Personal communication with Karen Blumen, Deputy Director, Allegheny County Department of Human Services.

144. *See* B. H. v. McDonald, No. 88 C 5599 (N.D. Ill., filed June 9, 1988).

145. Illinois Department of Children and Family Services, Executive Statistical Summary, September 2001, *Children Placed in Foster Care, Foster Care, available at* http://www.state.il.us/dcfs/com_communications_execstats.shtml (last visited Nov. 5, 2001).

families. In 2000, the most recent year for which data are available, Illinois took away twenty-two percent fewer children than it did in 1997.[146]

Furthermore, when Illinois must place children, forty-four percent of family foster home placements are with relatives[147] -- a form of care that also has been the target of smears by family preservation opponents.[148] And more than half of the state's adoptions are, in fact, formalizations of these kinship placements.[149] All of this was done while making children safer. Today, a child in Illinois is less likely to be re-abused in his own home than he is to be abused in foster care. And, contrary to the fears of those who smear kinship care, kinship placements are safer than placements with relatives.[150]

Illinois still has a long way to go. A federal judge recently ruled in favor of plaintiffs in a class action lawsuit alleging that the state is far too quick to label people child abusers, while offering little or no effective means of appeal.[151] And all of the progress made since 1997 has done little more than undo the damage wrought by the foster care panic. But the progress illustrates how much of the child removal caused by that panic was, in fact, unnecessary. And it illustrates that, contrary to the philosophy behind ASFA, an approach that emphasizes family preservation makes children safer.

ALTERNATIVES

Although ASFA encourages states to do a lot of bad things, by and large, it does not require them. In theory, the mandates of ASFA can be met just as easily with a renewed commitment to keeping families together as they can by embracing the "take the child and run" approach. There is nothing in ASFA that prohibits states from providing rent subsidies so children are less likely to lose their parents because those parents lack decent housing. And there is nothing to prohibit states from providing day care so single working parents do not lose their children because of "lack of supervision."

Intensive Family Preservation Services now has a record of safely and

---

146. *See id.*

147. *See id.*

148. *See* BARTHOLET, *supra* note 72.

149. *See* Jason Method, *Tragedy Spurred Illinois Reforms*, ASBURY PARK PRESS, Feb. 25, 2001.

150. Illinois Department of Children and Families, *Signs of Progress in Child Welfare Reform*, *available at* http://www.state.il.us/dcfs/signsplsub.shtml (last visited Nov. 2, 2001).

151. *See* Dupuy v. McDonald, No. 97 C 4199, (N.D. Ill. filed on March 30, 2001), *available at* http://www.hslda.org/docs/news/hslda/200104041.pdf (last visited Nov. 2, 2001).

effectively helping families that dates back more than two decades. The Casey Foundation's Family to Family initiative and the Edna McConnell Clark Foundation's Community Partnerships for Child Protection are pioneering new ways to tap the strengths of individual families and whole communities to keep families safely together.

Alabama, Pittsburgh, and to some extent Illinois are beginning to discover what the framers of ASFA never understood: the claim that child protection and family preservation are at odds is a false dichotomy. These communities are learning that there is no real child protection without family preservation.

Juvenile Delinquent Rehabilitation: Placement of Juveniles Beyond Their Communities as a Detriment to Inner-City Youths

Theresa A. Hughes*

INTRODUCTION

Unfortunately, life's opportunities are too often shaped by the neighborhoods we are raised in; clearly, there is a vast disadvantage for poor, inner-city, minority children. At the tender age of ten, Jose's future was drastically altered. His daily walk home from fifth grade on his way to the basketball court at Public School 384 carried him through the intersection of Schaeffer and Wilson Streets, in the notorious high crime neighborhood of Bushwick in Brooklyn, New York.[1] "An old guy," about thirty years of age, took notice of Jose, talking to him intermittently as he played basketball.[2] In time, the "old guy," later revealed to be "Mohammad," would regularly pass Jose a $50 or $100 bill "just for chillin' at the hoop."[3] This is a considerable amount of money to a financially underprivileged child, especially considering that it would take Jose's Latino immigrant father days to earn such salary at his manual labor job doing electrical work.[4]

Inevitably, Mohammad proposed what appeared to be an undemanding favor of his new 10-year-old friend, asking that Jose merely signal Mo-

---

* Hofstra University School of Law, Supervising Attorney, Child Advocacy Clinic, City University of New York School of Law, J.D. (1995), City University of New York at Brooklyn College, M.A. (1991), Manhattan College, B.A. (1987). The author extends heartfelt appreciation to Christopher Fanning for his continuous and unparalleled support, dedication, and encouragement. Sincere gratitude is offered to Jean Doherty for extraordinary technical expertise. Special acknowledgment to "Jose" for his inspiration, cooperation, and resilience.

1. Interview with "Jose" (alias used to preserve attorney-client confidentiality) at City-Challenge Aftercare Services, Office of Children and Family Services in Brooklyn, NY (June 26, 2000) [hereinafter Jose interview].
2. *See id.*
3. *See id.*
4. *See id.*

hammad, by tipping his baseball hat, from the basketball court if he happened to see a police car approaching.[5] Jose laughed nervously, "it was easy to see the cops in those big Chevrolets."[6]

Predictably, Jose became Mohammad's regular lookout for beat cops and undercover drug officers.[7] Jose was gently persuaded and gradually welcomed into Mohammad's drug family, and began to become familiar with the customers.[8] "Mohammad is like the older brother I ain't got."[9] Mohammad by now became "my man."[10]

Jose's appearance presented the ideal lookout for drug dealing, as he was quite small in stature, and as reports later reflected in his New York City Police Department code name, "J.D. [John Doe] Baby Face."[11] "I'm too little so the police don't pay no mind to me."[12] Early in his lookout days, Jose regularly attended school, inspired by Mohammad's daily $50 incentive, knowing that a truant officer would bring attention to the drug business.[13]

As Jose showed his allegiance, his responsibilities grew.[14] Soon equipped with a walkie-talkie and taught to use code words from the basketball court, like "Tohando," to inform Mohammad that an unfamiliar face, possibly a decoy, was near, Jose was given a standard work schedule.[15] "My man and their cats had different shifts, I'd be the one looking out and gettin' $200, $250, or $300 a day."[16] According to Jose, Mohammad ironically held a full-time day job as a counselor for juvenile delinquents at a local group home.[17]

With promises that Mohammad would "take good care" of him if ever arrested, Jose was persuaded, at the age of thirteen, to start hustling drugs himself.[18] Jose proved successful on the streets, as the police overlooked him for the older juveniles and adult dealers, his babyish look making him virtually invisible.[19] Every school day, from 3:00 - 7:00 p.m., and on Saturdays, from 7:00 a.m. onwards, Jose sold "bags and G-Packs (10 packs)" of

---

5. *See id.*
6. *Id.*
7. *See* Jose interview, *supra* note 1.
8. *See id.*
9. *Id.*
10. *Id.*
11. *Id.*
12. *Id.*
13. *See* Jose interview, *supra* note 1.
14. *See id.*
15. *See id.*
16. *Id.*
17. *See id.*
18. *See id.*
19. Jose interview, *supra* note 1.

heroin.[20] "I'd see them everyday so if I saw an unfamiliar face, I'd know it was a decoy and I wouldn't serve him. The customer would have to tell me what the name stamped on the dope was like 'unabomber,' 'no name,' or 'deaf row' and if he couldn't tell me the name, he's not gettin' it. The customers were white boys from Queens and New Jersey, rich guys that had money."[21]

Jose's school attendance inevitably declined, and he was thereafter left back one grade.[22] The benefits of working on the streets outweighed any interest in elementary school lessons, as he was now making up to $600 a day.[23] "Money be flowing like it ain't nothing."[24] Deeper and deeper Jose was dragged down into the drug business, assuming more responsibility, incurring greater risks and facing greater danger.[25] Jose was asked to store bundles of heroin at his home. "Mohammad would give me a big bag of dope with plastic, wrapped up with tape. I'd have to take a cab home, because I ain't walkin' with that. One day he gave me $30,000 worth of heroin and had me hide it in my parents' house for a whole week."[26]

Jose was not a drug user, "I never tried it. I ain't doing no dope, are you craze? I look at those dope heads and feel sorry for them. Besides, Mohammad would beat me up if I ever did it."[27] Predictably, Mohammad was not the noble protector Jose had naively embraced, as when Jose was arrested on drug charges, not only once, but twice, "my man" was absent.[28]

As a law guardian, I have represented hundreds of children and juvenile delinquents in New York State Family Courts. Jose was just one of 32,299 children under the age of sixteen arrested in New York that year.[29]

Jose was a typical client: an inner city kid growing up in a tough and poor neighborhood where gangs, guns, drugs, and violent crime are common.[30] Although he came from a two-parent household, ironically with a sister employed by Child Protective Services, Jose lived in a neighborhood ravaged by crime, and the Bushwick ghetto drug dealers penetrated the thick wall of family values, as they lured this ten year old child into a life of

---

20. *See id.*
21. *Id.*
22. *See id.*
23. *See id.*
24. *Id.*
25. *See* Jose interview, *supra* note 1.
26. *Id.*
27. *Id.*
28. *See id.*
29. Facsimile from Marge Cohen, New York State Division of Criminal Justice Services, Office of Justice Systems Analysis/Bureau of Statistical Services to Theresa Hughes, (Mar. 26, 2001) (on file with author).
30. *See* New York City Police Dep't, 83rd Precinct, *CompStat at* http://www.ci.nyc.ny.us/html/nypd/html/pct/cspdf.html (last visited Mar. 26, 2001).

crime.[31] Living in a neighborhood where there are high levels of crime and poverty increases the risk of involvement in serious crime for all children growing up there.[32] In Jose's case, it is essentially elements within the community that are responsible for the stripping of innocence and placing a child at high risk for committing crimes.

Subsequent to court proceedings, juvenile delinquency findings were made to drug charges, and the Department of Probation recommended that Jose be placed in a moderately structured setting.[33] Although probation stated that Jose only needed close supervision, counseling, and assistance with academics, and although Jose had no other prior delinquency findings, a Brooklyn Family Court judge denied Jose the opportunity to remain in his own neighborhood and enter into a community-based rehabilitation program which had accepted him.[34] Jose was ordered by the Court to be removed from the community and placed in public confinement hours north of, and a world apart from, his Bushwick community for up to an eighteen month period.[35] While serving his sentence, according to Jose, he attended school and played basketball, receiving no individual or group counseling, or vocational training.[36] Due to the mandated closing of the detention facility where Jose resided, approximately sixteen months later, two months short of the anticipated release, he was returned to aftercare within his community, where he was to report after school for the next eight weeks until his sentence was completed.

After being out of his neighborhood for such a considerable period of time, while being driven home to his parents' apartment in Bushwick, he passed the intersection of Schaeffer and Wilson Streets only to feel the familiarity, as he saw Mohammad in the same position which he had left him sixteen months prior, across from the basketball court, while the others continued to hustle drugs.[37] Of course, not much had changed, and the same problems that originally got Jose into trouble still existed.[38]

When I met Jose during the first week of his release to aftercare, the months he spent away had shown a good deal of change, as his baby face features had been replaced with tougher, detached, and more street smart features. Although he vowed that he was not going to be hustling, the

---

31. *See* Jose interview, *supra* note 1.

32. *See* PANEL ON JUVENILE CRIME: PREVENTION, TREATMENT, AND CONTROL, JUVENILE CRIME JUVENILE JUSTICE 79-83, 89-100 (Joan McCord et al. eds., 2001) [hereinafter JUVENILE PANEL].

33. *See* Jose interview, *supra* note 1.

34. *See id.*

35. *See id.*

36. *See id.*

37. *See id.*

38. *See id.*

temptation would repeatedly pose itself.[39] McDonald's was only offering $5.15 an hour, and his former high-paying job on the streets of Bushwick was still available.[40] Jose had not been rehabilitated from a life of delinquency, rather merely stowed away in a lock-up facility far away from his own neighborhood, as time passed in Bushwick awaiting his return.

PREMISE

As there are children at this very moment who are engaging in drug dealing, gang activities, and various other acts of delinquency, and as the forecast predicts that the youth population will increase over the next decade, there is virtually no one who does not want to see the level of juvenile crime decrease.[41] The debate rests in the method by which the social order attempts to decrease such crime. Should juvenile delinquency be treated within or outside of the child's community? If juvenile crime is directly related to social welfare issues surrounding these children's lives, such as poverty, inadequate housing and education, racism, child abuse, and drug addiction, "[e]radicating juvenile crime without addressing [the] relationship to these other issues is an impossibility."[42]

When the juvenile justice court was founded in the United States over one hundred years ago, it generally rested on the belief that children lacked the moral maturity of adults, and, as such, children were not entirely responsible for their actions.[43] Children were seen as being quite different from adults and needed to be protected, nurtured, and treated, rather than held completely responsible and punished for their acts.[44] The courts and society were sympathetic to the external influences, almost always circumstances beyond the child's control, which often propelled the child into criminal activity, such as poverty, abuse, and neglect.[45] The response to juvenile crime was not to punish, but rather to rehabilitate the child by availing him to treatment services designed to help him cope with negative social pressures in non-delinquent ways, because children could be turned from their

---

39. *See* Jose interview, *supra* note 1.

40. *See* U.S. Dep't of Labor, *The Minimum Wage at* http://www.dol.gov/dol/esa/public/minwage/main.htm (last visited Mar. 9, 2001).

41. *See* NAT'L CRIM. JUST. ASS'N, U.S. DEP'T OF JUSTICE, JUVENILE JUSTICE REFORM INITIATIVES IN THE STATES: 1994-1996 Program Report at iii (Oct. 1997).

42. Jennifer M. O'Connor & Lucinda K. Treat, *Getting Smart About Getting Tough: Juvenile Justice and the Possibility of Progressive Reform*, 33 AM. CRIM. L. REV. 1299, 1300 (1996).

43. *See id.* at 1303.

44. *See id.*

45. *See* Julianne P. Sheffer, Note, *Serious and Habitual Juvenile Offender Statutes: Reconciling Punishment and Rehabilitation within the Juvenile Justice System*, 48 VAND. L. REV. 479, 482 (1995).

deviant ways, not by threats of punishment, but by changing the child's thinking, goals, and values.[46]

However, the philosophical basis of the juvenile justice system has shifted, from the late nineteenth century common interest of rehabilitating the juvenile delinquent to the current "lock 'em up" culture.[47] Society has become preoccupied with punishment, and historical non-punitive rehabilitative notions have not been realized.[48] At some point society lost sight of the original goal of reducing juvenile crime and refocused on punishing the child for his wrongs. The youth are sent to state facilities with hostile environments, receiving little to no rehabilitation, and then "later released from incarceration where they are reintroduced to a society in which they cannot succeed."[49] The "lock 'em up" approach fails to significantly reduce rates of recidivism and in turn, recycles unhealed juvenile delinquents onto the streets from which they came.[50]

Many of the neighborhoods to which juveniles are returning in New York City are crime infested. Along with the many studies that identify inequality as a strong indicator of crime, the most noticeable signs of inequality, impoverishment, and despair are undoubtedly seen in the urban ghetto, which defines the high crime area.[51] Much of the problem with juveniles repeatedly committing acts of delinquency stems from the economic, geographic, and racial inequalities born by children on these inner-city streets.[52] New York City now holds an overall poverty rate of twenty-five percent, with thirty-seven percent of children living in poverty. The borough of Brooklyn boasts an appalling fifty-two percent of New York City homicide suspects being persons under age twenty-one.[53]

Some juvenile delinquency researchers have theorized that the increase in juvenile crime can be attributed to the mid 1980's inner-city crack epi-

---

46. *See id.*

47. *See generally* NAT'L CRIM. JUST. ASS'N, *supra* note 41, at 3.

48. *See* Major Richard L. Palmatier, Jr., *Criminal Offenses by Juveniles on the Federal Installation: A Primer on 18 U.S.C. § 5032*, ARMY LAW., Jan. 1994, at 3.

49. Amy M. Campbell, Survey, *Trying Minors as Adults in the United States and England: Balancing the Goal of Rehabilitation with the Need to Protect Society*, 19 SUFFOLK TRANSNAT'L L. REV. 345, 358 (1995).

50. *See* BRUCE FREDERICK, NEW YORK STATE DIVISION OF CRIMINAL JUSTICE SERVICES, FACTORS CONTRIBUTING TO RECIDIVISM AMONG YOUTH PLACED WITH THE NEW YORK STATE DIVISION OF YOUTH, 13 (1999).

51. *See* Abbe Smith, *They Dream Of Growing Older: On Kids and Crime*, 36 B.C. L. REV. 953, 967-68 (1995).

52. *See* JUVENILE PANEL, *supra* note 32, at 3.

53. *See* Poverty in New York City: A CSS Data Brief #1, Oct. 1999, *Brooklyn D.A.: Idle Hands Can Become Killers available at* http://www.apbnews.com/newscenter/breakingnews/2000/04/11/brooklyn0411 (last visited Mar. 10, 2001).

demic, holding that the expansion of the crack cocaine market led to drug dealers recruiting children as low-level sellers, carriers, and lookouts.[54] These children were recruited because they worked for less, took greater chances, and were more likely to escape detection and punishment.[55] These juveniles carried guns and other weapons which thereafter became staples in gang and drug market battles.[56]

As the rate of juvenile crime increased dramatically in the mid-1980's through the early 1990's, the rate of juvenile arrests increased.[57] From their neighborhoods within the urban ghetto of New York City, juvenile delinquents have been regularly removed by the courts and placed in institutional confinement, oftentimes in rural settings.[58] Without the familiarities of their families and schools and facing the daily negative peer and criminal influences, "from a rehabilitative perspective, there are inherent limits to what can be accomplished in such an artificial environment."[59] Such confinement and isolation from the community may punish the child, but only exacerbate the problem of juvenile crime, as evident from recidivist rates.[60]

Moreover, it costs a taxing $244.65 per day, or $89,297.25 per year, to house a youth in a New York State secure detention facility[61] – about three times as much as it costs to send a student to an Ivy League college.[62] In contrast, the cost of placing a child in a model community-based alternative-to-detention program in New York City is approximately $1,800 per year, a mere $4 per day.[63]

It is not only the youths who are troubled, but also the inner-city communities in which they live. In order to make our communities safer for all citizens, adults, and children alike, we must properly address the need for effective rehabilitation. Healing crime infested inner-city communities may not be a viable short-term objective. However, the goal of reducing crime can be reached through rehabilitation training while the juvenile lives within his own neighborhood, and has to face and reject the daily lure of criminal activity. In addition, it is crucial to give children real alternatives to selling

---

54. *See* NAT'L CRIM. JUST. ASS'N, *supra* note 41, at 4.
55. *See* JUVENILE PANEL, *supra* note 32, at 4.
56. *See id.* at 97-100.
57. *See id.* at 2.
58. *See* FREDERICK, *supra* note 50, at 22.
59. *Id.*
60. *See id.* at 1.
61. Telephone Interview with Pat Cantiello, Public Information Specialist, New York State Office of Children and Family Services, Public Information Department (Mar. 28, 2001).
62. *Justice Fellowship*, *at* http://www.justicefellowship.org (last visited Mar. 27, 2001).
63. *See* ANDREW GLOVER YOUTH PROGRAM, NEW YORK, NEW YORK, BROCHURE, 23 (Jan. 1999) [hereinafter ANDREW GLOVER BROCHURE].

drugs and engaging in delinquent activity.[64]

With an emphasis on New York City, this article will explore: the juvenile's right to effective rehabilitation; how this right is being overshadowed by society's demand for pure punishment; how predominant juvenile rehabilitation in residential state placement is failing youth as well as society; the most effective methods of discouraging juvenile delinquency as demonstrated by existing alternative-to-detention programs; and finally, this article will offer suggestions for effective community-based rehabilitation for inner-city youths.

THE JUVENILE'S RIGHT TO REHABILITATION IN NEW YORK STATE

In theory, once the state takes custody of a juvenile delinquent, "it must provide treatment to effectuate that rehabilitation."[65] In New York State, the Family Court has jurisdiction to take custody of delinquent children, between the ages of seven and fifteen, who have been found to have "committed an act that would constitute a crime if committed by an adult."[66]

Juvenile delinquency proceedings in New York State Family Court are considered quasi-criminal in nature and effect, and consequently must comply with the same due process safeguards applicable to adults in criminal proceedings. Therefore, children are entitled to full trials, consisting of a fact-finding and dispositional hearing.[67] Of concern to this article is the dispositional stage, wherein the court decides the appropriate placement for the child.[68] At the dispositional stage, the court must not think in terms of guilt or innocence, but of the child's need for protection and rehabilitation, considering more than the delinquent act itself no matter how extreme or violent it may have been.[69] The court has a great deal of discretion in deciding whether the child's needs should be met through supervision, treatment, or confinement.[70] However, the court must consider the "least restrictive alternative" when deciding whether to sentence the child to placement or probation, and to order a conditional discharge.[71]

Historically, the juvenile justice system's sentencing procedure, branded

---

64. *See* Smith, *supra* note 51, at 961.

65. Paul Holland & Wallace J. Mlyniec, *Whatever Happened to the Right to Treatment?: The Modern Quest for a Historical Promise*, 68 TEMP. L. REV. 1791, 1792 (1995).

66. N.Y. FAM. CT. ACT §§ 301.2(1), 302.1 (McKinney 2000); *see also* Martarella v. Kelley, 349 F. Supp. 575, 579 (S.D.N.Y. 1972).

67. *See* N.Y. FAM. CT. ACT §§ 305.3, 305.4, 351.1 (McKinney 2000).

68. *See* N.Y. FAM. CT. ACT §§ 353.3-.5 (McKinney 2000).

69. *See In re* Wooten, 284 A.2d 32, 34-35 (1971).

70. *See* N.Y. FAM. CT. ACT § 352.2 (McKinney 2000).

71. *See* N.Y. FAM. CT. ACT § 352.2(2) (McKinney 2000).

"disposition," was rendered following a relatively informal and flexible conference with the probation officer, the juvenile and his parent or guardian.[72] Steering the child away from a life of crime was the goal of the progressive reformers who led the juvenile court movement, with the view that a juvenile court judge was not to adjudicate and sentence the youth, but was to identify the conditions which had led him astray and to treat the child for those conditions. Thus, under the doctrine of *parens patriae* – "the State as parent," the original juvenile court statutes promised that the child who was removed from his family by a judge would receive the care, custody, and discipline that his parents should have provided.[73]

In contrast to criminal sentencing, which is substantially retributive in nature, the current delinquency dispositional phase is in theory still considered essentially rehabilitative in nature.[74] However, in 1976 the purpose of the New York State juvenile delinquency statutes was given a new thrust in that the court was now mandated to weigh the competing considerations of the needs and best interests of the juvenile with the need for protection of the community.[75] The inclusion of this new criteria – protection of the community – in effect weakened, if not directly conflicted with the historical purpose of the juvenile's needs and best interests,[76] creating a smoke screen for the unofficial third criteria: punishment of the juvenile.

HISTORICAL SHIFT FROM REHABILITATION TO PUNISHMENT

> [W]hat is strikingly clear from th[e] research is that the headlong rush to ever-greater incarceration in the name of 'getting tough on young thugs' is unjustified. For all but the truly violent few, investing in a continuum of graduated care makes better sense in every dimension – for our youth, for our communities, and, not least, for our pocketbooks.[77]

---

72. *See* Douglas J. Besharov & Merril Sobie, *Introductory Practice Commentaries*, N.Y. FAM. CT. ACT § 301.1 (McKinney 1999); *see also* Holland & Mlyniec, *supra* note 65, at 1791; Jessica Hardung, Comment, *The Proposed Revisions to Japan's Juvenile Law: If Punishment is Their Answer, They Are Asking the Wrong Question*, 9 PAC. RIM L. & POL'Y J. 139, 143-45 (2000).

73. *See Child or Adult? A Century Long View*, *at* http://www.pbs.org/wgbh/pages/frontline/shows/juvenile/stats/childadults.html (last visited Feb. 1, 2001); *see also* Holland & Mlyniec, *supra* note 65, at 1792.

74. *See* Besharov & Sobie, *supra* note 72.

75. *See* N.Y. FAM. CT. ACT §§ 301.1, 352.2, 301-385 (McKinney 1999); Besharov & Sobie, *supra* note 72; *see generally In re* Steven E.H., 477 N.Y.S.2d 563 (1984); *In re* Rudy S., 420 N.Y.S.2d 549 (1979).

76. *See generally* Linda F. Giardino, Note, *Statutory Rhetoric: The Reality Behind Juvenile Justice Policies in America*, 5 J.L. & POL'Y 223 (1996).

77. Barry Krisberg et al., *What Works with Juvenile Offenders? A Review of 'Graduated Sanction' Programs*, 10 CRIM. JUST. 20, 61 (1995).

Perhaps juvenile justice policy is itself delinquent or possibly the goal has actually been intentionally shifted away from the rehabilitative philosophy of the founders. What began as an American system designed to rehabilitate juveniles is moving increasingly toward a conservative system that punishes them.[78] Theories offered to justify the shift from rehabilitation to punishment include: punishment is a deterrent to future juvenile offenders; punishment incapacitates juvenile offenders and prevents them from committing future offenses; and punishment satisfies society's desire for accountability and retribution.[79] Especially during the last decade, society has rejected the juvenile court's traditional ideal of rehabilitation in favor of more punitive, offense-oriented sanctions, and, in effect, blurred the lines between juvenile and adult justice systems.[80] Even the Juvenile Justice and Delinquency Prevention Act of 1974, which focused on the needs of the youth, has been replaced by the demand for a more penalizing response, focusing on the crime alleged to have been committed.[81]

This new "get tough" approach is in direct contradiction to recent research. This research illustrates that it may be counter-productive to treat juveniles as adults.[82] Further, even though an ever-increasing number of juveniles are being detained and incarcerated, there is "evidence that most juveniles can be treated equally or more effectively in the community than in secure confinement, without jeopardizing community safety."[83]

In reality, this newly adopted "get tough" approach appears to be justifying punishment not on the grounds that it will deter, prevent, or rehabilitate an offender, but rather on the ground that the delinquent deserves to be incarcerated as appropriate punishment.[84] However, this theory flies in the face of the purpose of the juvenile justice system, reducing it to the mere vindication of private wrongs, and mirroring the adult court system.[85]

It is clear from statistics that stiffer sentences for juveniles result in increased crime.[86] However, with overwhelming pressure from society to "get tough," it is unlikely that our elected officials will lead the way to increase funding for alternatives-to-detention for juveniles.[87] "In State capitols [as

78. *See Child or Adult? A Century Long View*, *supra* note 73.

79. *See* Joseph F. Yeckel, Note, *Violent Juvenile Offenders: Rethinking Federal Intervention in Juvenile Justice*, 51 WASH. U.J. URB. & CONTEMP. L. 331, 333 (1997).

80. *See* JUVENILE PANEL, *supra* note 32, at 4; Sheffer, *supra* note 45, at 481.

81. *See* Juvenile Justice and Delinquency Prevention Act of 1974, Pub. L. No. 93-415, 88 Stat. 1109 (codified as amended in scattered sections of 5, 18 and 42 U.S.C.); JUVENILE PANEL, *supra* note 32, at 7.

82. *See* JUVENILE PANEL, *supra* note 32, at 4.

83. *Id.*

84. *See* Sheffer, *supra* note 45, at 481.

85. *See In re* R., 323 N.Y.S.2d 909 (1971).

86. *See* Campbell *supra* note 49, at 358.

87. *See* NAT'L CRIM. JUST. ASS'N, *supra* note 41, at 8.

well,] it's difficult to support expenditures that might reduce crime and prison costs in years to come when voters are clamoring for action now."[88]

Whether this recent trend toward stiffer sentences is a direct reaction to nationally publicized high-profile cases of violent juvenile crime, the result of a society that has grown less sympathetic and tolerant, or stems from other reasons altogether, the impact is ultimately on those in need of rehabilitation.[89] Either way, sheer punishment virtually ignores the goal of restoring juveniles to being productive members of society.[90] The instinct to punish is not about the failure of rehabilitation after years of experience. Many suppose that we have never made a serious, nationwide attempt to rehabilitate juvenile offenders, and further many of the few successful statewide rehabilitation programs have since lost their funding. [91]

The trend towards punishment and increased detention has resulted in neither safer New York communities, nor more effective juvenile rehabilitation. The juvenile housed in a facility that focuses on retribution is more likely to re-offend than one who is placed in a center with a goal of rehabilitation.[92] Traditional juvenile justice systems that remove the children from the community into detention facilities have failed to adequately rehabilitate juvenile offenders.[93]

DISPOSITION NEW YORK STYLE

When youth are adjudicated as juvenile delinquents and the prosecutor and/or the Department of Probation are recommending placement, decisions about such a placement are made by a private or state agency. Of the two, the preferred choice of the youth is virtually always a private agency. Nearly all private agencies, such as Lincoln Hall, located forty miles north of New York City in Westchester County, and St. Mary's Children and Family Services, are mandated residential and are located in non-urban settings. They are also less restrictive than state placements.[94] State facilities are more secure, often with the stereotypical barbed wire and fencing.[95] If

88. State Legislatures, *Ounce of Prevention*, 14-16, National Conference of State Legislatures, Denver, Colo. (May 1995).

89. *See* Jane Gross, *Sentencing to Better Their Lives; For Young Criminals, Focus Is Saving, Not Punishing*, N.Y. TIMES, July 5, 1997, at B25.

90. *See* Smith, *supra* note 51, at 960.

91. *See* FREDERICK, *supra* note 50, at 5-8.

92. *See* Hardung, *supra* note 72, at 159.

93. *See* Yeckel, *supra* note 79, at 336.

94. *See* Lincoln Hall Web Site, *Our History at* http://lincolnhall.org/history.htm (last visited Feb. 5, 2001); *see also* ST. MARY'S CHILDREN AND FAMILY SERVICES, SYOSSET, NEW YORK, BROCHURE (undated).

95. Telephone Interview with Joseph Dennison, Director, Office of Children and Family Services Bronx Residential Center, New York, New York (Feb. 2, 2001)

the youth is on remand status pending the outcome of disposition, he usually will be held in one of the local inner-city detention facilities such as Spofford, Bridges or Crossroads.[96] Here the youth is interviewed by the private facility's social worker, often times at the request of defense counsel, in hopes of securing a bed/placement in one of the private facilities. If the youth is rejected by a private facility, it is generally because either the underlying delinquent act was too serious, the youth is perceived as too aggressive to mix with the current facility's population or, most unfortunately, because the youth's family is dysfunctional, availing no viable release plan.[97]

If the youth is rejected from private placement, the court will usually place the adjudicated youth in the custody of New York State with the Office of Children and Family Services (OCFS), the organization charged with the care and "rehabilitation" of the child.[98] Once placed in a state facility, the youth is committed to OCFS custody for a definite term of twelve to eighteen months, renewable upon court order, in either a secure, limited secure or non-secure detention facility, a group home, day treatment or evening reporting center. Despite the thousands of New York City adjudicated delinquents, there are currently very few OCFS group homes offered. Currently, within New York City, OCFS only offers as "non-security community-based facilities" three evening reporting centers, two homes for pregnant girls and a sole group home and a City Challenge Home in Brooklyn, with no group home offerings in the boroughs of Manhattan and Queens.[99]

Whether in a state (OCFS) or private placement, these facilities are considered non-community-based, and the youth is mandated to reside at the facility on a full-time basis.[100] The youth does not live with his family, nor does he attend school or work within his own community despite data demonstrating that "incarcerated juveniles have higher rates of physical injury, and mental problems, and they have poorer educational outcomes, than do their counterparts who are treated in the community."[101]

[hereinafter Dennison interview].

96. *See id.*

97. *See id.*

98. *See id.*

99. *See* Dennison interview, *supra* note 95; Facsimile from the New York State Office of Children and Family Services to Theresa Hughes, Supervising Attorney, *Hofstra University School of Law Child Advocacy Clinic*, (Feb. 12, 2001) (on file with author).

100. *See* Dennison interview, *supra* note 95.

101. JUVENILE PANEL, *supra* note 32, at 5.

RACE, GEOGRAPHY, GENDER AND ECONOMICS

> A number of recent surveys have shown that there are profound racial disparities in the juvenile justice system, that Black and Hispanic youth are more likely to be tried as adults, receive longer sentences, and more likely to be in locked facilities.[102]

Race, geography, gender and economics are clearly factors in the likelihood of youth being brought into the juvenile justice system. These factors raise equality issues and have long-term fairness implications.[103] Data clearly shows that for most jurisdictions across the country, minority youths (especially Black) are disproportionately over-represented within the juvenile justice system.[104] For example, while Blacks constituted twelve and a half percent of the population in 1994, they accounted for nearly twenty-nine percent of the juveniles arrested, with more than half of the arrests for violent crime, including fifty-nine percent of the juvenile homicide arrests nationally.[105] In 1997, there were 1018 Black juveniles in residential placement for every 100,000 Black youths in the population compared to only 204 White youths per 100,000.[106] In 1998, although only fifteen percent of those under age eighteen were Black, these youths made up forty-two and three-tenths percent of the juvenile arrests for violent crime.[107] Along with being disproportionately arrested, minority youth are more likely to be placed in secure detention facilities, while White youth are more likely to be housed in private facilities or diverted from the juvenile justice system altogether.[108]

In a review by Pope and Feyerherm, it was found that juvenile court cases in urban settings are more likely to receive severe outcomes at various stages of the judicial process than are cases in non-urban areas, and because minority populations are concentrated in inner-cities, the effect may work to the disadvantage and overrepresentation of minority youth.[109]

In 1997, minorities accounted for sixty-three percent of juveniles committed nationally to residential placement, with only thirty-seven percent of

---

102. *See* PBS Web Site, *Is the System Racially Biased?* at http://www.pbs.org/wgbh/pages/frontline/shows/juvenile/stats/childadult.html (last visited Feb. 1, 2001).

103. *See* NAT'L CRIM. JUST. ASS'N, *supra* note 41, at 3.

104. *See* HOWARD N. SNYDER & MELISSA SICKMUND, U.S. DEP'T OF JUSTICE, JUVENILE OFFENDERS AND VICTIMS: 1999 NATIONAL REPORT 193 (1999).

105. *See* NAT'L CRIM. JUST. ASS'N, *supra* note 41, at 4.

106. *See* SNYDER & SICKMUND, *supra* note 104, at 193.

107. *See* JUVENILE PANEL, *supra* note 32, at 6.

108. *See* SNYDER & SICKMUND, *supra* note 104, at 193.

109. *See id.* at 193.

Whites entering placement.[110] In New York State, 152 White youths, per 100,000, were placed for every 1936 minorities placed.[111] The overrepresentation of minority youths within the juvenile system is partially explained by the risk factors associated with crime, such as living conditions, with more minority children residing in communities characterized by concentrated poverty and social disorganization.[112]

Although New York State's Family Court has softened its language, and prefers employing words such as "placement" rather than "incarceration," the reality is that 5891 youths were in fact incarcerated in OCFS facilities during the year 1998. They were taken out of their homes and communities and placed away in residential facilities without having the freedom to leave.[113] Nationwide too many children are being placed in these detention facilities, which are not providing needed treatment and rehabilitative services.[114] Staff are often underpaid and untrained to treat the youths properly.[115]

During 1998, OCFS saw 2382 new admissions, a five percent increase in placements since the prior year.[116] This amounts to forty-two of every 10,000 youths between the ages of twelve and seventeen having contact with the agency in New York State.[117]

Even though New York City comprises only five of sixty-two counties within New York State, these five lead all counties in juvenile placement admissions.[118] Although only forty percent of New Yorkers live within the city, [119] New York City youth were two and one-half times more likely to end up in an OCFS facility.[120] This translated into sixty-five percent of all

---

110. *See id.* at 195.

111. *See* SNYDER & SICKMUND, *supra* note 104, at 197.

112. JUVENILE PANEL, *supra* note 32, at 228-29, 238-42.

113. *See* NEW YORK STATE OFFICE OF CHILDREN AND FAMILY SERVICES, YOUTH IN CARE: 1998 ANNUAL REPORT, at ii (1999) [hereinafter YOUTH IN CARE REPORT].

114. *See* Stacey Gurian-Sherman, *Back to the Future: Returning Treatment to Juvenile Justice*, 15 CRIM. JUST. 30, 30 (2000).

115. *See id.*

116. *See* YOUTH IN CARE REPORT, *supra* note 113, at ii.

117. *See id.*

118. *See id.* at 6; *see also County Population Estimates for July 1, 1999 and Population Change for July 1, 1998 to July 1, 1999*, POPULATION ESTIMATES PROGRAM, POPULATION DIVISION, U.S. CENSUS BUREAU (Mar. 9, 2000) *available at* http://www.census.gov/population/estimates/county/co-99-1/99C1_00.txt (last visited Mar. 10, 2001).

119. *See* Susan A. Riedinger et al., *Income Support and Social Services for Low-Income People in New York*, Highlights, New York ISSS (The Urban Institute), (June 1999) at 1 *available at* http://newfederalism.urban.org/html/highlights/isss_ny.html (last visited April 13, 2001).

120. *See* YOUTH IN CARE REPORT, *supra* note 113, at 38.

youth taken into custody in 1998 being from New York City.[121]

Locked-up youth in New York are primarily inner-city Black and Latino males. Of these 2382 New York delinquents newly admitted into OCFS custody, a towering eighty-six percent were non-white, while only 336 were White youths.[122] Of the 1998 juvenile delinquents in New York State OCFS custody, males were the disproportionate majority over females, at the rate of eighty-six percent over fourteen percent, with only 335 females.[123] However, nationally, the rate of increase in arrest and incarceration has been larger in recent years for girls than boys, and the seriousness of the crimes that girls have committed has also increased.[124] Of the 2.5 million juvenile arrests nationally in 1999, twenty-seven percent were females.[125]

FAMILY AFFECT

Of the 1998 New York State facility admissions, seventy-two percent of the youths came from households with only one or no parent present.[126] More than half of all juvenile delinquents imprisoned in state institutions have immediate family members who have also been incarcerated.[127] Poor parenting practices have been found to be important risk factors for juvenile delinquency, such as: absence of parental supervision; exposure of the child to overt conflict or child abuse and neglect; inappropriate discipline; lack of emotional warmth; and parental stress.[128]

RECIDIVIST RATES – DETENTION AS AN ARTIFICIAL SETTING

> "We need to educate them and socialize them. Take them away from home and reprogram them, if you will."[129]

Despite the above callous analogy of Peter Reinharz, juvenile delinquents apparently do not respond well to encoding. It is evident that the recent focus on punishing juveniles in detention facilities away from their homes has

---

121. *See id.* at ii, 38 (1515 inner-city delinquents were taken into custody, as opposed to only 841 upstate New York delinquents being taken into custody).

122. *See id.* at 2, 4.

123. *See* YOUTH IN CARE REPORT, *supra* note 113, at 2.

124. *See* JUVENILE PANEL, *supra* note 32, at 11.

125. *See* Office of Juvenile Justice and Delinquency Prevention, *Juvenile Justice Facts and Figures*, *available at* http://ojjdp.ncjrs.org/ojstatbb/qa250.html (last visited March 27, 2001).

126. *See* YOUTH IN CARE REPORT, *supra* note 113, at ii.

127. *See* ANDREW GLOVER BROCHURE, *supra* note 63, at 9.

128. *See* JUVENILE PANEL, *supra* note 32, at 66-69.

129. Gross, *supra* note 89, at B25 (quoting Peter Reinharz, Chief Prosecutor, New York City Family Court).

resulted in increased recidivism rates and a lower likelihood that youths will be educated and rehabilitated.[130] In the past two decades the rate of juveniles placed nationally in institutions has increased substantially.[131] Locally, in New York State, most of the youths placed in OCFS custody were proven recidivists, at an alarmingly high rate of eighty-one percent of males and forty-five percent of females being arrested within thirty-six months of discharge from state custody.[132] With statistics such as these, the perception is that rehabilitation has failed.[133]

The Office of Justice Systems Analysis Report, "Factors Contributing to Recidivism Among Youth Placed with The New York State Division for Youth," attributes two potential problems consistent with high recidivism rates.[134] The first problem is the inconsistency among staff within service settings. The second problem is the discontinuity of programming among service settings, especially in the transition from residential confinement to aftercare.[135] Both residential staff and aftercare staff cite the lack of continuity between residential programming and aftercare as a problem potentially contributing to high recidivism rates. Residential staff suggested that their efforts may be ineffective because they do not carry through on aftercare, and because the youth faced difficult circumstances upon returning to their home communities.[136] For example, youths are released from confinement and sent back into their communities, losing attachments made during their confinement and having been taught non-delinquent ways in a synthetic environment that is non-reflective of their own neighborhood.[137] In order to be effective, the juvenile justice system must focus on the socio-economic background of the child.[138]

Youths are not readapting to their neighborhood surroundings upon being released from placement.[139] When these youths are transferred from a highly artificial and structured environment to standard community supervision, the constant, external control abruptly disappears and the youth is at a loss.[140] Many "experts now recommend a more gradual transition into the

130. *See* Hardung, *supra* note 72, at 154, 156-57.

131. *See* JUVENILE PANEL, *supra* note 32, at 6.

132. *See* FREDERICK, *supra* note 50, at 5, 21.

133. *See* Michael J. Dale, *The Supreme Court and the Minimization of Children's Constitutional Rights: Implications for the Juvenile Justice System*, 13 HAMLINE J. PUB. L. & POL'Y 199, 226-27 (1992).

134. *See* FREDERICK, *supra* note 50, at 17.

135. *See id.* at 17-18.

136. *See id.* at 18.

137. *See id.* at 21.

138. *See* Jeffrey K. Day, Comment, *Juvenile Justice in Washington: A Punitive System in Need of Rehabilitation*, 16 U. PUGET SOUND L. REV. 399, 412 (1992).

139. *See* FREDERICK, *supra* note 50, at 20-21.

140. *See id.*

community," with a period of intensive supervision.[141] However, intensive supervision that focuses primarily on surveillance has not been shown to decrease recidivism.[142]

If the long-term goal is to return these children to their families, the transition to the community must either start at an earlier stage or the transition must be eliminated with rehabilitation starting and ending within the child's own community.

In contrast to the high OCFS recidivist rates among youths in detention, "[m]ale [juveniles] who received their primary residential care in community-based facilities had a rearrest rate ... which was ten to twenty percentage points lower than the rates for most individual limited secure and nonsecure centers."[143] Perhaps this is so because the lessons learned in residence, may, in fact, not be applicable in the usual social environments that youth encounter upon return. Although some of "[t]he skills and attitudes promoted in residential programs may be useful in conventionally-oriented environments such as school or work; they may or may not be applicable at home; and they may be dangerous to the youth in relationships with peers 'on the street.'"[144] Treatment in a detention facility ignores the goal of treating the whole child in a multidisciplinary way, which includes addressing the youth's needs in both the family and community context.[145]

HOME COMMUNITY AS A FACTOR IN RECIDIVIST RATES

"[I]t is clear that youth released to some communities face a substantially greater risk of recidivism than similar youth released to other communities."[146] Where a youth resides affects the opportunities and resources that are available to him, and if surrounded by crime, poverty, joblessness and housing density, a child is more likely to see criminal behavior as an acceptable alternative.[147] It is the neighborhood where the child lives which influences the child's behavior, as it provides the most evident examples of the values that people hold.[148] Thus, communities in which criminal activity is widespread tend to establish criminal behavior as being within acceptable limits.[149] There is a clearly established relationship between criminal activity and poverty, as youth crime is notoriously concentrated in poor, urban ar-

141. FREDERICK, *supra* note 50, at 20.
142. *See id.*
143. *Id.* at 13.
144. *Id.* at 21.
145. *See* Smith, *supra* note 51, at 955.
146. FREDERICK, *supra* note 50, at 22.
147. *See* JUVENILE PANEL, *supra* note 32, at 79-80.
148. *See id.*
149. *See* JUVENILE PANEL, *supra* note 32, at 80.

eas.[150] As poverty levels increase, the rate of juvenile delinquency increases.[151] When juvenile homicide rates peaked recently, a quarter of all apprehended youths in the entire United States were arrested in just five counties, of which New York County was one.[152]

In a 1989 study comparing juvenile delinquents residing in different New York City neighborhoods, it was shown that

> youths from the [W]hite, working-class areas aged out of crime much faster than their [B]lack and Hispanic peers living in neighborhoods characterized by racial and ethnic segregation, . . . poverty, . . . joblessness, and single-parent households. The youths from the more disadvantaged areas had less access to employment and more freedom to experiment with illegal activity as a result of lower levels of informal social control in their immediate neighborhoods.[153]

The Office of Justice Systems (OJSA) Report, analyzing various New York State juvenile detention factors in 1999, speaks to factors associated with the risk of recidivism, such as gender and geographic region. "Across conditions, [one of the factors] most consistently associated with the risk of recidivism [was] . . . community characteristics."[154] Additionally, "[a]ggregate characteristics of localities were among the strongest predictors of recidivism. In every one of the twenty-two risk control models, one or more geographic factors remained significant even after accounting for available measures of youths' educational histories, family histories, household characteristics, selected mental health characteristics, and criminal histories."[155]

These statistics reinforce the notion that it is the ghetto that regularly brings into being the juvenile delinquent and that rehabilitation outside of the community sets the child up for failure upon return to his own neighborhood post detention. If the child is removed from his community for detention and subsequently replaced within the same community, facing the exact criminal temptations that originally led the child astray, reintegration is problematic and it is likely that child will become a repeat offender.

---

150. *See id.* at 89.

151. *See id.*

152. *See id.*

153. JUVENILE PANEL, *supra* note 32, at 95; *see generally* MERCER L. SULLIVAN, GETTING PAID: YOUTH CRIME AND EMPLOYMENT IN THE INNER CITY (1989).

154. FREDERICK, *supra* note 50, at 1.

155. *Id.* at 21.

SUCCESSFUL JUVENILE DELINQUENT REHABILITATION

> "If you send a kid upstate, you get a better educated criminal and a more angry person."[156]

In forming a dispositional/sentencing decision, the Family Court Judge is charged with considering the "least restrictive alternative," and may only send the youth to a restrictive placement as a last resort.[157] The current shortage, however, of alternatives-to-incarceration for New York City youths severely undercuts the decision-making process. The lack of community-based rehabilitation programs and finances to fund such programs has ultimately forced more punitive sanctions in secure facilities.[158] Upon release from one of its residential detention facilities, OCFS offers some after-care services: Evening Reporting Centers, which operate in the evenings and on weekends, while youth work or attend school in the daytime; Home-Based Intensive Supervision (H.B.I.S.), which offers supervision and services in the home community; and an Electronic Monitoring Program, which monitors adherence to curfew requirements.[159]

Although the majority of youths placed with OCFS are from New York City, only a very small number of these inner-city children have had the opportunity for one of the aforementioned aftercare programs.[160] For example, the H.B.I.S. program provides intensive supervision to youth in their communities, but at the end of 1998 only fourteen youths were participating in the program.[161] Furthermore, this program was not offered in New York City, but rather only in Monroe and Albany counties, which are located north of New York City. For New York City youth the only intensively supervised community-based program offered by OCFS is "City Challenge," a Brooklyn-based program. However, this program provided services to the insignificant number of thirty-four youths at the end of 1998.[162]

Although there were 55,360 juveniles under sixteen years of age arrested in New York City from 1995 to 1999, very few of these youth have the opportunity to attend an alternative-to-detention program due to the small

---

156. Telephone interview with Kerry Wallace, Director, Choices for Youth (Nov. 21, 2000) [hereinafter Wallace interview].

157. *See* John N. Kane, Jr., Note, *Dispositional Authority and Decision Making in New York's Juvenile Justice System: Discretion at Risk*, 45 SYRACUSE L. REV. 925, 948-49 (1994).

158. *See* JUVENILE PANEL, *supra* note 32, at 5.

159. *See* YOUTH IN CARE REPORT, *supra* note 113, at 31.

160. *See id.* at 33.

161. *See id.*

162. *See id.* at 35.

number of programs offered.[163] Although OCFS considers eleven out of their forty-two facilities to be community-based, such is a misnomer as youth are mandated to live within the facility. [164]

Genuine community-based rehabilitation programs in New York City include: the Center for Alternative Sentencing (CASES), the Church Avenue Merchants Block Association (C.A.M.B.A. a.k.a. "Choices for Youth"), the Dome Project, the Andrew Glover Project and The Youth Advocacy Project. Each of these non-profit organizations boasts high success and low recidivism rates, claiming that they are able to solve the problems that got the youths into trouble in the first place.[165] While participating and being monitored, each community-based program requires that the youth remain living at home with his family, attending his own school. Attending his local school, the youth gains academic credit unlike in an OCFS facility, where school credit does not transfer to the youth's local school upon release.

1. We run our program with a "Come-back-to-the-community-and-get-plugged-back-in-right-away" philosophy.[166]

Kerry Wallace, Director of Choices for Youth, declares that he believes in "education all the way."[167] Wallace proudly boasts that his program's recidivist rates have been impressively low – with only fifteen of the 300 youths served in the last three years re-committing delinquent acts. Success of the program can be attributed to close supervision, individual and group sessions, tutoring, vocational counseling and positive reinforcement as the youths learn to integrate into their own communities as law-abiding members.[168]

The program is currently licensed to take up to only seventy Brooklyn resident juveniles for the six to eighteen month program.[169] Although this center at one time ran two programs, "Youth Achieve," for non-substance abusers, and "Oasis," for substance abusers only, Youth Achieve lost its federal funding in 1999 and is now defunct.[170] Unfortunately, despite great

---

163. *See* Letter from Maureen E. Casey, Deputy Commissioner, *New York City Police Department Office of Management, Analysis, and Planning*, to Evan Przebowski, Research Assistant, *Hofstra University Child Advocacy Clinic* (Jan. 1, 2001) (on file with author).

164. *See* Facsimile from the New York State Office of Children and Family Services to Theresa Hughes, Supervising Attorney, *Hofstra University School of Law Child Advocacy Clinic*, (Feb. 12, 2001) (on file with author).

165. *See* ANDREW GLOVER BROCHURE, *supra* note 63, at 4.

166. Wallace interview, *supra* note 156.

167. *Id.*

168. *See id.*

169. *See id.*

170. *See id.*

demand, the program is now only addressing the needs of seventy juveniles with substance abuse issues.[171]

2. The staff take the time to find out what these young people are facing in their lives and to pinpoint the specific issues that will help them to move in a positive direction. . . . In my view, CASES offers a model for the administration of justice that is fair, safe and effective.[172]

As the oldest standing and largest alternative-to-incarceration program in New York City, CASES serves as the ideal model with intensive supervision, guidance, support and a recidivist rate of only twenty percent -- more than one-third lower than the rate of reconviction for comparable youths sentenced to placement in a facility.[173] With funding from the state, city and private donations, this community-based program serves approximately 260 youths at one time. Participants report to the program's Manhattan center virtually daily, attending educational classes, G.E.D. preparation, computer classes, employment counseling, internships, HIV peer education, and recreational programs.[174] The program requires strict adherence to rules and regulations, while youths attend school on a full-time basis or maintain employment. If a participant fails to meet the program's conditions, CASES returns the youth to court for a violation of the court's dispositional order and the youth may then face placement.[175]

3. The program seeks to modify the individual's behavior and concentrate on questions of socially acceptable behavior.[176]

The Dome Project, a twenty-eight-year-old Manhattan-based alternative-to-incarceration program, serves youth between the ages of eleven through twenty-one within New York City, with a goal of helping each participant maintain a life free of crime.[177] Despite an impressively low fifteen percent recidivist rate for the year 2000, state and private funding allows a caseload of only fifty-five juveniles, amounting to only 125 kids per year.[178] Chris-

---

171. *See id.*

172. CENTER FOR ALTERNATIVE SENTENCING AND EMPLOYMENT SERVICES (CASES), CASES BROCHURE 10, New York, N.Y. (quoting Michael Gage, Former Administrative Judge, New York City Family Court) [hereinafter CASES BROCHURE].

173. *See id.* at 3.

174. *See* Telephone Interview with Laurie Held, Center for Alternative Sentencing and Employment Services, (Feb. 2, 2001) [hereinafter Held interview].

175. *See* CASES BROCHURE, *supra* note 172, at 3.

176. *See Programs at the Dome Project*, WHAT IS THE DOME PROJECT? (The Dome Project, Inc., New York, N.Y.), at 1.

177. *See id.*

178. *See* Telephone Interview with Christine Venuti, Juvenile Justice Coordina-

tine Venuti, the programs juvenile justice coordinator, attributes the high success rate to the intense services offered such as individual counseling, family sessions, group workshops, recreation, close monitoring of school attendance, job placement assistance, and one-to-one attention given to each individual's needs.[179]

4. Community-based rehabilitation brings kids back into the community, as opposed to jails where they are learning to perfect their criminal skills.[180]

The Youth Advocacy Project, a component of the Center for Community Alternatives, specifically addresses the problems that lead Brooklyn-based youth to their delinquent behavior by providing intensive services to youth under the age of sixteen years.[181] Youth attend the program three times per week after school to receive counseling and family support, group services, home and school monitoring and periodic staff home visits.[182] Funding limitations, however, allow only seven to thirteen youths to attend the program concurrently.[183]

5. I am not a believer in kids walking the streets. I am a believer that some need jail, but the majority need support. You can send a kid to a city or state facility at the cost of $60,000 or more a year, where he will be dumped into a placement with no services, or to our intensive community-based program for $1,800. Community-based programs are the solution.[184]

The privately funded Andrew Glover program, is fighting to make the community safer by teaching kids to become members of the work force rather than repeating crimes.[185] Seven out of ten youths enrolled in the program stop delinquent and criminal activity and go on to become productive members of society.[186] Program Director Angel Rodriguez asserts that he brings to the court "more than a rap sheet and complaint," allowing kids from the Lower East Side of Manhattan to have an advocate in court who has worked with the youth in his own home, on the streets, and in the play-

---

tor, *The DOME Project* (Feb. 8, 2001) [hereinafter Venuti interview].

179. *See id.*

180. Telephone Interview with Patrick Thomas, Director, *Youth Advocacy Project* (Feb. 9, 2001) [hereinafter Thomas interview].

181. *See id.*

182. *See id.*

183. *See id.*

184. Telephone Interview with Angel Rodriguez, Director and Founder, *Andrew Glover Youth Program* (Feb. 8, 2001) [hereinafter Rodriguez interview].

185. *See id.*

186. *See* ANDREW GLOVER BROCHURE, *supra* note 63, at 5.

ground.[187] The program maintains twenty-four hour supervision, 365 days a year, by assigning each youth a "Youthworker," keeping the program doors open until late at night, equipping his ten person staff with beepers, requiring curfew checks and regular monitoring. Youthworkers must be available twenty-four hours a day and the program must operate on the streets where the youths spend most of their time.[188] The Andrew Glover Program helps with family problems and counseling, provides job readiness classes, health and art classes, regular speakers (such as judges, police officers, and recruiters,) refers the youth for drug counseling, and offers employment and educational assistance.[189] Aside from community-based rehabilitation programs, Rodriquez claims, "the only other way to get out [of the inner-city ghetto] is those blue buses that take the kids upstate."[190]

These programs mirror the philosophy of the founders of the juvenile justice system. Despite the powerfully low recidivist rates and the lower cost, these types of community-based programs are the minority in juvenile rehabilitation in New York City benefiting only a few hundred youths, while the majority is placed in state detention facilities.[191]

PROPOSALS

Based on current trends of incarceration and placement, it is apparent that society does not favor a community-based solution to juvenile crime. Ironically, national public opinion surveys show that a significant majority favor rehabilitation and treatment, rather than punishment for juveniles.[192] A 1991 University of Michigan survey revealed that seventy-one percent of those surveyed believed that all but the most serious juvenile offenders should be supervised in community-based programs that focus on rehabilitation, while only twenty-nine percent favored residential correctional institutions.[193] The majority stated that they wanted their tax dollars invested in programs incorporating the restorative justice philosophy, which emphasizes the use of restitution to victims and service to the community.[194] They also wanted young offenders to have access to job training, and to be placed in community-based programs that focus on education and counseling

187. *See* Rodriguez interview, *supra* note 184.

188. *See id.*

189. *See id.*

190. *Id.*

191. *See* JUVENILE PANEL, *supra* note 32, at 176-82.

192. *See* Mabel Arteaga, Note, *Juvenile Justice with a Future . . . for Juveniles*, 2 CARDOZO WOMEN'S L.J. 215, 241 (1995).

193. *See* Robert E. Shepherd, Jr., *What Does the Public Really Want?*, 11 CRIM. JUST. 51, 51 (Spring 1996).

194. *See id.*

within the local community.[195] These survey results mirror recent studies of delinquency which have shown that successful treatment techniques have had larger positive effects when utilized within the community rather than institutional settings, demonstrating that the chances of rehabilitating troubled youths are higher for those who participated in community-based programs as opposed to those sent to institutional settings.[196] However, these opinions are not reflected by current juvenile justice policy and the government has repeatedly failed to allocate resources to effective rehabilitative programs.[197]

"Few states . . . claim to have the resources to tackle all aspects of youth violence and crime prevention at once."[198] Additionally, policymakers often want a guarantee that a rehabilitation program will produce a definitive outcome, because voters want immediate results.[199] The President's Crime Prevention Council (PCPC), based partially on the notion that youth do not have the decision-making capacity of adults,[200] recommends that states searching for the right balance of services for delinquents target their efforts to the types of programs that currently exist in the jurisdiction, and add programs that have been shown to have a positive impact on youth misconduct.[201]

A child is a unique being and requires a separate and special response to criminal behavior.[202] As "juvenile crime necessarily results from individual, family, and environmental factors, it is unwise to attempt to prevent juvenile crime by removing juveniles from the conditions to which they will eventually return."[203] The juvenile court's placement system should be reevaluated in light of the demonstrated success of the New York City alternative-to-detention programs.[204] There should be more community-based juvenile rehabilitation programs integrating the following services and resources to help the youth while he resides within his own neighborhood.

1. Family Involvement - Rehabilitation must integrate the family in order

---

195. *See id.*

196. *See* Krisberg, *supra* note 77, at 25, 58.

197. *See* Holland & Mlyniec, *supra* note 65, at 1794.

198. NAT'L CRIM. JUST. ASS'N, *supra* note 41, at 10.

199. *See id.* at 11.

200. *See* JUVENILE PANEL, *supra* note 32, at 1-4; J.J. Quadrel, et al., *Adolescent (In) Vulnerability*, 48 AM. PSYCH. 102-16 (1993).

201. *See* NAT'L CRIM. JUST. ASS'N, *supra* note 41, at 10-11.

202. *See* Campbell, *supra* note 49, at 356.

203. Roger J.R. Levesque, Alan J. Tomkins, *Revisioning Juvenile Justice: Implications of the New Child Protection Movement*, 48 WASH. U. J. URB. & CONTEMP. L. 87, 100-01 (1995).

204. Only a small portion of juveniles are arrested for serious crimes. *See* JUVENILE PANEL, *supra* note 32, at 102-03; *see also* O'Connor & Treat, *supra* note 42, at 1343-44.

to be effective. Since the family is the first social group that a child is exposed to, its structure is critical in the development of the youth's character.[205] It is "[t]he home environment [that] is especially critical [and] could be the influence that determines what type of environment is dominant in the youth's [life]."[206] Rehabilitationists firmly advocate that providing a nurturing environment for juvenile delinquents will develop new, positive self-images that will, in turn, result in lower juvenile crime rates.

2. Educational Assistance & Support - Since juvenile delinquency is associated with poor school performance, truancy and leaving school at a young age, the rehabilitation program must monitor school performance and behavior and provide needed support and referrals.[207]

3. Individual and Family Counseling - Family counseling plays an essential role in rehabilitating the youth, as families often contribute significantly to the juveniles' troubling behavior.[208] Also, reducing alcohol and drug abuse among parents may improve their ability to parent, which will reduce family-related risk factors for delinquency. Counseling, parenting skills and referrals may be helpful for both the youth and the parents.[209]

4. Recreation - There is a crime-averting effect to youth recreation facilities.[210] Recreation should be used as a tool that enables young people to gain necessary life skills; it provides an outlet, and if coupled with comprehensive services, enables youths to make good and appropriate choices.[211] Midnight basketball programs in themselves are simply not enough, but rather must be joined with strong youth development programs which give youth the self-empowerment skills to be successful.[212] Leaders who are involved with decision-making need to work on a goal-oriented agenda and engage youth in risk behavior modification, adventure and entrepenurial programs. [213]

5. Community Cooperation - "Mobilizing communities – including youth – and developing stronger ties between community residents, service providers, and law enforcement officials have proven to be critical components of crime prevention."[214] As the youth will be reporting to and coun-

---

205. *See* Hardung, *supra* note 72, at 142.

206. FREDERICK, *supra* note 50, at 21.

207. *See* JUVENILE PANEL, *supra* note 32, at 9.

208. *See* Levesque & Tomkins, *supra* note 203, at 101.

209. *See* JUVENILE PANEL, *supra* note 32, at 9.

210. *See id.* at 130.

211. Telephone interview with Lorraine Lubicich, Special Programs Consultant, Police Athletic League, New York City (Mar. 9, 2001) [hereinafter Lubicich interview].

212. *See id.*

213. *See id.*

214. NAT'L CRIM. JUST. ASS'N, *supra* note 41, at 10 (quoting THE PRESIDENT'S CRIME PREVENTING COUNCIL, PREVENTING CRIME AND PROMOTING RESPONSIBILITY: 50 PROGRAMS THAT HELP COMMUNITIES HELP THEIR YOUTH 4 (Sept. 1995)).

seled at a local site, it is essential not to have opposition from the community. Also, the participation of local volunteers and mentors, comprised of community members, parents and police, plays an active and positive role in providing the youth with appropriate direction.[215]

6. Peer Group Reinforcement - As employed by currently operating programs, peer group reinforcement provides an effective vehicle for emphasizing non-criminal behavior.[216] However, such peer groups must maintain adult supervision. Studies have shown that peer groups without adult supervision are not only ineffective, but may at times be harmful, since high-risk youth are particularly likely to reinforce one another's deviant behavior when they are grouped together for intervention.[217]

7. Street Supervision - Around the clock supervision provides access for the youths to a support system while monitoring their behavior. It has also proven to be an integral part of successful rehabilitation.[218]

8. Various services should be made available through the coordination of local programs, including, but not limited to: vocational training, therapy, mental health services, parental involvement, mentoring, community service, and drug and alcohol rehabilitation.[219]

CONCLUSION

Rehabilitation is a promising word in this era of alarmingly high recidivist rates.[220] Perhaps it is time to genuinely revisit the late nineteenth century objective of earnestly trying to transform youth into law-abiding citizens.[221] Denying a child a nurturing and familiar environment during the rehabilitation process will likely result in future criminal activity.[222] Rehabilitation requires treatment of juveniles as opposed to punishment.[223] Treatment acknowledges the socioeconomic factors that bear upon the youth's life; focuses on instilling positive values; provides an incentive to invest in programs; and is at least successful, if not more, in deterring future crime.[224] The formula for successfully responding to juvenile delinquency with treatment has been successfully utilized by the previously detailed community-based programs, as well as many others outside New York City.

---

215. *See* Hardung, *supra* note 72, at 142.
216. *See* Rodriguez interview, *supra* note 184.
217. *See* JUVENILE PANEL, *supra* note 32, at 147-48.
218. *See* Rodriguez interview, *supra* note 184.
219. *See* JUVENILE PANEL, *supra* note 32, at 150.
220. *See generally* FREDERICK, *supra* note 50, at 1.
221. *See* Marygold S. Melli, *Juvenile Justice Reform in Context*, 1996 WIS. L. REV. 375, 377, 393 (1996).
222. *See generally* FREDERICK, *supra* note 50, at 1.
223. *See generally* Day, *supra* note 138.
224. *See* O'Connor & Treat, *supra* note 42, at 1343-44.

Replication of these programs is needed through a genuine commitment from society coupled with a financial assurance by government and private donors.

The Panel on Juvenile Crime: Prevention, Treatment and Control was charged with numerous tasks including analyzing the full range of juvenile crime research.[225] The panel, at the outset, recommends that community-based rehabilitation programs be expanded, and that the federal government provide states with federal funding and incentives in order to reduce the use of secure detention by developing community-based alternatives that should replicate the existing, successful programs.[226] The panel further held that research showed that treating the youth within the community does not in fact compromise public safety, but rather may improve it through reduced recidivism.[227] As individual, social and community conditions, and peer interactions influence the behavior of delinquent youth, effective rehabilitation calls for services addressing all of these controls.[228]

The infinite list of compelling reasons for finally investing in community-based juvenile rehabilitative programs as true alternatives-to-detention starts with the simple right of each child to be given the opportunity to learn to continue walking as he passes through the intersection of Schaeffer and Wilson Streets.

225. *See* JUVENILE PANEL, *supra* note 32, at 14.
226. *See id.* at 7.
227. *See id.*
228. *See id.* at 8-9.

COMMENT

PGA v. ADA: A Skins Match, Winner Take All in the Battle for Equality for Disabled Individuals in Professional Golf

"This is where I want to be. I have a passion to play golf. When I wake up, I want to work on my game. And so I'm grateful that I still have that opportunity." [1] – Casey Martin.

"We have the highest regard for Casey Martin. But frankly, we don't think this case is about Casey Martin. We think this case is about whether or not a sports organization, like the PGA Tour, should be able to make its own rules." [2] – Tim Finchem (PGA Tour Commissioner).

"They have the right to make their rules. But they need to comply with what the United States government says. And they say for people with disabilities, you need to make accommodation. And they simply refuse to do that." [3] – Casey Martin.

"I think that saying that [the PGA is] losing the right [to make its own rules] is overstating it just a little, because reasonable accommodations have to be determined on an individualized basis, on a case by case basis. And so therefore deciding that Mr. Martin is entitled to this . . . really does not decide that question for any other golfer." [4] – Jane Korn (Law Professor, Univ. of Arizona).

"Sometimes tradition has to give way to the reality of equal opportunity for people, [t]raditions and rules that do not permit . . . equal opportunity

---

1. *ESPN: SportsCenter* (ESPN television broadcast, Jan. 14, 2001).
2. *Id.*
3. *Id.*
4. *Id.*

are rules and traditions that have to go by the wayside."[5] – Senator Tom Harkin (D-Iowa).

Q: "Wouldn't this open a Pandora's box, in all sports, to change the landscape of competition?"

A:"I don't think so."

Q:"Why?"

A:"Because there aren't that many people with disabilities playing professional sports; I mean think about it . . . golf is one sport where you can make accommodations for people."

Q:"What if you lose? What happens?"

A:"If I lost, obviously my first appeal would be to the tour . . . one last chance to see if they would be willing to let me ride, even after this situation; if they refuse it wouldn't look that good on the short term for my golf because physically . . . my leg just isn't in a position to start walking, but you never know . . . for all intents and purposes this would be the end of my career."[6]

I. INTRODUCTION

The fate of Casey Martin, Ford Olinger and disabled athletes everywhere, hangs in the balance as the Supreme Court deliberates the issues presented before it on January 17, 2001, during oral arguments in the case of *Martin v. PGA Tour, Inc.*[7] The Supreme Court has the task of deciding whether the Americans with Disabilities Act (ADA) of 1990 is applicable to professional sports organizations. If the Court answers that question in the affirmative, it would open the door for disabled athletes everywhere to use the ADA to gain access to a wide variety of sports competitions. If the Court concludes that the ADA does not apply to the Professional Golf Association (PGA), it would shut the door on the expanding progression of the ADA and limit its application significantly.

In Part II of this article, a brief history of disability legislation is presented, and the specific vehicle used by Casey Martin to gain access to the courts is explained.[8] Part III outlines the decisions of the reviewing courts

---

5. *ESPN: SportsCenter* (ESPN television broadcast, Jan. 17, 2001). Senator Harkin was speaking from the United States Supreme Court steps directly after oral arguments were heard by the Supreme Court. He was standing side-by-side with Casey Martin and Senator Bob Dole. Both Senators Harkin and Dole were sponsors of the ADA.

6. *Id.* (interview between Mark Gumbell, an ESPN journalist, and Casey Martin).

7. 204 F.3d 994 *aff'd* 121 S. Ct. 1879 (2001).

8. *See infra* notes 46-83 and accompanying text.

in deciding *Martin v. PGA Tour, Inc.* and *Olinger v. U.S. Golf Ass'n.*[9] Part IV of this article attempts to answer the biggest question of this debate: whether the ADA, through its public accommodation language, is applicable to the PGA.[10] This part also looks at the affirmative defenses the PGA and other sports organizations would have if it is determined that the ADA applies.[11] The final part of this paper attempts to rationalize the outcome of this case from the viewpoint of the U.S. Supreme Court.

A. Casey Martin

Casey Martin is currently a twenty-eight-year-old Oregon resident who has been playing golf since he was six years old.[12] He continued his golfing aspirations at Stanford University, where he played a top-notch game.[13] At Stanford, he played with future greats such as Tiger Woods and Notah Begay,[14] before joining the Nike tour and turning pro in 1995.[15]

Martin suffers from Klippel-Trenaunay-Weber Syndrome (KTW), which is a rare congenital, degenerative circulatory disorder, [16] developed as a child.[17] Also referred to as Parkes-Weber syndrome, KTW was reported in less than 150 cases in 1984.[18] It affects every person differently, making a cure nearly impossible to come by and worsening the effects because there is no standard treatment.[19] Martin has circulatory problems in his right leg

---

9. *See infra* notes 84-158 and accompanying text.

10. *See infra* notes 159-234 and accompanying text.

11. *See infra* notes 247-295 and accompanying text.

12. *See* Mike Cullity, *Casey Martin v. PGA Tour Inc.*, GOLFWEEK, *at* http://www.golfonline.com/news/golfweek/1998/january/caseybio0124.html (visited Jan. 24, 1998).

13. *See id.* ("[1990-91], posted five top-15s, including T8 at San Francisco invitational and T4 at NCAA West Regional . . . [1991-92], posted six top-20s, his best being a T6 at Pac-10 Championships, where he led Stanford to its first Pac-10 title since 1977 . . . [1993-94], had six top-5s in 15 starts, including a victory at Northwestern Classic and second places at Ping/GOLFWEEK Preview and Nike Southwestern").

14. *See* The Hartford, *Casey Martin: The Bring it on Spirit, available at* http://www.thehartford.com/corporate/special/bio.html (visited Feb. 22, 2001).

15. *See id.* The Hartford is Martin's endorsing sponsor since 1998, the sponsorship was renewed for three years in 2000. *See id.*

16. *See* Martin v. PGA Tour, Inc., 204 F.3d 994, 996 (9th Cir. 2000).

17. *See* The Hartford, http://www.thehartford.com/corporate/special/bio.html, *supra* note 14.

18. *See* Mike Cullity, *Casey Martin v. PGA Tour, Inc.*, GOLFWEEK *at* http://www.golfonline.com/news/golfweek/1998/january/syndrome0124.html.

19. *See id.*

as a result of KTW, which causes the leg to be significantly weakened and swollen.[20] KTW has put Martin at risk for tibial fractures with almost any daily motion which includes swinging a golf club, and can result in excessive bleeding, putting his life in danger.[21]

As of 1998, Martin's condition significantly worsened to the point where he was unable to walk the length of a golf course without having great difficulty because of the associated pain.[22] Recently, in the Kemper 2000 Insurance Open, Martin's cart was unavailable at the start of the opening round.[23] This caused Martin to have to walk over 200 yards, in considerable pain, to reach his ball after he shot it into the rough.[24] That was the longest he had walked in two years.[25] Martin originally petitioned the PGA for permission to use a cart to play in PGA sponsored tournaments.[26] At first, the PGA denied his request, but on February 11, 1998, Federal Magistrate Thomas Coffin granted Martin's petition for a preliminary injunction, forcing the PGA to allow him use of a cart, using the ADA of 1990 as his basis for suit against the PGA.[27] The Ninth Circuit Court of Appeals affirmed the Magistrate's ruling in May of 1999.[28] The U.S. Supreme Court is now going to review Martin's victory[29] in lieu of an adverse ruling decided one day after the Martin case by the Seventh Circuit.[30] In *Olinger v. U.S. Golf Ass'n*,[31] the Seventh Circuit took an opposite approach to the *Martin* court, and denied Ford Olinger the right to use a cart in United States Golf Association (USGA) sponsored events via the ADA.[32] This circuit split, coupled with the highly publicized *Martin* ruling and compounded by the golfing habits of Justices on the bench, has drawn

---

20. *See id.*

21. *See id.*

22. *See* The Hartford, http://www.thehartford.com/corporate/special/bio.html, *supra* note 14.

23. *See* Joseph White, *Long Walk for Casey Before Cart Arrives*, *at* http://golfonline.com/tours/2000/pga/Kemper/0604.html (last visited Feb. 23, 2001).

24. *See id.*

25. *See id.*

26. *See* The Hartford, http://www.thehartford.com/corporate/special/bio.html, *supra* note 14.

27. *See generally* Martin v. PGA Tour, Inc., 984 F. Supp. 1320 (D. Or. 1998).

28. *See generally* Martin v. PGA Tour, Inc., 204 F.3d 994 (9th Cir. 2000).

29. *See* PGA Tour, Inc. v. Martin, 121 S. Ct. 30 (2000), 530 U.S. 1306 (2000) (Mem.) (granting PGAs' petition for writ of certiorari on Sept. 26, 2000).

30. *See generally* Olinger v. U.S. Golf Ass'n, 205 F.3d 1001 (7th Cir. 2000) (*vacated* 121 S. Ct. 2212 (2001)).

31. *See id.*

32. *See id.* at 1007.

some speculation as to the real reason this case is being reviewed.[33] The issues presented upon appeal are centered around disability legislation, namely the ADA of 1990, which was the basis for Casey Martin[34] and Ford Olinger[35] bringing their suits.

B. Ford Olinger

Ford Olinger is a thirty-three-year-old professional golfer from Warsaw, Indiana.[36] He suffers from a congenital hip disorder known as bilateral avascular necrosis.[37] Like Casey Martin, Olinger has great difficulty walking a golf course and can barely finish an eighteen hole round.[38] On May 20, 1999, Ford Olinger was denied use of a cart in USGA tournaments by a district court judge in Indiana.[39] Olinger lost his appeal in a Seventh Circuit decision one day after the *Martin* decision was handed down.[40] Ford Olinger is somewhat disturbed by the two differing results regarding similar issues; "I want everyone to know that I wish Casey the best of luck . . . . But if the USGA wants a fair field for everyone, there's another guy with a cart and its not me. I don't know how you can justify that."[41] When asked how he felt about Casey Martin being allowed to ride in a cart he said: "I'm in the same shoes . . . . What's that difference?"[42] Olinger now awaits the decision of the Supreme Court in the *Martin* case to see whether an appeal is viable.[43] Ford Olinger continues to persist in his efforts to play in professional golf tournaments and the United States

---

33. *See generally* Jeffrey A. Rosenthal, *Should High Court Review Case of the Disabled Golfer?*, N.Y.L.J., Aug. 7, 2000, at 9.

34. *See generally* Martin v. PGA Tour, Inc., 204 F.3d 994 (9th Cir. 2000).

35. *See generally Olinger*, 205 F.3d at 1001.

36. *See* John Fineran, *Long Walk Ahead for Olinger*, SOUTH BEND TRIBUNE, May 22, 1999, at 17.

37. *See id.*

38. *See* J.R. Ross, *Disabled Pro Fails to Qualify for U.S. Open Sectional, at* http://www.golf.com/news/19990524/gol/19990524.olinger-noqualf.html (last visited Feb. 24, 2001).

39. *See* Olinger v. U.S. Golf Ass'n, 55 F. Supp. 2d 926, 938 (N.D. Ind. 1999).

40. *See* Olinger v. U.S. Golf Ass'n, 205 F.3d 1001, 1007 (7th Cir. 2000) (This decision was handed down on March 7, 2000, while the *Martin* decision was handed down on March 6, 2000).

41. Fineran, *supra* note 36.

42. Tom Coyne, *Walker Doesn't Help Olinger in U.S. Open Local Qualifier*, GOLFWEEK *at* http://www.golfweek.com/articles/2000/homepages/12095.asp (last visited Feb. 24, 2001).

43. *See* Fineran, *supra* note 36. The idea of an appeal was quashed thanks to the vacated decision, 121 S. Ct. 2212.

Golf Association has approved use of a walker in his efforts.[44] However, Olinger is not optimistic that he will be able to continue playing golf in this way; "If I can't ride I can't play, I'm not coming out here and doing this. I'm hurting too much right now. My chest is hurting because of all the drugs I'm taking. There's just no way."[45]

II. HISTORY OF DISABILITY LEGISLATION

The move towards gaining equality for disabled Americans was started as early as 1920, with the implementation of the Smith-Fess Act.[46] This act was designed to help vocationally disabled veterans of World War I make up for lost experience while they were at war.[47] By making this Act permanent by incorporating it into the Social Security Act of 1935, Congress began a push for the equality for disabled individuals in every day life.[48] Ultimately this effort starting in 1935, culminated into what is now known as the Rehabilitation Act of 1973.[49] The scope of the Rehabilitation Act of 1973 included a prohibition of discriminatory acts against people with disabilities and broadened disability legislation to include all federal contractors, any entity that receives federal grant money, and people who partake in federal programs.[50] This culmination of discriminatory prohibitions broadening a disabled individual's chance to achieve gainful employment was a landmark step in the fight for disabled individuals' freedom from discrimination.

The Supreme Court corralled the Rehabilitation Act in 1986 in *United States Department of Transportation v. Paralyzed Veterans,*[51] where it held

---

44. *See* Coyne, *supra* note 42.

45. *Id.* When trying to qualify for the U.S. Open without use of his cart, but with the approved walker, Olinger shot 19 over and failed to qualify. *See id.* "The knees started to lock in and then the hips started aching, then I had to watch every step, because it jolts me pretty good." This was after the seventh hole at six over par. *Id.*

46. *See* HENRY H. PERRITT, JR, ADA HANDBOOK 106 (2d ed. 1991).

47. *See id.*

48. *See id.* In 1943, scope of the Smith-Fess Act was expanded to the mentally and physically ill. In 1965, federal funding was allotted to state rehabilitation programs. In 1967, federal money being put into the program was increased and disabled migrant workers and their families were incorporated into the rehabilitation program. The 1968 amendments gave each state $1million dollars minimum funding and the rehabilitation program again expanded to include family services and new employment opportunities. *See id.* at 106-07.

49. *See* PERRITT, *supra* note 46, at 107.

50. *See id.* at 105.

51. 477 U.S. 597 (1986).

that airline companies were not subject to the prohibitions of the Act simply because they used an airport which was federally funded.[52] The Supreme Court refused to broaden the scope of the Rehabilitation Act of 1973 to include entities that are intertwined with federal monies, as opposed to entities actually receiving federal funds.[53] The Court reasoned that the airport was the direct recipient of federal monies, not the airline companies.[54] This was the first hint that the Rehabilitation Act was not going to be sufficient to cover disabled individuals in a wide spectrum of areas. The Congressional move towards expanding the Rehabilitation Act of 1973 began in the late 1980's.[55] Senator Tom Harkin (D-Iowa) led the revolution in expanding the Rehabilitation Act of 1973 so that it included private entities as well as those entities which were covered under the original act.[56] The expansion effort led by the distinguished Senator from Iowa ultimately led to the enactment of one of the most powerful Acts to impact discrimination on disabled individuals to date, The Americans with Disabilities Act of 1990.[57]

The ADA of 1990 is a comprehensive approach complimenting the Rehabilitation Act of 1973 and expanding the scope of disabled citizens' rights.[58] The ADA does not replace, limit, or change the Rehabilitation Act of 1973, as indicated by legislative history[59] and in the text of the ADA itself, "[N]othing in this Act shall be construed to apply a lesser standard than the coverage required or the standards applied under [the Rehabilitation Act]."[60] Safeguards were also built into the ADA so that the application of both Acts would not conflict in any way.[61]

---

52. *See id.* at 610.

53. *See id.* at 598, 613.

54. *See id.* at 609.

55. *See* 136 CONG. REC. S9684, S9691 (1990) (outlining provisions authored by Sen. Tom Harkin, chairman of the subcommittee on disability beginning in 1986, including: The Developmental Disabilities Assistance Program and Bill of Rights Act Amendments of 1987; establishing a National Institute on Deafness and Communication Disorders; the Technology Related Assistance for Individuals with Disabilities Act of 1988; the Protection and Advocacy for Mentally Ill Individuals Amendment Act of 1988). *Id.*

56. *See* 136 CONG. REC. S9684, S9691-92.

57. 42 U.S.C. § 12101 et seq. (1995).

58. *See* JANE K. ALPER ET AL., THE ADA 79 (Massachusetts Continuing Legal Education ed. 1992).

59. *See* Menkowitz v. Pottstown Mem'l Med. Ctr., 154 F.3d 113, 120 (3d Cir. 1998).

60. 42 U.S.C.A. § 12201 *reprinted in*, HENRY H. PERRITT, ADA HANDBOOK 117 (2d ed. 1991).

61. *See* 42 U.S.C. § 12117(b) (1995). This section, entitled *Coordination*, provides that the agencies responsible for enforcing the Rehabilitation Act and

The House of Representatives and the Senate passed the Act on July 12, and July 13, 1990 respectively, with President George H. Bush signing the Act into law on July 26, 1990, at a White House ceremony in honor of the Act's passage.[62]

The ADA is broken up into four titles; Title I deals with discrimination of disabled individuals in employment by private entities, which by far has seen the most use;[63] Title II deals with state and local government services;[64] Title III encompasses discrimination of disabled individuals with regards to goods and services provided by private entities;[65] and Title IV deals with discrimination in the realm of the telecommunications industry.[66] The application of Title III to professional sports is the main focus of this article.

A.Title III

The ADAs' Title III provision is the heart of a highly publicized debate centered around the Martin and Olinger legal battle. The legislature enacted Title III to " . . . provide disabled individuals the opportunity to participate in or benefit from all that places of public accommodation have to offer."[67] When applied correctly the statute allows for a disabled individual to access and enjoy certain goods, services, facilities, privileges, advantages, or accommodations provided by an entity.[68] The operative language of the statute, however, has some virtual hoops that must be jumped through before permanent relief can be granted.

Inherent in all statutes are elemental requirements that are necessary in order for it to apply to a given situation and afford relief. The ADA Title

---

ADA shall create policies to the degree necessary to ensure cohesive implementation of the two respective acts. *Id.*

62. *See* PERRITT, *supra* note 46, at 22.

63. *See generally* Michael Faillace, *Title I of the ADA: Statutory Requirements, Legislative History, Regulations, Technical Assistance Manual, Relevant Case Law under the ADA, and 1973 Rehabilitation Act, and Practical Recommendations*, 527 PRACTICING LAW INST. LITIG. AND ADMIN. PRACTICE, 481 (1995).

64. *See* Tanya R. Sharpe, Comment: *Casey's Case: Taking a Slice Out of the PGA Tour's No-Cart Policy*, 26 FLA. ST. U. L. REV. 783, 787 (1999).

65. *See* 42 U.S.C. § 12182(a) (1995).

66. *See* Sharpe, *supra* note 64, at 787.

67. Grant P. Fondo, *Access Reigns Supreme: Title III of the ADA and Historic Preservation*, 9 BYU J. PUB. L. 99, 106-07 (1994).

68. *See* 42 U.S.C. § 12182(a) (1995). The Section reads: "No individual shall be discriminated against on the basis of *disability* in the full and equal enjoyment of the goods, services, facilities, privileges, advantages, or accommodations of any place of *public accommodation* by any person who owns, leases (or leases to), or operates a place of public accommodation." *Id.* (emphasis added).

III has three so-called elements.[69] The first is that the individual must be *disabled* within the meaning of the statute itself.[70] Second, the goods, services, facilities, privileges, advantages, or accommodations that the disabled individual seeks the right of entry to must be offered by, what the code defines as, a *public accommodation* under Title III.[71] Finally, if the first two requirements are shown to be true as provided by the Code, the accommodation requested by the disabled individual to gain access to the goods, services, facilities, privileges, advantages, or accommodations *must be reasonable*. Provided that the above statutory requirements are met, the public entity must provide the requested accomodation. However, there are two affirmative defenses available to the public entity to counter the requested accomodation. First, the public entity can show that the modification requested by the disabled individual *fundamentally alters* the nature of those goods, services, facilities, privileges, advantages, or accommodations that are offered by the entity.[72] Secondly, if the entity can show that providing the requested modification to allow a disabled individual access to the goods, services, facilities, privileges, advantages, or accommodations would cause that entity to sustain an *undue administrative* or *financial burden*, then the requested modification will be denied.[73] These requirements have a far greater depth than previously outlined and require vigorous analysis to determine the requisite applicability.

The first element, that a person seeking relief under the ADA be "disabled" within the meaning of the statute, requires the least discussion by far. The applicable provision of the ADA is as follows: "The term 'disability' means, with respect to an individual . . . a physical or mental impairment that substantially limits one or more of the major life activities of such an individual . . . a record of such an impairment . . . or . . . being regarded as having such an impairment."[74] This provision, albeit self explanatory, has had some limited debate associated with it, but for the pur-

---

69. *See generally* 42 U.S.C. § 12182 (1995).
70. *See* 42 U.S.C. § 12102(2) (1995).
71. *See* 42 U.S.C. § 12182(a) (1995).
72. *See* 42 U.S.C. § 12182(b)(2)(a)(ii) (1995). Discrimination is:

> a failure to make reasonable modifications in policies, practices, or procedures, when such modifications are necessary to afford such goods, services, facilities, privileges, advantages, or accommodations to individuals with disabilities, [*U*]*nless* the entity can demonstrate that making such modifications would *fundamentally alter* the nature of such goods, services, facilities, privileges, advantages, or accommodations.

Id.

73. *See* 28 C.F.R. § 35.164 (1995).
74. 42 U.S.C. § 12102(2) (1995).

poses of this article it does not provide a source of controversy.[75]

The second element to a claim under the ADA requires that the providing entity to which the disabled person seeks to gain access be a *public accommodation* as defined within Title III.[76] A public accommodation is defined by a list provided by the drafters of the Act.[77] This list is not exhaustive or to be seen as limiting the scope of applicability of the ADA with regard to public accommodations.[78] This element of the ADA has seen numerous interpretation problems and is a fundamental problem underlying the Casey Martin and Ford Olinger debate.[79]

Thirdly, the ADA allows an entity to show, by way of an affirmative defense, that a request by a disabled individual will fundamentally alter the nature of the goods, services, facilities, privileges, advantages, or accommodations that the entity is offering. The "fundamentally alter" concept originally arose under the Rehabilitation Act[80] by common law creation.[81] If it is found that the requested modification will fundamentally alter the nature of what the entity provides than the language of the ADA's Title III supplies an escape valve; the entity will not have to provide the modification.[82] This is an important provision because the wording of the statute does not require that the disabled individual show that their request will not fundamentally alter the nature of the goods, services, facilities, privileges, advantages, or accommodations. Instead, the burden of proof is upon the challenged entity to show to the court that the modification is improper because it will fundamentally alter the nature of the "thing" that it pro-

---

75. *See* Martin v. PGA Tour, Inc., 204 F.3d 994, 996 (9th Cir. 2000). "There is no dispute that Martin is profoundly disabled." *Id*; *see also* Olinger v. U.S. Golf Ass'n., 205 F.3d 1001 (7th Cir. 2000). "Everyone agrees he [Olinger] is a 'disabled' person within the meaning of the ADA." *Id.* at 1001.

76. 42 U.S.C. § 12182(a) (1995).

77. 42 U.S.C. § 12181(7). "The following private entities are considered public accommodations for the purposes of this subchapter, if the operations of such entities affect commerce . . . a motion picture house, theatre, concert hall, stadium, or other place of exhibition or entertainment . . . a gymnasium, health spa, bowling alley, golf course, or other place of exercise or recreation . . . ." *Id.*

78. There are a total of fifty-seven specific public accommodations listed in this section and additional phrases like "or other place of recreation" and "other establishment serving food or drink," which indicate that this list is not intended to be final or exhaustive. *See* 42 U.S.C. § 12181(7) (1995).

79. *See infra* Part IV(b) for the discussion of the problems of defining public accommodation.

80. *See supra* notes 48-56 and accompanying text.

81. *See* Southeastern Community College v. Davis, 442 U.S. 397, 414 (1979) (codified in 42 U.S.C. § 12182(b)(2)(a)(ii) (1995)); *see also infra* notes 267-93 and accompanying text.

82. *See* 42 U.S.C. § 12182(b)(2)(a)(ii) (1995).

vides.[83]

III. THE REVIEWING COURTS

A. Casey Martin v. PGA Tour Inc.

The Ninth Circuit Review

The PGA appealed the district court's ruling that it was a public accommodation under the ADA and that modifying the walking rule would not fundamentally alter the nature of the game.[84] The district court granted Casey Martin use of a cart in PGA sponsored events through a permanent injunction.[85] The Ninth Circuit sided with Martin and upheld the lower court's decision.[86]

To decide the PGA's appeal from the district court's decision, the Ninth Circuit focused on two issues.[87] The first issue presented and decided by the court of appeals, was whether Title III of the ADA applies to the PGA, namely whether the PGA qualifies as a public accommodation within the meaning of the Act.[88]

The Ninth Circuit chose to apply the public accommodation requirement of Title III broadly to include the PGA within its umbrella.[89] "There is nothing ambiguous about this provision; golf courses are public accommodations."[90]

The PGA contends that with regard to the spectator portions of the golf course, it does fall within the realm of public accommodation defined by Title III, but that the course itself where the players compete does not fall within this classification as defined by the ADA.[91] This is the basis of the

---

83. *See id.* (stating that discrimination includes "a failure to make reasonable modifications . . . unless the entity can demonstrate that making such modifications would fundamentally alter the nature of such goods, services ").

84. *See generally* Martin v. PGA Tour, Inc., 984 F. Supp. 1320 (D. Or. 1998) (granting partial summary judgment to Martin on the issue of the applicability of the ADA to the PGA), *aff'd*, 121 S. Ct. 1879 (2001).

85. *See id.*

86. *See* Martin v. PGA Tour, Inc., 204 F.3d 994, 1002 (9th Cir. 2000), *aff'd*, 121 S. Ct. 1879 (2001).

87. *See id.*

88. *See id.* at 997.

89. *See id.* at 997, 999.

90. *Id.* at 997. The Court was referring to 42 U.S.C. § 12181(7) which provides that "[t]he following private entities are considered public accommodations for purposes of this title . . . ." and 42 U.S.C. § 12181(7)(L), which adds "a gymnasium, health spa, bowling alley, golf course, or other place of exercise or recreation." *Id.*

91. *See Martin*, 204 F.3d at 997.

mixed use facility argument by the PGA.[92] The PGA relies on two examples that are given within the ADA itself.[93] The first example is that of a hotel which is separated into two parts; one section being hotel and the other residential housing.[94] The former half of the hotel being governed by 42 U.S.C. § 12181(7)(A), which pertains to "an inn, hotel, motel, or other place of lodging . . . "[95] The residential part of the hotel was to be governed by The Fair Housing Act.[96] The Ninth Circuit dealt with this argument by saying that "the residential wing has never functioned as a hotel. A golf course during a tournament, however, is serving as a golf course."[97] The second example used by the PGA in their argument is that of a factory that provides tours to the public; the tour path itself would be a public accommodation as defined by the ADA,[98] while the sections simply being viewed by the public from the designated routes are not.[99] The Ninth Circuit concluded that this example would only be applicable if Casey Martin was "a spectator seeking to use his golf cart within the competitors' area of a tournament."[100]

The Ninth Circuit turns to the PGA's assumption, with its arguments, that the tournaments it conducts are not public.[101] The PGA argues to the Ninth Circuit that its tournaments are held as competition between the best golfers in the nation and that this in itself makes the tournaments outside the scope of Title III.[102] The court reasons that this is not an acceptable argument. The court references the ADA and its application to secondary, undergraduate, and post-graduate schools.[103] The court uses this reference by the ADA to conclude that "[t]he competition to enter the most elite private universities is intense, and a relatively select few are admitted"[104] and that does not preclude the application of ADA Title III.[105] Justice Canby

---

92. *See infra* notes 197-219 and accompanying text.
93. *See Martin*, 204 F.3d at 998.
94. *See id.* (citing 28 C.F.R. Ch. I pt. 36, app. B at 623 (1999)).
95. 42 U.S.C. § 12181(7)(A) (1995).
96. *See Martin*, 204 F.3d at 998.
97. *Id.* at 998.
98. *Id.* (citing 28 C.F.R. Ch. I pt. 36, app. B at 624 (1999)).
99. *See id.*
100. *Id.*
101. *See Martin*, 204 F.3d at 998.
102. *See id.*
103. *See id.*
104. *Id.*
105. *See Martin*, 204 F.3d at 998. In footnote 7, the court says that the ADA does not limit its application to members of the public, but that the ADA states: "No individual shall be discriminated against." *Id.* at n.7 (quoting 42 U.S.C. §12182(A)). In addition, the court does relent in saying that the individual has to be entitled to access the area; therefore, an individual not admitted to the school

opined that there is a similar process in the PGA. Anyone who pays a $3000 application fee and has two letters of recommendation can attempt to qualify for the PGA.[106] Canby further contends the golf courses that these qualifying rounds are played on are public accommodations.[107] The applicant plays in a series of tournaments to determine whether the applicant moves to the next stage. The final stage is qualifying for the PGA. The court does not allow the theory that when "the performance of athletes becomes so excellent"[108] that the place in which they play changes in legal character.[109] The court decided that "[d]espite the surface plausibility of this argument, it too narrowly construes the nature of a public accommodation."[110]

Justice Canby of the Ninth Circuit looks next at the idea of a reasonable modification, which is required by a public accommodation when it has been determined that such a modification is necessary for a disabled individual to be afforded "goods, services, facilities, privileges, advantages, or accommodations,"[111] unless the public accommodation can show that the modification would fundamentally alter the nature of the good or service.[112] The court goes on to say that the reasonableness aspect is a non-issue; it is reasonable to allow golfers to use carts, especially since they are allowed in other competitions that the PGA sponsors, and it would solve Martin's problem.[113]

The court theorizes that the real problem in this case is whether the use of a cart by Casey Martin "will 'fundamentally alter'" the nature of the

---

would not be covered. *See id.*

106. *See id.* at 999.

107. *See id.* At one point the PGA admitted that "'the competition areas of some amateur sporting events may well constitute places of public accommodation under Title III of the ADA when virtually any member of the public can participate.'" *Id.*

108. *Id.*

109. *See id.*

110. *Id.* at 997.

111. *Id.* at 999. The applicable provision of the ADA to which the court is referring to is:

> a failure to make reasonable modifications in policies, practices, or procedures, when such modifications are necessary to afford such goods, services, facilities, privileges, advantages, or accommodations to individuals with disabilities, *unless* the entity can demonstrate that making such modifications would *fundamentally alter* the *nature* of such goods, services, facilities, privileges, advantages, or accommodations.

42 U.S.C. § 12182(b)(2)(A)(ii) (1995) (emphasis added).

112. *See Martin*, 204 F.3d at 999.

113. *See id.*

game as it is played.[114] The court summons Rule 1-1 of The Rules of Golf to summarize what the nature of the game is; "The game of golf consists in playing a ball from the teeing ground into the hole by a stroke or successive strokes in accordance with the Rules."[115] The Ninth Circuit correctly points out that the PGA is not holding a game of golf, it is holding a specific competition with specific rules;[116] these rules include walking the course.[117] The PGA Tour Rule Committee can waive the walking rule and on occasion it has chosen to do so.[118] The Court is now charged with determining whether a decision to issue a permanent injunction for a permanent waiver of the walking rule would fundamentally alter the nature of the game.

The PGA contends that the question of whether or not a golf cart for Martin would fundamentally alter the game was answered when the district court concluded that the walking rule was intended to inject fatigue into the game.[119] The district court took a wholly subjective approach to answering the "fundamentally alter" question. They found that the question must be asked of whether letting Martin ride in a cart would fundamentally alter the game, and therefore it is an extremely fact-based question.[120] The district court found that Martin had to still walk twenty-five percent of the time, that Martin still had significant pain when walking and moving from the cart to his ball and back and that Martin "easily endures greater fatigue even with a cart than his able-bodied competitors do by walking."[121] The Ninth Circuit concludes that these findings were not erroneous and that there was sufficient evidence to support these findings.[122] The key component to the competition in a PGA event, which is shot making, will be undisturbed by Martin's use of a cart.[123] Justice Canby stated further that the use of a cart will afford Martin access to a type of competitive play that he

---

114. *See id.*

115. *Id.* (quoting from Rule 1-1 of The Rules of Golf established by The USGA and the Royal and Ancient Golf Club of St. Andrews, Scotland).

116. *See Martin*, 204 F.3d at 999.

117. *See id.*

118. *See id.* at 1000. The PGA Tour Rules Committee has waived the walking rule to save time when players have lost a ball and need to get back to the tee box. The Committee has also allowed carts to be used between holes when the distance to be traversed is long. *Id.*

119. *See id.* The PGA contends that this rule is "substantive" and therefore it is not subject to exceptions to accommodate disability. *Id.*

120. *See Martin*, 204 F.3d at 1001.

121. *Id.* at 1000. (quoting Martin v. PGA Tour, Inc., 994 F. Supp. 1242, 1252 (D. Or. 1998)).

122. *See id.*

123. *See Martin*, 204 F.3d at 1000.

would otherwise not be able to enjoy and that this is the major function that Congress wanted the ADA to serve.[124] This ends the Circuit Court's discussion on whether the use of a cart fundamentally alters the nature of the competition in the negative. The court then briefly discusses the PGA's last contention that the subjective analysis they would be required to undertake, to determine who would need a cart and who would not, would be far too burdensome and, therefore, they should not have to abide by it.[125]

The PGA asserts that, even if they are held to be within the scope of the ADA, and even if a cart for Martin does not fundamentally alter the nature of the competitive game of golf, the administrative burden imposed on the PGA is so great that the exception should not apply.[126] In a short determination of this issue the court wrestles with some twisted problems.

The PGA lays the foundation for its argument by citing *Sandison v. Michigan High School Athletic Ass'n*,[127] where the Sixth Circuit found that it would be too great a burden to require coaches and hired physicians to evaluate what factors would make a student's age an unfair advantage.[128] In *Sandison*, a high school student was seeking waiver of an eligibility requirement, which limited a student's eligibility by age, to play interscholastic sports, arguing that a learning disability made him repeat a grade.[129] This grade repetition caused him to be too old to play sports in his senior year of high school.[130] The Ninth Circuit in *Martin* rejects this argument and sides with the dissenting opinion of Judge Richard Arnold in *Pottgen v. Missouri State High School. Ass'n*,[131] who rejected an objective analysis of the burden factor. Judge Canby, in writing the opinion for this case, agrees with Judge Arnold that individual analysis must be undertaken in order to adjudge whether there is an undue burden placed on the organization in accomodations for this specific individual.[132] The court in *Martin*

---

124. *See id.* Justice Canby looks at legislative intent described in a case decided through his own circuit. *Id.*

125. *See id.* at 1001.

126. *See id.* The PGA relies on *Sandison v. Michigan High Sch. Athletic Ass'n*, 64 F.3d 1026 (6th Cir. 1995), to make their argument. This case was about an upper age limit for high school athletics. The court in this case said that it would be an undue burden for the school and physicians to determine whether a student's age, coupled with various factors, gave him an unfair advantage. *Id. See also* Pottgen v. Missouri State High Sch. Activities Ass'n, 40 F.3d 926, 931 (8th Cir. 1994).

127. 64 F.3d 1026 (6th Cir. 1995).

128. *See id.* at 1035.

129. *See id.* at 1028.

130. *See id.*

131. 40 F.3d 926 (8th Cir. 1994).

132. *See Martin*, 204 F.3d at 1002. The Court quotes language from the dissenting opinion by Judge Arnold that suggests that it does not matter if the school

says that the PGA in this case refused to analyze the disability of the golfer, and that with regard to this specific case the district court had no problems determining that the requested modification of the rule would not fundamentally alter the nature of the game of golf.[133] The Ninth Circuit does not address whether it would be an administrative or financial burden to make this analysis for 100 golfers or even ten golfers wishing a modification of the rule, but yet limits its analysis to each individual supplicant.[134]

The court concludes its opinion by holding that a golf course is a public accommodation under Title III of the ADA, that the district court did not err in determining that Casey Martin's request for a golf cart was a reasonable modification and that it did not fundamentally alter the nature of the tournaments.[135] Based on this decision by the Ninth Circuit Court of Appeals coupled with the opinion of the Seventh Circuit in *Olinger*, the PGA appealed this adverse ruling to the Supreme Court of the United States.[136]

B. Ford Olinger v. United States Golf Association

Seventh Circuit Review

On March 7, 2000, just one day after the *Martin* decision was handed down, the Seventh Circuit handed down an opinion that would create a controversy that the U.S. Supreme Court would ultimately vacate.[137]

The initial question of Ford Olinger's disability status was discarded in much the same way that the Court in *Martin* did.[138] Judge Evans spends the first three and one-half pages of his Seventh Circuit opinion discussing

---

had a thousand requests to review the eligibility requirements, if it was not a burden to review one, then it is not a burden to review them all. *See id.* "Nor will the record support the argument that a waiver of the rule in Mr. Washington's case would place an undue administrative or financial burden on the IHSAA." *Id.* (quoting Washington v. Indiana High Sch. Athletic Ass'n., 181 F.3d 840, 852 (7th Cir. 1999)). This theory by Judge Arnold is adopted by the Ninth Circuit and seems to eliminate the idea that the administrative and financial burden could be in the number of waivers of a specific rule that an organization has to look at and not the burden on individual waiver analysis. *See id.*

133. *See id.* at 999-1000.

134. *See id.*

135. *See id.*

136. PGA Tour, Inc., v. Martin, 204 F.3d 994 (9th Cir. 2000), *cert. granted*, 121 S. Ct. 30 (2000) (mem.).

137. *See* Olinger v. United States Golf Ass'n., 205 F.3d 1001, 1007 (7th Cir. 2000), *vacated*, 121 S. Ct. 2212 (2001).

138. *See id.* at 1001. "Everyone agrees he is a 'disabled' person within the meaning of the ADA." *Id.*

information that is legally irrelevant to the question at hand.[139]

Eventually, the opinion begins its analysis with the question of whether or not the course the tournaments are played on are subject to Title III of the ADA.[140] This analysis by Judge Evans is relatively short because he chooses to answer the question on a simpler, narrower basis than having to analyze the more difficult question inherent within the ADA "public accommodation" requirement.[141] The Court reluctantly agrees with the United States Golf Association that golf courses during tournament play are "mixed-use"[142] facilities and the "inside the ropes" playing area is not subject to the ADA,[143] but refuses to include this in the holding of the case.[144] Instead, the focus of Judge Evans' opinion in this case falls on whether the use of a golf cart in USGA sponsored tournaments "funda-

---

139. *See id.* at 1001-04. (discussing such topics as who has won past USGA tournaments as well as the awesome scope of the tournaments as a whole). "Its venues are true meccas of tournament golf, places liked Winged Foot (Mamaroneck, New York); Medinah (Medinah, Illinois); Shinnecock Hills (Southampton, New York); Merion (Ardmore, Pennsylvania); and Congressional (Bethesda, Maryland)." *Id.* at 1002 n.1. "Past winners of the U.S. Open include legends in the game: Bobby Jones (1923, 1926, 1929, and 1930); Gene Sarazen (1922, 1932); Byron Nelson (1939); Ben Hogan (1948, 1950, 1951, and 1953); Arnold Palmer (1960); Jack Nicklaus (1962, 1967, 1972, and 1980); Gary Player (1965)[;] Hale Irwin (1974, 1979, and 1990); Tom Watson (1982); and Lee Trevino (1968, 1971)." *Id.* at 1002.

140. *See Olinger*, 205 F.3d at 1004.

141. *See id.* at 1005.

142. *See infra* notes 197-219 and accompanying text for a discussion on this topic.

143. *See Olinger*, 205 F.3d at 1005. "The USGA concedes that a golf course is listed as one of the examples of a place of public accommodation under the ADA . . . [b]ut . . . [only] when it is used for 'exercise or recreation' . . . ." *Id.* The USGA relies heavily on the Justice Department regulations implementing Title III for their argument. *See id.* at 1005. This provision was also cited by the PGA in a "mixed-use" facility argument against the application of the ADA to their tournaments, but was rejected by the Ninth Circuit. *See also infra* notes 198-220 and accompanying text for a discussion on this topic.

144. *See Olinger*, 205 F.3d at 1007. "While there may be some logic to this contention, we hesitate to embrace it for we can resolve this appeal, as did the District Court, on a more narrow ground." *Id.* at 1005. "The focus of our opinion has been on one question: *Must* the USGA allow Ford Olinger to compete while riding in a golf cart instead of walking." *Id.* at 1007. It seems as though this question is not limited to the discussion of whether the alteration "fundamentally alters" the nature of the game, as the Seventh Circuit would like to believe, but it is the broad question which requires the analysis of all parts of Title III of the ADA. This includes the question of whether a golf course during a USGA tournament is subject to the ADA's Title III provisions.

mentally alters" the nature of the tournament itself.[145]

The Seventh Circuit buttresses its conclusion heavily upon the concept of whether the use of a golf cart "fundamentally alters" the nature of the tournament offered by the USGA.[146] The court concludes that since the Rehabilitation Act of 1973[147] circuit courts have not required entities to "change their basic nature, character, or purpose insofar as that purpose is rational, rather than a pretext for discrimination."[148]

The Court relies on the district court's factual findings that the nature of the game would be fundamentally altered if Ford Olinger were permitted to ride in a golf cart during tournament play.[149] These factual findings included "'conditions that now affect a golfer's performance, but which lie beyond the golfer's ability to control--the fatigue born of the hills, of heat, of humidity--would lessen in importance to the competition.'"[150] The district court came to these conclusions, and the Appeals Court agreed, through various testimonies of select doctors and professional golfers, including Ken Venturi.[151] Venturi gave testimony about the days when he played in the 1940's – 1960's, and various situations where golfers, including himself, could have desperately needed the use of carts, but stated that the USGA would not have even thought about making an accommodation for these golfers, and even if they had, the golfers probably would have refused and risked serious injury to be able to walk the course.[152]

---

145. *See Olinger*, 205 F.3d at 1005.

146. *See id.*

147. *See supra* notes 48-56 and accompanying text.

148. *Olinger*, 205 F.3d at 1005, *cited in* Sandison v. Michigan High Sch. Athletic Ass'n, 64 F.3d 1026, 1035 (6th Cir. 1995); Pottgen v. Supra Missourri State High Sch. Activities Ass'n, 40 F.3d 926 (8th Cir. 1994).

149. *See* Olinger, 205 F.3d at 1006.

150. *Id.* (quoting Olinger v. United States Golf Ass'n, 55 F. Supp. 2d 926 (1999)). Also included in the district court's findings regarding the effect of altering the competition by Ford Olinger using a cart was "'the point of an athletic competition,'" which is "to decide who, under conditions that are about the same for everyone, can perform an assigned set of tasks better than (not as well as) any other competitor. The set of tasks assigned to the competitor in the U.S. Open includes not merely striking a ball with precision, but doing so under greater than usual mental and physical stress." *Id.*

151. *See Olinger*, 55 F. Supp. 2d at 936. One golfer's testimony the appeals court found particularly persuasive was that of Ken Venturi. The testimony consisted of a situation in the 1964 U.S. Open where, against the better judgment of his physician, Venturi went out to play the afternoon 18 and was suffering from dehydration in 100°C heat and 97% humidity. "In the stifling heat and humidity, battling dehydration, Venturi walked the course and, on the verge of collapse, won the tournament." *Olinger*, 205 F.3d at 1006.

152. *See Olinger*, 205 F.3d at 1007; *see also supra* note 151 and accompanying text. Venturi went on to testify about the great Ben Hogan. Hogan had a se-

The Seventh Circuit also relied on testimony by Dr. Theodore Holland at trial to find that Ford Olinger's use of a golf cart during tournament play would fundamentally alter the nature of the competition provided by the USGA.[153] The Doctor testified that golf is a demanding sport that includes physically maneuvering the course, expending energy and stamina, which tests a player's physical endurance levels.[154]

With these evidentiary findings by the district court, the Seventh Circuit concluded that to order the USGA to allow a player to ride in a golf cart would "fundamentally alter" the nature of the competition as the USGA has offered it and therefore they are not required to do so.[155]

The Seventh Circuit briefly, in one paragraph, talks of the administrative burdens that would be placed upon the USGA if they were required to evaluate who is, and is not, entitled to use a golf cart because of their disability.[156] The court of appeals sided with the district court's rationale: ". . . the USGA 'would need to develop a system and a fund of expertise to determine whether a given applicant truly needs, or merely wants, or could use but does not need, to ride in a cart to compete.' The district court thought that this should be unnecessary. We agree."[157]

The Seventh Circuit affirmed the district court ruling, which denied Ford Olinger use of a golf cart in USGA tournaments because it would "fundamentally alter" the nature of the competition as defined by the USGA.

> And without question, most players would prefer to walk while playing competitive, championship-caliber golf. Surely a player like Olinger would gladly trade in his cart if he could walk a course without pain. But the decision on whether the rules of the game should be adjusted to accommodate him is best left to those who hold the future of golf in trust. Because the law does not force the USGA to make the accommodation Olinger seeks, the judgment of the district court is AFFIRMED.[158]

---

vere car accident in 1949 and was told he would probably never walk again, let alone play golf. Despite all of this Hogan went on to play the following year's U.S. Open and win, while walking. *See id.*

153. *See* Olinger v. U.S. Golf Ass'n, 205 F.3d 1001, 1006 (7th Cir. 2000).

154. *See id.* Dr. Theodore Holland's own words were "[t]here is a lot more to getting . . . around those 72 holes than just hitting the shots." *Id.*

155. *See id.* at 1007.

156. *See id.*

157. *Id.*

158. *Id.*

IV. UNITED STATES SUPREME COURT GRANTS CERTIORARI

Casey Martin & Ford Olinger

The Battle Continues

A. Introduction

The issues before the Supreme Court relevant to the Martin and Olinger cases are cloudy at best.[159] There are several definite issues that will be focused upon in settling the dispute brought by the PGA. First, is the legal question of whether the PGA operates a place of public accommodation, therefore subjecting them to the application of the ADA.[160] If the Supreme Court answers that question in the affirmative, then the next question would likely be whether the use of a golf cart by Martin would "fundamentally alter" the nature of the game.[161] This second question raises more disputes among analysts because it seems to be a clearly factual rather than legal question.[162] The final issue, which should only be reached if the previous two were answered "yes" and "no" accordingly, is whether the administrative and financial burden would be too great upon the PGA in carrying out Martin's and others' requests so as to alleviate the PGA from compliance.[163] This question should only be reached in a last ditch effort by the PGA to escape providing Casey Martin's requested accommodation.

B. "Public Accomodation"

Public accommodation is defined as "[a] business establishment that provides lodging, food, entertainment, or other services to the public; esp. (as defined by the Civil Rights Act of 1964), one that affects interstate commerce or is supported by state action."[164] Applied broadly, this definition could mean anything and everything that touches interstate commerce, with some very limited exceptions.[165] The history and use of "public ac-

---

159. *See generally* Jeffery A. Rosenthal, *Should High Court Review Case of Disabled Golfer: Despite Publicity, Many Factors, Including Primarily Factual Nature of Dispute, Suggest a Negative Answer*, 224 N.Y.L.J. Aug. 7, 2000 at 9.

160. *See id.* at 10.

161. *See id.*

162. *See id.*

163. *See* Martin v. PGA Tour, Inc., 204 F.3d 994, 1001 (9th Cir. 2000) *accord Olinger*, 205 F.3d at 1007.

164. BLACK'S LAW DICTIONARY 1242 (7th ed. 1999).

165. *See generally* United States v. Lopez, 115 S. Ct. 1624 (1995) (holding

commodation" may provide insight to the questions presented in this article.

The ADA's definition encompasses the phraseology used above in including in the definition of public accommodation; "[t]he following private entities are considered public accommodations for the purposes of this subchapter, if the operations of such entities affect commerce . . . ."[166] In adopting this definition it is necessary to look into the idea of public accommodations affecting interstate commerce.

The two leading cases in the area were decided by the Supreme Court on the same day, December 14, 1964.[167] These cases opened the door for Congress to enact broader legislation using the commerce clause[168] of the Constitution as their vehicle, that otherwise could not have been enacted.[169]

Heart of Atlanta Motel was an action by the owner of a motel seeking to enjoin the enforcement of the "public accommodation" portion of the Civil Rights Act of 1964.[170] The policy of the motel was to refuse to rent any of the 216 rooms available to people of color.[171] The language inherent in the Civil Rights Act of 1964 prohibited discrimination towards people of color in any place of public accommodation.[172] The Act defines public accommodations as "[e]stablishments affecting interstate commerce . . . ."[173] The Civil Rights Act of 1964 then proceeds to devote a section to delineate what criteria an entity needed to fit in order to affect interstate commerce.[174]

---

that Congress could not extend its power through the commerce clause to regulate gun possession in and around schools).

166. 42 U.S.C. § 12181(7) (1995); *See also supra* note 77 for a partial list of entitities.

167. *See* Heart of Atlanta Motel, Inc. v. United States, 379 U.S. 241 (1964); *See also* Katzenbach v. McClung, 379 U.S. 294 (1964).

168. U.S. CONST., art. I, § 8, cl. 3. "The Congress shall have the Power . . . [t]o regulate commerce with foreign Nations, and among the several States, and with the Indian Tribes." *Id.*

169. *See generally Heart of Atlanta Motel, Inc.* 379 U.S. 241; *Katzenbach*, 379 U.S. 294.

170. *See Heart of Atlanta Motel, Inc.*, 379 U.S. at 242.

171. *See id.* at 243.

172. *See* 42 U.S.C. § 2000a (a) (1994). "All persons shall be entitled to the full and equal enjoyment of the goods, services, facilities, privileges, advantages, and accommodations of any place of *public accommodation* . . . without discrimination or segregation on the ground of race, color, religion, or national origin." *Id.* (emphasis added).

173. 42 U.S.C. § 2000a(b) (1994).

174. *See* 42 U.S.C. § 2000a(c) (1994).

> The operations of an establishment affect commerce within the meaning of this subchapter if: (1) It is one of the establishments described in paragraph (1) of subsection (b) of this section;[175] (2) in the case of an establishment described in paragraph (2) of subsection (b) of this section, it serves or offers to serve interstate travelers or substantial portion of the food which it serves, or gasoline or other products which it sells, has moved in commerce;[176] (3) in the case of an establishment described in paragraph (3) of subsection (b) of this section, it customarily presents films, performances, athletic teams, exhibitions, or other sources of entertainment which move in commerce . . . .[177] For the purposes of this section, 'commerce' means travel, trade, traffic, commerce, transportation, or communication among the several states, or between the District of Columbia and any State . . . .[178]

It was this criteria that the petitioner sought to have stricken from the Civil Rights Act of 1964 as unconstitutional, and therefore inapplicable to his establishment, so that he might continue to refuse people of color from frequenting his motel.[179]

The Supreme Court opined that congressional power was broad enough under the commerce clause of the Constitution to reach the motel.[180] In defining what is subject to congressional power, through interstate commerce, using the commerce clause the Court cites an 1824 case, *Gibbons v. Ogden*,[181] which began the expansion of Congressional power in this area.

---

175. Section (b) includes "any inn, hotel, motel, or other establishment which provides lodging to transient guests, other than an establishment located within a building which contains not more than five rooms for rent or hire and which is actually occupied by the proprietor of such establishment as his residence." 42 U.S.C. § 2000a (b)(1) (1994).

176. The list is continues: "any restaurant, cafeteria, lunchroom, lunch counter, soda fountain, or other facility principally engaged in selling food for consumption on the premises, including, but not limited to, any such facility located on the premises of any retail establishment; or any gasoline station." 42 U.S.C. 2000a(b)(2) (1994).

177. The list continues: "any motion picture house, theatre, concert hall, sports arena, stadium *or other place of exhibition* or entertainment." 42 U.S.C. 2000a(b)(3) (1994) (emphasis added).

178. 42 U.S.C. § 2000a(c) (1994).

179. *See generally Heart of Atlanta Motel, Inc.*, 379 U.S. at 241.

180. *See id.* at 261.

181. 9 Wheat. 1, 196, 197 (1824). In this case, the Supreme Court held that commerce,

> is something more: it is intercourse between nations . . . and parts of nations, in all its branches, and is regulated by prescribing rules for carrying on that intercourse. To what commerce does this power extend? . . . [I]t may very properly be restricted to that commerce which concerns more states than one . . . . What is this power? It is the power to regulate; that is to prescribe the

The Court uses the broad power of Congress to use the commerce clause in regulation given by Justice Clark in *Ogden* to relieve the problem in *Heart of Atlanta Motel*. In *Ogden*, the opinion expanded the use of the commerce clause by deciding that Congress had the power to regulate all business which impacted interstate commerce, even those businesses which appear to be local in origin.[182]

Katzenbach v. McClung[183] was the second case decided in the same day on the same issue. In *Katzenbach*, Ollie's Barbecue, a family owned restaurant, serviced white customers in its sit-down area and provided a take-out service only for people of color.[184] In *Katzenbach*, the Court applied their opinion of *Heart of Atlanta Motel* and stretched it somewhat further. In light of the fact that the restaurant acquired much of its food from outside the state through various wholesalers it was deemed that their activity affected interstate commerce so as to subject them to the provisions of the Civil Rights Act of 1964.[185]

The Supreme Court in both of these cases upheld the application of the Civil Rights Act of 1964 to establishments that had some, albeit small, impact on interstate commerce.[186] Whether it was because travelers moving from state to state were impacted[187] or supplies for a restaurant traveled interstate[188] the Supreme Court held that it was within Congressional power to regulate those businesses.[189] Surely, these two famous cases should have some impact on the ADA.

The ADA adopts very similar language to that of the Civil Rights Act of

---

> rule by which commerce is to be governed. This power, like all others vested in Congress, is complete in itself, may be exercised to its utmost extent, and acknowledges no limitations other than are prescribed in the constitution . . . . [T]he determinative test of the exercise of power by the Congress under the Commerce Clause is simply whether the activity sought to be regulated is 'commerce which concerns more States than one' and has a real and substantial relation to the national interest.

Heart of Atlanta Motel, Inc., 379 U.S. at 254, 255 (quoting Gibbons v. Ogden, 9 Wheat. 1, (1824)).

182. *See Heart of Atlanta Motel, Inc.*, 379 U.S. at 257-58.

183. 379 U.S. 294 (1964).

184. *See Katzenbach*, 379 U.S. at 296.

185. *See id.* at 304-05.

186. *See generally Heart of Atlanta Motel, Inc.*, 379 U.S. 241; *Katzenbach*, 379 U.S. 294.

187. *See Heart of Atlanta Motel, Inc.*, 379 U.S. at 258.

188. *See Katzenbach*, 379 U.S. at 304.

189. *See generally Heart of Atlanta Motel, Inc.*, 379 U.S. at 260; *Katzenbach*, 379 U.S. at 304-05.

1964.[190] Given the similarity in language as well as the incorporation of the public accommodation language in both instruments, it would be a poor assumption to believe that the wide breadth of *Heart of Atlanta Motel* and *Katzenbach* in applying a congressional discrimination statute would not apply today for the ADA.[191]

The PGA, under the above standard, certainly falls within the category of "affecting commerce" and thus should be held to the standards of the ADA. Nationally televised events, nationwide sponsorship of tournaments, players from around the world in competition; Certainly if an entity can be held accountable under the public accommodation language by affecting interstate commerce through purchases of supplies from a wholesaler who attains them from out of the state, the PGA can be as well.

There is another argument made by the PGA that, although the statute is broadly defined using *places of public accommodation that effect interstate commerce*, the statute itself provides limiting language.[192] The statute provides a list of private entities that are covered within the act, and a golf course is named as one of those entities subject to the ADA.[193] The question is whether the language "or other places of exercise or recreation" is intended to limit the ADAs' application to golf courses as a place of recreation and not a place of competition; or is there no difference?[194] This question should be answered by looking at legislative intent in application

---

190. 42 U.S.C. § 12182(a) (1995); *compare with* 42 U.S.C. 2000a (a) (1995). A side by side comparison of the two provisions would lead one to believe that it was a simple edit, substituting "race, color, religion, or national origin" for "on the basis of disability."

191. Both Acts seek to eliminate discrimination among individuals in the United States and both are worded similarly to achieve broad application. In a senate hearing it was said:

> For too long valuable resources available to this nation from individuals with disabilities have been wasted needlessly . . . . [S]enseless discrimination, intended or not, which subjected persons with disabilities to isolation and robbed America of the minds, the spirit, and the dedication we need to remain a competitive force . . . . Today we are going to unlock these resources through the ADA and bring Americans with disabilities into the mainstream of the economic structure of this country.

136 CONG. REC. S9684, S9685 (1990) (debate by Senator Hatch (Utah).

192. *See* Martin v. PGA Tour, Inc., 204 F.3d 994, 997 (9th Cir. 2000).

193. *See* 42 U.S.C.A. § 12182(7)(L) (1995). "[G]olf course, or other place of exercise or recreation." *Id.*

194. *See* 42 U.S.C.A. § 12181(7)(L) (1995).

of the ADA.[195] Senator Harkin (D-Iowa), stood on the steps of the Supreme Court the day of oral arguments with Casey Martin and Senator Bob Dole, as he applauded the applicability of the ADA to Casey Martin's case.[196]

The answer to this question should be answered in a way as to not limit the application of the ADA. Based on the Supreme Court's application in applying the broad power of the commerce clause it would be oxymoronic to pick a few words out of the statute, and interpret them as limiting the application of the ADA to the PGA.

1. Mixed Use Facility

But what of the arguments that the PGA raises in contention that they fall outside the scope of the ADA's "public accommodation" language through exceptions provided within the ADA itself?[197]

The first of the PGA's arguments deal with the contention that they fall within the classification of a "mixed-use" facility.[198] The conceptual framework outlining a "mixed-use" facility can be found in the Justice Department Regulations implementing Title III of the ADA.[199] The Justice Department specifically claims that "[m]any facilities . . . are mixed use facilities."[200] Two examples are given in this appendix to the ADA by the Justice Department that the PGA cites as applicable to their case.[201]

The first is classified as a "residential hotel," which has two distinct uses. First, a part of the hotel is used by transient guests which falls under the "places of lodging" provision of the ADA.[202] The second half of the hotel is used for long-term residence, which the Justice Department claims could be covered by the Fair Housing Act,[203] not the ADA. This consti-

195. *See* Olmstead v. Zimring, 527 U.S. 581, 588-602 (1999) (Justice O'Connor spends this entire section interpreting the legislative intent of the ADA in applying Title II).

196. *See supra* note 5; *see also* 136 CONG. REC. S9684, S9691. (Senator Tom Harkin was the leading voice and constructionist of the ADA).

197. *See* Martin v. PGA Tour, Inc., 984 F.Supp. 1320, 1326 (D. Or. 1998).

198. *See* Martin v. PGA Tour, Inc., 204 F.3d 994, 997 (9th Cir. 2000); *see also* 28 C.F.R. pt. 36, app. B at 623 (1999).

199. *See* 28 C.F.R. pt. 36, app. B at 623 (1999). This provision of the Justice Department's regulations conclude that an entity can be both a public accommodation and a private entity not subject to the ADA provisions. *Id.*

200. *Id.* at 623.

201. *See* 28 C.F.R. pt. 36, app. B at 624 (1999).

202. *See id.*

203. Fair Housing Act of 1968 *codified in* 42 U.S.C. § 3601 (1995).

tutes, as the PGA argues, a "mixed-use" facility.

The second example provided by the Justice Department is that of a commercial facility that runs public tours through its grounds.[204] It is maintained that with relation to the tour itself, any part of the facility to which the public has access, the ADA would apply and would govern the standard of accessibility to the tour.[205] All portions of the entity that are not accessible to the public would not fall under the provisions of the ADA as a "public accommodation,"[206] and accessibility to these areas would not be covered within the statute.[207]

The PGA uses these examples provided by the Justice Department to claim "mixed-use" facility status, claiming that the portions of the course in which the players are competing, "within the ropes,"[208] is equal to the "long-term residence" and "non-public tour" portions of their facility, namely the country club.[209] This argument is easily discarded by the Martin Court by noting that the PGA falsely believes that the competition ground cannot be a place of public accommodation simply because it is highly competitive.[210] The Ninth Circuit decision goes onto state that the ADA provides applicability to "'secondary, undergraduate, or postgraduate private schools'" and that simply because they are highly competitive does not mean that once accepted, the students cease to become members of the public using the university as a public accommodation.[211]

Courts have generally dealt with the "mixed-use" argument within the realm of zoning, as opposed to "mixed-use" facilities within the Justice Department Application to the ADA.[212] A few courts, including the Martin Court, have dealt with the "mixed-use" facility argument.

Jankey v. Twentieth Century Fox Film Corporation[213] is the Ninth Cir-

---

204. *See* 28 C.F.R. pt. 36, app. B at 624 (1999).

205. *See id.*

206. *See id.*

207. *See id.*

208. *See* Martin v. PGA Tour, Inc., 204 F.3d 994, 998 (9th Cir. 2000). The argument outlined by the PGA here is that the 'within the ropes' portion of the golf course is the actual area of the course in which the players compete. *Id.* However, it seems that golfers have played their lie from among the crowd implying that sometimes the crowd is 'within the ropes as well'. *Id.*

209. *See id.*

210. *See id.*

211. *See id.*

212. *See generally* Phillips v. Borough of Keyport, 107 F.3d 164 (3d Cir. 1997); *see also* Town of North Bonneville, Wash. v. U.S., 11 Cl.Ct. 694 (Cl.Ct., 1987); *Cf.* Centerfold Club, Inc. v. City of St. Petersburg, 969 F.Supp. 1288 (M.D.Fla., 1997); *Cf.* Schenck v. City of Hudson Village, 937 F.Supp. 679 (N.D.Ohio 1996).

213. 212 F.3d 1159 (9th Cir. 2000).

cuit's leading decision on "mixed-use" facilities. This case dealt with a review de novo[214] of plaintiff's claim that he was discriminated against by defendant in violation of Title III of the ADA because he was denied access to a filming lot where the public was not permitted.[215] The plaintiff wished to gain access to the commissary, studio store and an automatic teller machine.[216] The question on appeal was whether an entity listed under 42 U.S.C. section 12181(7), can be subject to the act even though its facilities are not open to the public.[217] The court rejects petitioners argument that because Twentieth Century Fox falls within a defined category specified in 42 U.S.C. section 12181(7) that it is automatically subject to the provisions of the ADA. "This argument . . . ignores the plain language of § 12187, which states: 'The provisions of [Title III] shall not apply to private clubs or establishments exempted from coverage under Title II of The Civil Rights Act.'"[218]

What of the contention that the USGA and PGA fall within the private club exemption[219] outlined by the Ninth Circuit in Jankey?

2. Private Club Exemption

The PGA might possibly contend that it falls within the *private club exemption* of the ADA, and therefore the Act does not apply.[220] The Code of Federal Regulations, which defines private clubs, refers to the Civil Rights Act of 1964[221] for entities that would qualify for exemption under the ADA.[222] Following these directions, a look at the Civil Rights Act of 1964 leads us to this:

> Private establishments
>
> The provisions of this subchapter shall not apply to a private club or other establishment not in fact open to the public, except to the extent that the fa-

---

214. "An appeal in which the appellate court uses the trial court's record but reviews the evidence and law without deference to the trial court's rulings." BLACK'S LAW DICTIONARY 94 (7th ed. 1999).

215. *See* Jankey v. Twentieth Century Fox Film Corp., 212 F.3d 1159, 1160 (9th Cir. 2000).

216. *See id.* at 1161.

217. *See id.*

218. *Id.*; *See generally* Louie v. Ideal Cleaners, et al., No. C-99-1814, 1999 WL 1269191 (N.D.Cal 1999).

219. *See* 42 U.S.C.A. § 12187 (1995).

220. *See id.*

221. *See* 42 U.S.C. § 2000a (1995).

222. *See* 28 C.F.R. § 36.104 at 507 (2000). "*Private club* means a private club or establishment exempted from coverage under title II of the Civil Rights Act of 1964 (42 U.S.C. 2000a(e))." *Id.*

> cilities of such establishment are made available to the customers or patrons of an establishment within the scope of subsection (b) of this section.[223]

The Supreme Court in *Daniel v. Paul*[224] denied private club status to a recreational facility that provided swimming, boating and miniature golf, among other things, to members, but denied membership to people of color.[225] The Court concluded that the recreational facility was not a private club: "[i]t is simply a business operated for a profit with none of the attributes of self-government and member-ownership traditionally associated with private clubs."[226] In another similar case, the Supreme Court again denied private club status to a swimming pool association.[227] In *Tillman v. Wheaton-Haven Recreation Association*,[228] the Court looked at the swimming pool associations by-laws, which allowed anyone living within a three-quarter mile radius of the pool to become a member without any recommendations from already existing members.[229] The association had denied membership to a black couple who had bought a home within the prescribed area.[230] The Court held that where the membership was open to every white person in the geographic area, that the association was public enough as to fall outside the confines of the *private club* exemption.[231]

In light of the *Tillman* case, the PGA's argument to establish itself as a private club is a valiant, but failing effort. The PGA gives out memberships, as the district court in *Martin* recognizes, but anyone in the public who can pay an application fee and submit two letters of recommendation can attempt to qualify for membership, unless of course you are disabled.[232]

With the *private club exemption* eliminated from the PGA's arsenal of weapons against the ADA, the applicable subsection for private club exemptions under the Civil Rights Act of 1964[233] sends the inquiry back to subsection (b), to the discussion of public accommodations as defined by

---

223. 42 U.S.C. § 2000a(e) (1995).

224. 395 U.S. 298 (1969).

225. *See id.* at 302.

226. *Id.* at 301.

227. *See* Tillman v. Wheaton-Haven Recreation Ass'n Inc., 410 U.S. 431, 438-39 (1973).

228. *See id.* at 431.

229. *Id.* at 433.

230. *Id.* at 434.

231. *See id.* at 438.

232. *See* Martin v. PGA Tour, Inc., 204 F.3d 994, 999 (9th Cir. 2000); *see also supra* note 106 and accompanying text.

233. *See* 42 U.S.C. § 2000a(e) (1995).

the Civil Rights Act of 1964, discussed and answered above.[234]

C. Reasonable Accomodations

According to the language of the ADA, the accommodation sought by a disabled individual must be a reasonable one.[235] Courts seem to take a position that the burden of showing the reasonableness of a suggested accommodation falls on the plaintiff seeking to have that accommodation implemented.[236] There is no accepted test that the courts may turn to determine if the requested accommodation is reasonable.[237] In general, the inquiry lies on a specific case-by-case analysis of each plaintiffs' specific circumstances; whether or not the modification would be reasonable for that specific plaintiff. [238] The Supreme Court has even stressed that it is

---

234. *See* 42 U.S.C. § 2000a(b) (1995); *see also supra* notes 164-197 and accompanying text.

235. *See* 42 U.S.C. § 12182(b)(2)(a)(ii) (1995). "[F]ailure to make reasonable accommodations . . . when such modifications are necessary . . . to individuals with disabilities . . . ." *Id.* This language follows 42 U.S.C. 12182(b)(2)(A), which states "[f]or purposes of subsection (a), discrimination includes . . . ." *Id.*

236. *See* Stern v. University of Osteopathic Medicine and Health Sciences, 220 F.3d 906, 908 (8th Cir. 2000) (holding that where student failed to show that accommodation was reasonable through the use of expert testimony his claim should be denied); *see also* Garcia-Ayala v. Lederle Parenterals, Inc., 212 F.3d 638, 648 (1st Cir. 2000) (holding that the burden to show reasonable accommodation rests with the plaintiff); *see also* Johnson v. Gambrinus Company/Spoetzl Brewery, 116 F.3d 1052, 1059 (5th Cir. 1997) (holding that plaintiff may meet his burden of proving that accommodation is reasonable by introducing that it is reasonable in the general sense); *see also* Guckenberger v. Boston University, 974 F.Supp. 106, 146 (D. Mass 1997) (holding that the plaintiff has the initial burden of proving the modification requested is reasonable in the general run of cases); *see also* McPherson v. Michigan High School Athletic Association, 119 F.3d 453, 461 (6th Cir. 1994) (holding that plaintiff's claim depends on a showing that defendant could have reasonably accommodated him).

237. *See* Staron v. McDonald's Corp. and Burger King Corp., 51 F.3d 353 (2d Cir. 1995). "Although neither the ADA nor the courts have defined the precise contours of the test for reasonableness, it is clear that the determination of whether a particular modification is 'reasonable' involves a fact specific, case-by-case inquiry. . . ." *Id.* at 356. *See also* Johnson v. Gambrinus Company/Spoetzl Brewery, 116 F.3d 1052 (5th Cir. 1997). "[W]e have not been consistent in our interpretation . . . regarding the burden of proof on reasonableness." *Id.* at 1059 n.4.

238. *See Staron*, 51 F.3d at 356; *see also* Garcia-Ayala v. Lederle Parenterals, Inc., 212 F.3d 638 (1st Cir. 2000). Under the section titled "Individualized Assessment" in this case the court struggles with the fact that the lower court failed to focus on the facts from the individual standpoint. *Id.* at 647; *see also* Matthews v. National Collegiate Athletic Association, PAC 10 Athletic Conference, 79 F.Supp. 1199, 1204 (E.D. Wash. 1999). "When any court applies the ADA, it must examine facts specific to the case and the disabled individuals circumstances, and

essential to give "individualized attention" to disability claims.[239] Most courts have concluded that this requires each accommodation to be viewed in light of the plaintiff presenting the claim.[240] After this inquiry is made and found to be supportive in favor of the plaintiff, according to the statutory language of the ADA, the public entity must provide the individual with that reasonable accommodation.[241] The ADA does, however, provide a statutory escape clause to the implementation of the accommodation requested by the disabled individual.[242] Once the burden on the plaintiff has been met in providing evidence that the accommodation is a reasonable one, the burden shifts to the defendant to prove either that providing the accommodation would result in undue hardship or that providing the accommodation would fundamentally alter the nature of whatever the entity is trying to provide.[243]

1.Initial Burden

The initial burden to prove that the accommodation requested is reasonable lies with the plaintiff. The trial court in *Olinger* recognized that "[t]he USGA does not appear to challenge the reasonableness of Mr. Olinger's requested accommodation in the general sense, and the golf cart has become so ubiquitous in the sport that any such challenge would seem doomed."[244] Therefore the initial burden of *Olinger*, to show that the ac-

---

the accommodations that might allow him to participate." *Id.* at 1204.

239. *See Staron*, 51 F.3d at 356 (summarizing the Supreme Court decision in *School Bd. of Nassau Co. v. Arline*, 480 U.S. 273, 287 (1987)).

240. *See* Bryan L. Koepke, *The ADA and Professional Golf-Breaking Par for the Sake of Equality*, 38 WASHBURN L.J. 699, 718-19 (1998).

241. *See* 42 U.S.C. § 12182(b)(2)(A)(ii) (1995); *see also* Johnson v. Gambrinus Co./Spoetzl Brewery, 116 F.3d 1052, 1059 (5th Cir. 1997); Garcia-Ayala v. Lederle Parenterals, Inc., 212 F.3d 638, 646 (1st Cir. 2000).

242. *See* 28 C.F.R. § 35.164 (1995) (explaining the burdens of proof and what is and is not required under 42 U.S.C. § 12182(b)(2)(A)(iii)).

243. *See* 42 U.S.C. §12182(b)(2)(A)(iii) (1995) ("[U]nless the entity can demonstrate that taking such steps . . . would result in an undue burden"); 42 U.S.C. § 12182(b)(2)(A)(ii) (1995) ("[U]nless the entity can demonstrate that making such modifications would fundamentally alter the nature of the goods, services, facilities, privileges, advantages, or accommodations"); *see also* 28 C.F.R. § 35.164 "This subpart does not require a public entity to take any action that it can demonstrate would result in a fundamental alteration in the nature of a service, program, or activity or in undue financial and administrative burdens. . . . [A] public entity has the burden of proving that compliance with this subpart would result in such alteration or burdens." *Id. See generally* School Bd. of Nassau Co. v. Arline, 480 U.S. 273 (1987).

244. Olinger v. United States Golf Ass'n., 55 F. Supp. 2d 926, 934 (N.D. Ind. 1999).

commodation was reasonable in a general sense, was met at the trial court level. Likewise, the trial court in *Martin* made quick work of the determination that the requested modification was reasonable, therefore satisfying Martin's initial burden to show that accommodation requested was reasonable, in the general sense.[245] With both of the plaintiff's initial burdens out of the way it is now up to the defendants to present evidence of an affirmative defense, or else provide the accommodation requested.[246]

2. Affirmative Defense # 1: Administrative Burden

The administrative and financial burden aspect of the escape clause has been subject to a variety of interpretations. Effectively, if the PGA can show that it would be too costly, or would require too much manpower to implement the requested accommodation they can escape providing it. One source concludes that the burden cannot be limited to a financial one to render accommodations unreasonable and that it extends to accommodations that are substantially disruptive, unduly expensive or harmful to business needs.[247] Another court ends the determination at the "parade of horribles" vision that the Justices feared in *McPherson v. Michigan High School Athletic Ass'n.*[248]allowing a claim under the ADA to be sustained.[249] A parallel list of factors to be considered in determining undue hardship in employment cases can be found in the Code of Federal Regulations; they include, "(1) the nature and cost of the accommodation needed; (2) the overall financial resources of the facility involved . . . (4) the effect on expenses and the impact on the operation of the facilities; (5) the overall financial resources of the covered entity . . ." and others that are

---

245. *See* Martin v. PGA Tour Inc., 994 F. Supp. 1242, 1248 (D. Or. 1998) "'Plaintiff meets this burden by introducing evidence that the requested modification is reasonable in the general sense, that is, in the general run of the cases.' With respect to this element, the use of a cart is certainly not unreasonable in the game of golf. As noted hereafter, the Rules of Golf do not require walking . . . ." *Id.* (quoting Johnson v. Gambrinus Co., 116 F.3d 1052, 1059 (5th Cir. 1997)).

246. *See* 42 U.S.C. §§ 12182(b)(2)(A)(ii)-(iii) (1995).

247. *See Garcia-Ayala*, 212 F.3d at 649-50; *see also* Brief of the United States Golf Association as Amicus Curiae in support of Appellant, Olinger v. United States Golf Ass'n. 55 F. Supp. 2d 926 (N.D. Ind. 1999), *reprinted in* VA. J. SPORTS & LAW 110, 126-28 (1999).

248. 119 F.3d 453 (6th Cir. 1997).

249. *See id.* ("[R]equiring a waiver under these circumstances would impose an immense financial and administrative burden on the [defendant], by forcing it to make 'near impossible determinations'. . . . The Plaintiff would have us require waivers for all learning-disabled students who remain in school more than eight semesters. That, of course, would have the potential of opening the floodgates for waivers . . . "). *Id.* at 462.

relevant to employment cases.[250] In *Sandison v. Michigan High School Athletic Ass'n*,[251] the court notes that "the word 'modification' connotes moderate change . . . "[252] which leads one to believe any change that is more than moderate could be deemed an undue burden to the entity.[253]

To apply these somewhat sketchy doctrines to the claim against the golfing industry is not as difficult as one would surmise. Both in the *Martin* and *Olinger* cases the Courts did not spend a large amount of time considering these issues.

In the Ninth Circuit's opinion in *Martin*, it is argued by the PGA that it would be "far too burdensome for [them] to determine whether disabled individuals using carts would have an advantage over non-disabled walking competitors."[254] The *Martin* court wholly disagrees with that assertion and contends that the PGA's reliance on *Sandison* was made in error; the court also cites *McPherson* and *Pottgen v. Missouri State High School Activities Ass'n*,[255] in which the courts ruled similarly.[256] The *Martin* court

---

250. *See* Diane L. Kimberlin & Linda Ottinger Headley, *ADA Overview and Update: What Has the Supreme Court Done to Disability Law?* 19 REV. LITIG. 579, 637 (2000) (citing 29 C.F.R. § 1630.2(p)(2)).

251. 64 F.3d 1026 (6th Cir. 1995).

252. *Id.* at 1037. The court in *Sandison* was also confused as to who carried what burden of proof in Rehabilitation Act claims as noted in note 3, going only as far as to place "the burden of proving 'reasonable accommodation' on federal employers who are under an affirmative action obligation." *Id.* at 1031 (quoting Hall v. U.S.P.S., 857 F.2d 1073, 1080 (6th Cir. 1988)).

253. *Id.* at 1037.

254. Martin v. PGA Tour, Inc., 204 F.3d 994, 1001 (9th Cir. 2000).

255. 40 F.3d 926 (8th Cir. 1994).

256. In all three cases the plaintiffs sought to lift the eligibility requirements that their school had imposed in order to play high school athletics. In *Sandison*, the plaintiffs challenged a rule that 19-year olds were unable to compete in interscholastic sports. When plaintiffs were in elementary school, they were held back due to a learning disability and ultimately when they entered high school were two class levels behind where they normally should have been. When they reached their senior year of high school they had turned 19 and were ineligible to play interscholastic sports. The Sixth Circuit concluded that because the age eligibility rule was necessary and therefore *essential* that the plaintiffs were not "otherwise qualified individuals" as the ADA requires. *See* Sandison v. Michigan High School Athletic Ass'n., 64 F.3d 1026, 1028, 1036, 1037 (6th Cir. 1995). Likewise, in *McPherson*, the plaintiff brought suit against a school district for not waiving an eight semester eligibility rule that prohibited students who had completed eight semesters in high school from playing interscholastic sports. Plaintiff, who was learning disabled, had to repeat his junior year in high school. When he reached his senior year, he had already completed his eighth semester and therefore was ineligible to play interscholastic sports. Again, the Sixth Circuit relied on its decision in *Sandison* that the eight semester rule was necessary and the plaintiff was not an 'otherwise qualified individual.' However, it also added that it would be an

concludes that those cases are based on the challenged rules being necessary, and therefore essential to the program, and on that basis the plaintiffs were not "qualified individuals" as the ADA language, appearing in Title I, requires them to be.[257] *Sandison* and *McPherson* both mention that the undue burden associated with individual determinations concerning the age eligibility requirement would be exactly what the ADA would have been addressing in its escape clause, but the *Martin* court chose to ignore this.[258] They chose to look at both the necessity of the rule[259] and the undue burden it would place on the entity from an individual standpoint and conclude that the burden is not intolerable, and therefore not undue, for the PGA.[260]

The *Olinger* court took quite a different spin on deciding whether the requested modification would be an undue financial or administrative burden on the USGA. As the Ninth Circuit in *Martin* deferred to the trial court on rulings of undue financial and administrative burden, so too did the

---

undue administrative and financial burden for the school to have to evaluate each student who requested a waiver. *See* McPherson v. Michigan High School Athletic Ass'n., 119 F.3d 453, 455, 461-63 (6th Cir. 1997). Finally, in *Pottgen*, a student received a preliminary injunction allowing him to play baseball in violation of the age limitation placed upon interscholastic sports in Missouri. The Missouri State High School Activities Association appealed the decision and won its case on appeal. The Eighth Circuit concluded that the eligibility requirement based on age was necessary and therefore essential to the program. Based on the language of the ADA the student was not an "otherwise qualified individual" as required. *See* Pottgen v. Missouri State High School Activities Ass'n., 40 F.3d 926, 928-29 (8th Cir. 1994).

257. *See Martin*, 204 F.3d at 1001-02. The ADA requires that an individual asserting a claim under its auspices must be an "otherwise qualified individual." *See* 42 U.S.C. § 12131(2) (1995). Under the statute, "[t]he term 'qualified individual with a disability' means an individual with a disability who, with or without reasonable modifications to rules, policies, or practices, the removal of architectural, communication, or transportation barriers, or the provision of auxiliary aids and services, *meets the essential eligibility requirements* for the receipt of services or the participation in programs or activities provided by a public entity." *Id.* (emphasis added). The *Martin* court argues that in all three cases cited by the PGA the decisions were made because the plaintiff was deemed not to be a "qualified individual with a disability." *Martin*, 204 F.3d at 1002.

258. *See Sandison*, 64 F.3d at 1035.

259. For further discussion on the necessity argument of the PGA and the "Walking Rule," *see infra* notes 296-305.

260. *See Martin*, 204 F.3d at 1002. "Nothing in the record establishes that an individualized determination would impose an intolerable burden on the PGA. . . . [T]he district court appeared to have little difficulty making the factual determination that providing Martin with a golf cart would not give him an unfair advantage over his competitors." *Id.*

Seventh Circuit in the *Olinger* case.[261] As a matter of law on this issue there is no argument; the legal standard is clear. Therefore, the inquiry was a factual one. The standard of review on appeal for a factual inquiry by the trial court is a "clearly erroneous" standard.[262] Both the *Olinger* and the *Martin* courts did not disturb the decisions on these points by the lower courts.[263]Unfortunately, the decisions by the lower courts in both cases were not fact-based determinations. A fundamental difference existed in the way each district court approached the question of undue burden of the modification. The United States District Court, Northern District of Arizona in *Olinger*, chose to look at the requested modification as it would unduly burden the entity as a whole; to set up a process for deciding who was eligible for cart use.[264] This approach took a look at what ripples would flow through the golfing world if golfers were allowed to utilize carts because of their respective disabilities; the court concluded that the impact would be far too burdensome for the USGA.[265] Both administra-

---

261. *See* Olinger v. United States Golf Ass'n, 205 F.3d 1001, 1007 (7th Cir. 2000). "The district court also offered a second rationale for ruling in favor of the USGA, the administrative burdens of evaluating requests to waive the walking rule and permit the use of a golf cart. As the court explained, the USGA 'would need to develop a system and a fund of expertise to determine whether a given applicant truly needs, or merely wants, or could use but does not need, to ride a cart to compete.' The district court thought that this should be unnecessary. We agree." *Id.* (quoting the district court in *Olinger v. United States Golf Ass'n*, 55 F. Supp. 2d 926, 937 (N.D. Ind. 1999)).

262. *See* JACK H. FRIEDENTHAL & MARY KAY KANE, CIVIL PROCEDURE 612-25 (3d ed. 1999). "[I]t is clear that the use of a clearly erroneous standard reflects a very narrow scope of review . . . . [T]he trial judge is thought to have an advantage over the appellate court because of his opportunity to view the witnesses; demeanor evidence is of course unavailable to the appellate court." *Id.* at 621-22.

263. *See Olinger*, 205 F.3d at 1007; *see also* Martin v. PGA Tour, Inc., 204 F.3d 994, 1002 (9th Cir. 2000).

264. *See Olinger*, 55 F. Supp. 2d at 937. "Once the inquiry moves beyond Mr. Olinger, the issue broadens and becomes far more difficult. Ford Olinger must ride a golf cart to play, and even with a cart he is likely to be more fatigued at the day's end than a healthy Tiger Woods or a healthy David Duvall. Perhaps it will be easy to say the same thing about next year's applicant for cart use—but how will that applicant be compared to Ford Olinger? Will next year's applicant for permission to ride have a competitive advantage over Mr. Olinger if allowed to ride? Will Mr. Olinger have a competitive advantage over next year's applicant if both are allowed to ride? If either would have an advantage over the other, would one be allowed to ride? Or both? Or neither? . . .A court deciding a case under the ADA must make an individualized decision concerning the plaintiff, but a court evaluating a requested accommodation's impact on the nature of the activity or program also must give full consideration to that impact." *Id.*

265. *See Olinger*, 55 F. Supp. 2d at 937.

tively and financially, the court felt as though the USGA could not put a system in place to evaluate golfers on the basis of their eligibility for alterations based on their disability, this was the reasoning followed in *Sandison*, *McPherson* and *Pottgen*.[266] The United States District Court, District of Oregon in *Martin*, chose to look at how the golf industry, specifically the PGA, would be impacted by the requested modification of Casey Martin, and Casey Martin alone.[267] The difference between the two court's analysis is whether the modification requested would be an undue administrative or financial burden on the entities' services in general or whether the burden from that individual would be undue.

This poses an interesting question for the Supreme Court if they happen to reach this stage of the analysis. The trial records of *Martin* and *Olinger*, based on the district courts' opinions are devoid of sufficient factual evidence to determine this issue.[268] Surely with the resources available to the PGA and the United States Golf Association, it would be difficult for them to show that it would be too heavy a burden to make an inquiry as to the eligibility of disabled golfers. This is a small price to pay for equality.

3. Affirmative Defense #2: Fundamentally Alter

The second aspect of the escape clause allows an entity to deny reasonable accommodations when the accommodations requested would fundamentally alter the nature of the goods, services, facilities, privileges, advantages, or accommodations that the public accommodation is providing.[269] This provision has been deemed an affirmative defense, thereby putting the burden of proof on the defendant entity, by the Code of Federal Regulations and accompanying case law.[270] The concept of "fun-

266. *See* Sandison v. Michigan High Sch. Athletic Ass'n., 64 F.3d 1026, 1035 (6th Cir. 1995); McPherson v. Michigan High Sch. Athletic Ass'n., 119 F.3d 453, 462 (6th Cir. 1997); Pottgen v. Missouri State High Sch. Activities Ass'n., 40 F.3d 926, 931 (8th Cir. 1994).

267. *See* Martin v. PGA Tour, Inc., 994 F. Supp. 1242, 1253 (D. Or. 1998) (The majority of the court's analysis focuses on the individual circumstances of Casey Martin's disability impacting the golf world).

268. *See generally* Martin v. PGA Tour, Inc., 994 F. Supp 1242 (D. Or. 1998); Olinger v. United States Golf Ass'n., 55 F. Supp. 2d 926 (N.D. Ind. 1999).

269. *See* 42 U.S.C. § 12182(b)(2)(A)(ii) (1995); *see also* 28 C.F.R. § 35.164 (1999).

270. *See* 28 C.F.R. 35.164 (1999); *see also* Garcia-Ayala v. Lederle Parenterals, Inc., 212 F.3d 638, 649 (1st Cir. 2000) (holding the burden of proof rests with defendant employer to show undue hardship); Johnson v. Gambrinus Co. & Spoetzl Brewery, 116 F.3d 1052, 1065 (5th Cir. 1997) (holding that brewery owner must make accommodation unless it can prove an affirmative defense).

damentally alter" was first introduced in a Supreme Court decision interpreting the Rehabilitation Act of 1973.[271] The concept is now codified at 42 U.S.C. 12182(b)(2)(A)(ii) in the ADA of 1990 and remains applicable to the Rehabilitation Act of 1973 through the Supreme Court's case of *Southeastern Community College v. Davis*.[272] In this case, the Supreme Court decided that an educational institution does not have to change the essence of its program in order to allow a hearing disabled plaintiff access to their nursing program because it would fundamentally alter the program, and the Act cannot be read to require such a change.[273] Since its codification in the ADA of 1990, the fundamental alteration concept has been a vital determining factor for whether or not reasonable accommodations must be made.

The PGA and the USGA both relied on the fundamental alteration concept as their affirmative defense to providing the plaintiffs, Martin and Olinger, their requested accommodations; namely golf carts.[274] The respective trial courts ultimately arrived at conflicting decisions on this issue.[275] The "fundamentally alter" concept requires a factual analysis of the accommodation requested and the impact that it would have on the service, or in this case the competition, that the entity is offering.[276] The problem that arose above in the undue burden analysis of the trial courts is present in this test as well. Should the analysis be based individually or on the system as a whole?

The United States District Court for the District of Oregon used the factual allegations of both the plaintiff, Casey Martin, and the defendant,

---

271. *See* Southeastern Cmty. Coll. v. Davis, 442 U.S. 397, 407-08 (1979).

272. *Id.*

273. *See id.* "Whatever benefits respondent might realize from such a course of study, she would not receive even a rough equivalent of the training a nursing program normally gives. Such a *fundamental alteration* in the nature of a program is far more than the 'modification' the regulation requires." *Id.* at 410 (emphasis added).

274. *See* Martin v. PGA Tour, Inc., 994 F. Supp. 1242, 1249 (D. Or. 1998); *see also* Martin v. PGA Tour, Inc., 204 F.3d 994, 999-1002 (9th Cir. 2000); Olinger v. United States Golf Ass'n, 55 F. Supp. 2d 926, 933-38 (N.D. Ind. 1999); Olinger v. United States Golf Ass'n, 205 F.3d 1001, 1005-07 (7th Cir. 2000).

275. *See Martin*, 994 F. Supp. at 1253 (holding that allowing Casey Martin to ride the PGA fairways during tournaments does not constitute a fundamental alteration of the nature of the game); *see also Olinger*, 55 F. Supp. 2d at 938 (holding that the requested accommodation of traversing the fairways of a USGA competiton would fundamentally alter the nature of the competition).

276. *See* 28 C.F.R. § 35.164; 42 U.S.C. § 12182(b)(2)(A)(ii) (1995). "[Discrimination includes] a failure to make reasonable accommodations . . . unless the entity can *demonstrate* that making such modifications would fundamentally alter the nature of such goods . . . ." *Id.*

PGA, to make an individualized assessment on the question of whether allowing Martin to ride a cart would fundamentally alter the nature of the game.[277] "The type of evidence that satisfies this burden [of the defendant] focuses on the specifics of the plaintiff's or defendant's circumstances and not on the general nature of the accommodation."[278] In the trial court's opinion, the majority of Circuit Court authority leads to an individualized assessment of the requested accommodation that purportedly fundamentally alters the nature of the entities' service.[279]

The PGA's main argument is that walking injects a "fatigue factor" into the game of golf and that by allowing Casey Martin to ride in a cart would allow him an advantage over the other competitors who must walk.[280] The PGA introduces Ken Venturi's 1964 U.S. Open performance, in which he was suffering from a high fever and exhaustion when he finished the course nearing death, to illustrate the effects walking introduces into the game of golf.[281] The court rebuts the contention using testimony from Dr. Gary Klug, a professor in physiology at the University of Oregon who undercuts the PGA's contention by concluding that the fatigue endured on the golf course is largely psychological.[282] The court then goes on to conclude that the majority of golfers, given the chance to ride in a cart, decline and choose to walk.[283] The court then completes its analysis by examining how riding a cart benefits Casey Martin, and had this to say: "If the majority of able-bodied elect to walk in 'carts optional' tournaments, how can anyone perceive that plaintiff has a competitive advantage by using a cart given his condition."[284] The court proceeds to say that Martin's condition provides more fatigue than the average golfer suffers even with the cart and concludes: "As plaintiff easily endures greater fatigue even with a cart

---

277. *See* Martin v. PGA Tour, Inc., 994 F. Supp. 1242, 1249 (D. Or. 1998).

278. *Id.*

279. *See id.* at 1249 n.10.

280. *See id.* at 1250-51.

281. *See id.* at 1250; *see also* Olinger v. United States Golf Ass'n, 205 F.3d 1001, 1006-07 (7th Cir. 2000) (going into great detail of Ken Venturi's experience in the 1964 U.S. Open).

282. *See Martin*, 994 F. Supp. at 1251. The court goes into detail of Dr. Klug's testimony that walking the course burns about 500 calories, which is equivalent to a Big Mac. *See id.* He also testifies that one would be fatigued more quickly if standing on a freeway in the middle of rush hour without moving than if someone claimed fatigue after thirty minutes on a bicycle and then was offered money to continue riding, which would cause the fatigue to dissipate. *Id.* at 1251.

283. *See id.* The court makes this determination from the fact that in the Senior PGA Tour and PGA Qualifying Tournament where carts are permitted, the majority of golfers choose to walk. *Id.* Testimony from Eric Johnson, a Nike Tour Pro, reiterated the same conclusions the court makes. *Id.* at 1251 n.13.

284. *Id.* at 1251.

than his able-bodied competitors do by walking, it does not fundamentally alter the nature of the PGA Tour's game to accommodate him with a cart."[285] Thus, using an individualized inquiry into the "fundamentally alter" concept, the district court decided that the game is not fundamentally altered with cart use in Martin's case, and Martin's case alone.[286]

The United States District Court for the Northern District of Indiana in *Olinger*, using an overview approach to the "fundamentally alter" concept, concluded that to waive the walking rule in anyone's case would fundamentally alter the nature of the service provided by the USGA.[287] The court rejected the contention that it was the game itself that was in danger of being altered; in fact it was the program, namely the U.S. Open.[288]

The *Olinger* court accepted testimony from the USGA's physiological expert, Dr. James Ripp who concluded that "[t]he person who walks the course . . . performs the same amount of physiological work as if he had run eleven-minute miles for 2.5 hours, and studies indicate that such exertion causes cognitive and psychomotor tasks to deteriorate . . . ."[289] and a competitor riding a cart would have an advantage over one who was not.[290] In its opinion, the court admitted that this analysis compared two able-bodied golfers and did not take into account Ford Olinger's disability, which in itself causes fatigue.[291] The court went on to conclude that in evaluating a "requested accommodation's impact on the nature of an activity or program [a court] also must give full consideration to that impact."[292]

Finally, using the broad scope of the "fundamentally alter" concept, the court concluded that to allow Ford Olinger the use of a cart in the U.S. Open would fundamentally alter the nature of the program that the USGA was trying to provide.[293] The court opined that for their decision to require that someone have the discretion to allocate advantages that others are not

---

285. *Id.* at 1252.

286. *See id.* at 1253.

287. *See* Olinger v. United States Golf Ass'n, 55 F. Supp. 2d 926, 934-37 (N.D. Ind. 1999).

288. *See id.* at 934.

289. *Id.*

290. *See id.* This evidence seems to be in direct contradiction to Dr. Gary Klug's testimony in the *Martin* case. *See Martin*, 994 F. Supp. at 1250. In *Olinger*, the court decided to exclude Dr. Klug's testimony because it lacked sufficient grounding in scientific principles. *See Olinger*, 55 F. Supp. 2d at 933 n.4.

291. *See Olinger*, 55 F. Supp. 2d at 937. "Ford Olinger must ride a golf cart to play, and even with a cart he is likely to be more fatigued at the day's end than a healthy Tiger Woods or a healthy David Duvall." *Id.*

292. *Id.*

293. *Id.*

afforded would fundamentally alter the nature of the competition.[294] "The only alternative would be to take the discretion out of the accommodation, and allow all U.S. Open competitors to ride golf carts if they so choose. To do so, however, would be to remove the stamina . . . from the set of qualities designed to be tested in this competition."[295] This concluded the analysis of the court in the *Olinger* case.

4. Qualified Individual with a Disability

There is another aspect which affects the application of the ADA to professional sports, a question which is not currently before the Supreme Court. In Title II of the Act, there is language that requires an individual to be a *qualified individual with a disability* in order to have standing to sue under the Act.[296] The PGA attempts to use *Sandison*[297] and selected portions of the Act, to infer that Casey Martin is not a qualified individual with a disability because he does not meet all the essential eligibility requirements, namely walking, that the sport requires.[298] Under Title II it is required that an individual, without the requested modification, meet all the essential eligibility requirements for the receipt of services.[299]

The PGA argues that the walking rule they impose is an essential eligibility requirement, and therefore Casey Martin, since he cannot meet it, has no standing to sue.[300] The Ninth Circuit briefly dealt with this question by looking into the rules of golf and the number of waivers the PGA has given to the walking rule, and determined it not to be as essential an eligibility requirement as they believed.[301] However, this argument can be quickly vacated for a different reason; the language of the definition of *qualified individual with a disability* itself limits application of that defini-

---

294. *See id.*

295. *Id.*

296. *See* 42 U.S.C. § 12132 (1995). "Subject to the provisions of this subchapter, *no qualified individual with a disability* shall, by reason of such disability, be excluded from participation in . . . or be subjected to discrimination by any such entity." *Id.*

297. *See* Sandison v. Mich. High Sch. Athletic Ass'n, 64 F.3d 1026, 1034 (6th Cir. 1995); *see also supra* note 256 and accompanying text.

298. *See* 42 U.S.C. § 12131(2) (1995); *see also supra* note 259 and accompanying text.

299. *See* Cleveland v. Policy Mgmt. Sys. Corp., 526 U.S. 795, 803 (1999) (noting that "the ADA defines a 'qualified individual' to include a disabled person 'who . . . can perform the essential functions' of her job 'with reasonable accommodation.'" *Id.*; *see also* Olmstead v. Zimring, 527 U.S. 581, 590 (1999); Steele v. Thiokol Corp., 241 F.3d 1248, 1252 (10th Cir. 2001).

300. *See* Martin v. PGA Tour, Inc., 204 F.3d 994, 1000 (9th Cir. 2000).

301. *See id.* at 1000-01.

tion to that specific subsection, namely Title II, dealing with public, not private entities under Title III.[302] This distinction is subtle, but courts have determined that interchangeability between Titles of the ADA is not acceptable.[303] Furthermore, the definition of disability is provided in the section immediately proceeding Title I,[304] which applies to all Titles. Title II is the only Title that supplements that definition with another definition, thus, the limitation seems to be on Title II alone.[305]

V. PGA TOUR, INC. V. CASEY MARTIN: THE DECISION IS MADE

A. Introduction

Shortly after this article was completed, but prior to its publication, the Supreme Court of the United States rendered a decision on the issues around which this article is centered.[306] This section will outline the high court's rendition of the issues and provide an overview of the analysis from the seven-member majority and the published dissent. It will be shown how the Supreme Court obviously ignored the major issues necessary in reaching a decision to expand disability law, as well as the careful critique of these omissions given by the two dissenters.

B. A Flawed Majority Opinion

The Supreme Court outlined the issues as two-fold. First, whether Casey Martin and other disabled golfers[307] fall within the protections of the ADA as it relates to professional golf courses.[308] Secondly, the Court asked whether Casey Martin and other disabled golfers can be summarily denied access to a professional tournament, via the PGA denying use of a cart, by claiming the cart fundamentally alters the nature of the tournaments, with respect to the other golfers who are required to walk.[309]

---

302. *See* 42 U.S.C. §§ 12131-12132 (1995).

303. *See* Bloom v. Bexar County, Texas, 130 F.3d 722, 726 (5th Cir. 1997).

304. 42 U.S.C. § 12102(2)(a) (1995). "The term 'disability' means, with respect to an individual (A) a physical or mental impairment that substantially limits one or more of the major life activities of such individual . . ." *Id.*

305. *See* 42 U.S.C. § 12131(2) (1995).

306. *See* PGA Tour, Inc., v. Martin, 121 S. Ct. 1879 (2001).

307. There is no dispute, even at the Supreme Court level, that Casey Martin is disabled within the meaning of the ADA. *See id.* at 1885.

308. *See id.* at 1884. This portion of the opinion deals with the question of whether the PGA Tour and their golf courses are public accommodations within the meaning of the ADA, and if so, how extensive that coverage is. *See id.*

309. *See id.* In answering this question, the Court will decide whether the use of a golf cart fundamentally alters the nature of the competition so as to render the request for a cart unreasonable for purposes of the ADA, and if consequently the

i. The Facts as Outlined by the Supreme Court: Entry to the Tour

Dispensing with the preliminary facts regarding the PGA Tours non-profit status, the Court begins its discussion with annual income generated from television, admissions, concessions, and contributions. This amount, mostly dispersed as prize money, totals nearly 300 million dollars.[310]

Justice Stevens focuses heavily on two major areas of fact as they relate to Casey Martin's case: First, the ease in which a member of the public can enter the professional world of golf by earning his tour card, and second, the "Rules of Golf" as they are presently interpreted by the court.[311] The first of these factual areas is described much the same way as it was in the Seventh[312] and Ninth[313] Circuits, and will not be repeated here. However, it will be discussed in the analysis portion of this opinion.[314] The second factual basis relied on by Justice Stevens, also previously discussed, will be briefly reiterated because of the heavy reliance upon the Rules of Golf as groundwork for the "public accommodation" portion of his opinion.[315]

The "Rules of Golf," as outlined by the Court in this opinion, are broken into three separate sets of regulations.[316] First, are the "Rules of Golf" authored by the United States Golf Association (USGA) and its Scottish equivalent, the Royal and Ancient Golf Club of Scotland.[317] The "Rules of Golf" govern amateur play across the country as well as all major organizations who hold tournaments for professional golfers.[318] "Those rules do not prohibit the use of golf carts at any time."[319]

Second, the "'Conditions of Competitions and Local Rules,' often described as the 'hard card,'"[320] which are directly applicable to the PGA's Tours.[321] The Nike and PGA Tour hard cards provide that golfers should walk at all times and prohibits the use of golf carts during tournament play,

---

ADA is applicable.

310. *See id.*

311. *See PGA Tour, Inc.*, 121 S. Ct. at 1884-85.

312. *See Olinger*, 205 F.3d at 1002-03.

313. *See Martin*, 204 F.3d at 996.

314. *See infra* notes 325-355 and accompanying text.

315. *See* PGA Tour, Inc., v. Martin, 121 S. Ct. 1879, 1893-96 (2001).

316. *See id.* at 1884.

317. *See id.*

318. *See id.* at 1884. The PGA Tour, USGA, the Ladies PGA (LPGA), and the Senior Women's Golf Association all use this basic set of rules. *See id.*

319. PGA Tour, Inc. v. Martin, 121 S.Ct. 1879, 1884-85 (2001).

320. *Id.* at 1885.

321. *See id.*

but allows the use of carts during open qualifying rounds.[322] Finally, are the 'Notices to Competitors' which govern rules for a given event, not to be applicable to any other tournament.[323] These 'event specific' rules may or may not authorize use of carts.[324] It is the use of these three specific set of rules governing tournament level play that Justice Stevens uses to rebut the PGA's contention that it does not fall within the gambit of the ADA.

ii. The Analysis: Oops! Did We Miss Something?

The Supreme Court, through Justice Stevens opens two wide entrances for professional sports to escape the auspices of Title III. First, the heavily weighted discussion focusing on the Rules of Golf from Q-School to Professional Tournament play suggests that an alteration to the Rules and practices of professional sports, specifically professional golf, may have a direct impact on whether or not denying access to disabled individuals will violate the act. Second, and more noticeably abused, is the very clear language of Title III of the ADA which limits the application of 42 U.S.C. § 12181(7)(L), which contains "or other places of exercise or recreation" at the end of the section, thus implying that professional sports may not be covered. [325] The latter issue will be dealt with first since it has been the theme of this paper to address the public accommodation aspect before dealing with whether the accommodation requested fundamentally alters the aspect of the program.

A. Public Accommodation

Whether or not the PGA, the USGA, or any other professional sports organization is subject to the provisions of the ADA depends on the language contained within the statute.[326] Justice Stevens is very happy to include golf courses, as used during a PGA sponsored tournament, within the

322. *See id.* "The PGA Tour hard card provides: 'Players shall walk at all times during a stipulated round unless permitted to ride by the PGA Tour Rule Committee. The Nike Tour hard card similarly requires walking unless otherwise permitted. Additionally, as noted, golf carts have not been permitted during the third stage of the Q-School since 1997. Petitioner added this recent prohibition in order to approximate a PGA Tour event as closely as possible.'" *Id.*

323. *See id.* These rules inform players on how a specific hazard, man-made or natural, should be dealt with in lieu of the Rules of Golf. *Id.*

324. *See* PGA Tour, Inc. v. Martin, 121 S.Ct. 1879, 1885 (2001). For example, carts may be authorized if there are unusually long distances between holes and they are needed to speed up play. *Id.*

325. 42 U.S.C. § 12181(7)(L) (1995).

326. *See supra* notes 164-196 and accompanying text.

provisions of 42 U.S.C. section 12181(7)(L).[327] In his opinion, it is clear and unambiguous that golf courses are covered by the plain and literal words contained in this provision:[328] "The events occur on 'golf course[s],' a type of place specifically identified by the Act as a public accommodation."[329] This represents the sum total of Stevens first justification calling for Title III applicability to the PGA. Stevens offers a second reason for coverage upon the PGA:

> In addition, at all relevant times, petitioner 'leases' and 'operates' golf courses to conduct its Q-School and tours. As a lessor and operator of golf courses, then, petitioner must not discriminate against any 'individual' in the 'full and equal enjoyment of the goods, services, facilities, privileges, advantages, or accommodations' of those courses.[330]

Stevens uses this language taken from 42 U.S.C. § 12182(a) in conjunction with the PGA allowing any citizen to pay and tryout for a PGA tour card,[331] to justify coverage of the golf courses the PGA leases and operates.[332] Apparently, according to Stevens' opinion, there is a direct correlation to coverage within the provisions of the ADA and how open to the public the service actually is.[333] Stevens concludes this part by saying "[i]t would therefore appear that Title III of the ADA, by its plain terms prohibits petitioner from denying Martin equal access to its tours on the basis of his disability."[334] At no point in this explanation does Stevens successfully argue away the apparently plain and unambiguous limiting language of the statute, "or other places of exercise or recreation."[335]

Lastly, the Court dispenses quickly with the PGA's contention that the

---

327. "[A] gymnasium, health spa, bowling alley, golf course, or other place of exercise or recreation." 42 U.S.C. § 12181(7)(L) (1995).

328. *See* PGA Tour, Inc., v. Martin, 121 S.Ct. 1879, 1890 (2001). Stevens dispenses with this provision in a single paragraph. *See id.*

329. *Id.* at 1890; *see also* 42 U.S.C. § 12181(7)(L) (1995).

330. *PGA Tour, Inc.*, 121 S.Ct. at 1890 (citations omitted). Stevens concludes that the nature of citizens paying to qualify as discussed previously is a privilege covered by the Act. *See id.* What if the PGA were to change its qualifying rules so as to significantly limit the qualifying, so that ordinary individuals may not pay to qualify. *See infra* notes 362-64 and accompanying text.

331. *See supra* notes 104-07 and accompanying text.

332. *See PGA Tour, Inc.*, 121 S.Ct. at 1890.

333. *See id.* Stevens discusses that thousands of individuals from the general public can participate in this privilege offered by the PGA and that Casey Martin is without a doubt one of these individuals. *Id.*

334. *Id.*

335. *See generally PGA Tour, Inc.*, 121 S.Ct. at 1889-93 (the limiting language is found directly after 'golf course'). *See* 42 U.S.C. § 12181(7)(L) (1995).

ADA was intended to deal with "clients or customers,"[336] and that Martin is certainly not a client or a customer.[337] Justice Stevens quickly decides that even if the PGA has correctly deciphered the statute, that Casey Martin cannot be anything else but a client or a customer.[338]

> In our view, petitioners tournaments (whether situated at a 'golf course' or at a 'place of exhibition or entertainment') simultaneously offer at least two 'privileges' to the public—that of watching the golf competition and that of competing in it. Although the latter is more difficult and more expensive to obtain than the former, it is nonetheless a privilege that petitioner makes available to members of the general public.[339]

Again, Justice Stevens stonewalls another PGA viewpoint by pointing to the open nature of the tournaments and competitions the PGA conducts. Following this logic the Court notes that this arrangement is unlike other professional sports where players are under employment contracts with their respective teams, suggesting that perhaps if the PGA changes the nature of their agreements with professional golfers the outcome of this case might be different.[340]

Justice Stevens' entire argument for the majority on the public accommodation question rests upon the setup and organization of the PGA Q-School and its relative wide-open nature of gaining access to professional tournaments.[341] This opinion does nothing for the advancement of disabled access to professional sports. In fact, it clears up gray areas which now allow sports organizations to correctly set up or change their existing setup, so as not to fall within the provisions of the ADA. What has been done here, to find a way to allow Casey Martin to continue riding his cart in PGA sponsored events, is a victory to Martin, but a tremendous defeat to all other disabled athletes countrywide.

B. Reasonable Accommodation, According to the Majority

After summarily discarding the PGA's contention that the ADA does not

---

336. *PGA Tour, Inc.*, 121 S. Ct. at 1891. This language is derived from 42 U.S.C. § 12182(b)(1)(A)(iv), which states "for purposes of clause (i) through (iii) of this subparagraph, the term individuals or class of individuals refers to the clients or customers of the covered public accommodation . . . ." 42 U.S.C. § 12182(b)(1)(A)(iv) (1994).

337. *See PGA Tour, Inc.*, 121 S. Ct at 1891.

338. *See id.* at 1891-92.

339. *Id.* at 1892.

340. *See id.* at 1892 n.33.

341. *See generally* PGA Tour, Inc. v. Martin, 121 S. Ct. 1879 (2001).

apply to their tournaments, the Court decides, what they believe is the only other major area of controversy: whether allowing Casey Martin to ride in a golf cart during PGA sponsored tournaments fundamentally alters the nature of the competition so as to render the request unreasonable.[342]

In deciding this issue Justice Stevens again relies on the nature of the PGA itself, specifically the Rules of Golf and the traditional method of play, again suggesting that an alteration at the fundamental level of golfing rules would shield the PGA from the ADA.[343] The majority opinion concludes, by way of the "Rules of Golf" that the walking rule is not indispensable to the game of golf.[344]

> There is nothing in the Rules of Golf that either forbids the use of carts, or penalizes a player for using a cart. That set of rules, as we have observed, is widely accepted in both the amateur and professional golf world as the rules of the game. The walking rule that is contained in petitioner's hard cards, based on an optional condition buried in an appendix to the Rules of Golf, is not an essential attribute of the game itself.[345]

Once again, the majority uses the fact that in the first two rounds of Q-School, competitors have the option of using a golf cart to compete.[346] This, along with the evidence that the Senior PGA Tour, and on occasion other PGA and NIKE Tour events authorize the use of carts for varying reasons was enough to sway the court in Martin's favor.[347]

The PGA contends that the walking rule injects the element of fatigue into the game of shot-making and a modification for Casey Martin would tip the balance in his favor, making the playing field unequal.[348] The Court makes quick work of this argument by explaining how the game of golf is one "in which it is impossible to guarantee that all competitors will play under exactly the same conditions or that an individual's ability will be the sole determinant of the outcome."[349] This statement, coupled with testimony from an expert on fatigue, stating that walking a golf course burns

---

342. *See PGA Tour, Inc*, 121 S.Ct. at 1893.

343. *See id.* at 1894. Justice Stevens tracks the history of the game of golf and notes that "[g]olf carts started appearing with increasing regularity on American golf courses in the 1950's. Today they are everywhere. And they are encouraged." *Id.* What Justice Stevens fails to mention is that these alterations are for laymen engaged in the game of golf, not professionals.

344. *See id.* at 1895.

345. *Id.* at 1894-95.

346. *Id.* at 1895.

347. *See* PGA Tour, Inc. v. Martin, 121 S. Ct. 1879, 1895 (2001).

348. *See id.*

349. *Id.*

less calories (500) than the nutritional equivalent of a Big Mac surely cannot be outcome effective.[350]

Accepting the PGA's argument that the walking rule is outcome affective, the majority shuffles onto another point. Even if the walking rule is outcome affective, and dispensing with it would fundamentally alter the nature of the game, the PGA needed to assess Martin's individual circumstances in deciding whether to grant his accommodation.[351] This is a powerful point that the majority opinion lays out. However, the Court only speaks of the applicable legal standard. No factual application is given to aid in future determinations. It is not clear from the Court's plain words what the PGA did, or failed to do, in failing to comply with this requirement.[352] How is one supposed to know what action does or does not constitute compliance? In sum, the Court uses poor factual evidence to determine whether a fundamental alteration of the game of golf would be present if Casey Martin were to ride a cart.

The majority waives off the last point made by the PGA as easily as the rest. The PGA argues that the substantive rules for its highest level tournaments are not within the coverage of Title III and may not be altered, for fear of an uneven playing ground.[353] Here, the Court says "that provision [of Title III] carves out no exception for elite athletics, and given Title III's coverage not only of places of 'exhibition or entertainment' but also of 'golf courses' its application to petitioner's tournaments cannot be said to be unintended or unexpected."[354] Again, as the dissent points out, the majority has mis-construed the statute into covering areas of competitive athletics in a way in which it was never intended to do.[355]

iii. Dissent: Scalia Painfully on Point Again

Justice Scalia authors a crippling dissenting opinion, which employs a logical construction of Title III's limiting language. Scalia criticizes the Stevens interpretation of 42 U.S.C. § 12181 (7), which opens the twelve categories of public accommodations covered to any use of those facilities for any reason.

---

350. *See id.* at 1896.

351. *See id.*

352. *See* PGA Tour, Inc. v. Martin, 121 S.Ct. 1879, 1896 (2001). Justice Stevens throws this out as an "even if" scenario. However, if the "even if" scenario is flawed when accepting all else as in petitioner's favor, then does it not follow that petitioner wins the battle?

353. *See id.*

354. *Id.* at 1896.

355. *See infra* note 359 and accompanying text.

Section 12181(7) lists twelve specific types of entities that qualify as "public accommodations," with a follow-on expansion that makes it clear what the "enjoyment of the goods, services, etc." of those entities consist of—and it plainly envisions that the person "enjoying" the "public accommodation" will be a customer. For example, Title III is said to cover an "auditorium" or "other place of public gathering,".[356] Thus, 'gathering' is the distinctive enjoyment derived from an auditorium; the persons 'gathering' at an auditorium are presumably covered by Title III, but those contracting to clean the auditorium are not. Title III is said to cover a "zoo" or "other place of recreation,".[357] The persons 'recreating' at the 'zoo' are presumably covered, but the animal handlers bringing in the latest panda are not.[358]

The logic employed by Scalia is simple and straightforward and a conclusion the majority fails to discuss in rendering its opinion. Scalia opines that the ADA, Title III was only intended to cover "clients or customers" and not everyone else, as the majority points out.[359] Scalia goes on to buttress his comment on the limiting language of the ADA by quoting the Department of Justices's Code of Federal Regulations, which further indicates the necessary conclusion Scalia reaches.[360]

The dissent does not consider a professional golfer a 'client or customer' as the majority so readily concludes. Scalia compares professional golfers, such as Casey Martin, to baseball players at Yankee Stadium in which the principle objective of both is to make money not to 'enjoy' goods or services.[361] Scalia concludes that based on the plain language of the statute and

---

356. 42 U.S.C. § 12181(7)(L) (1995).

357. 42 U.S.C. § 12181(7)(I) (1995).

358. PGA Tour, Inc. v. Martin, 121 S.Ct. 1879, 1899 (2001) (Scalia, J., dissenting).

359. *See id.* The majority and dissent are in agreement here, but only as far as this is an alternative argument for the majority, an "even if" scenario, where the majority concludes Casey Martin would still fit in this category because he is a 'client or customer' that is protected. *See supra* notes 336-40 and accompanying text.

360. *See id.* at 1899. "And finally, the regulations promulgated by the Department of Justice reinforce the conclusion that Title III's protections extend only to customers. 'The purpose of the ADA's public accommodations requirements,' they say, 'is to ensure accessibility to the goods offered by a public accommodation.'" *Id.* (quoting 28 C.F.R. ch. 1, pt. 36, App B, p.650 (2000)).

361. *See id.* at 1900. "To be sure, professional ballplayers participate in the games, and use the ballfields, but no one in his right mind would think that they are customers of the American League or of Yankee Stadium." *Id.* The problem with Scalia's argument here, although seemingly correct in outcome, is that players at Yankee Stadium are under *employment* contracts, more suitably dealt with under Title I of the ADA or more accurately, within the individual contract itself.

common sense that golf courses are not covered within Title III as used by the PGA in highly skilled competition.[362]

Scalia moves on to the second aspect, which he feels the court ignores completely, that is the question of modification and fundamental alteration. Scalia points out that Title III and the Code of Federal Regulations that coincides, specifically does not require the entity to alter its inventory or provide specialty goods solely for the benefit of the disabled individual.[363] Therefore, the question of whether or not walking is essential or if it is a reasonable modification should never be reached.[364] "The PGA TOUR cannot deny respondent access to that game because of his disability, but it need not provide him with a game different (whether in its essentials or in its details) from that offered to everyone else."[365] Even believing that the analysis should stop here, Scalia moves further along the majority opinion.

After brushing over the reasonable modification question the court moves forward to tackle the question of whether the modification fundamentally alters the program.

> The Court attacks this fundamental alteration analysis by asking two questions: first, whether the 'essence' or an 'essential aspect' of the sport of golf has been altered; and second, whether the change, even if not essential to the game, would give the disabled player an advantage over others and thereby 'fundamentally alter the character of the competition.'[366]

Scalia points out to the reader that the first question raised by the majority is based on false assumptions. "Nowhere is it written that PGA TOUR golf must be classic 'essential' golf. Why cannot the PGA TOUR, if it wishes, promote a new game, with distinctive rules . . . ?"[367]

The second question is left unanswered by Scalia because he believes that the majority has made it a determinative question of individualized standards completely arbitrary.[368] In his conclusion, Scalia announces the lesson to be learned from the majority opinion:

The lesson the PGA TOUR and other sports organizations should take from this case is to make sure that the same written rules are set forth for

---

362. *See id.*

363. *See PGA Tour Inc., v. Martin*, 121 S. Ct. at 1901.

364. *See id.*

365. *Id.* at 1902.

366. *Id.*

367. *Id.* "If members of the public do not like the new rules – if they feel that these rules do not truly test the individual's skill at 'real golf' . . . they can withdraw their patronage. But the rules are the rules." *Id.*

368. *See id.* at 1904.

all levels of play, and never voluntarily to grant any modifications. The second lesson is to end open tryouts. I doubt that, in the long run, even disabled athletes will be well served by the incentives that the Court has created.[369]

VI. CONCLUSION

The hopes and dreams of Casey Martin, Ford Olinger and all professional athletes that are challenged through their respective disabilities should be realized through the ADA. The expansion of rights that the Supreme Court has initiated through congressional legislation and the nature in which they have interpreted that legislation leads to the conclusion that it is simply a matter of time before disabled individuals enjoy the same access that minorities have struggled for centuries to realize. The expansion of disability legislation has become more encompassing with each congressional turn in its efforts to vindicate the rights of disabled individuals in the United States.[370] Sports organizations should be corralled into the ADA with the intended sweeping language that it employs.[371]

The affirmative defenses that the ADA affords, namely proving that providing the requested accommodation would be an undue financial or administrative burden or that the requested accommodation would fundamentally alter the nature of the service, should be a case-by-case, fact specific inquiry.[372] It is viable that Casey Martin and Ford Olinger will have differing results in their two respective cases. [373] The differentiation between the presentation of the evidence and the weight afforded it by the

---

369. *PGA Tour, Inc., v. Martin*, 121 S. Ct. at 1904. Scalia's final paragraph is worth reproducing here:

> Complaints about this case are not 'properly directed to Congress.' They are properly directed to this Court's Kafkaesque determination that professional sports organizations, and the fields they rent for their exhibitions, are "places of public accommodation" to the competing athletes, and the athletes themselves "customers" of the organization that pays them; its Alice in Wonderland determination that there are such things as judicially determinable "essential" and "nonessential" rules of a made-up game; and its Animal Farm determination that fairness and the ADA mean that everyone gets to play by individualized rules which will assure that no one's lack of ability (or at least no one's lack of ability so pronounced that it amounts to a disability) will be a handicap. The year was 2001, and "everybody was finally equal."

Id. at 1905.

370. *See supra* notes 46-83 and accompanying text.

371. *See supra* notes 159-234 and accompanying text.

372. *See supra* notes 269-295 and accompanying text.

373. *See supra* notes 84-158 and accompanying text.

fact-finder, is exactly what our system of justice is all about.

It is a matter of time before the expansion of disabled individual's rights are realized in the broad scope that was intended by Congress in implementing the ADA. The Supreme Court has chosen the opportunity to make it sooner rather than later even though the impact is minimal.

Damien Bielli

NOTES

The Evolutionary War on First Amendment Rights and Abortion Clinic Demonstration*

I. INTRODUCTION

A little over two-hundred years ago, the United States Constitution recognized a person's right to free speech in the First Amendment of the Bill of Rights.[1] Yet it was only twenty-seven years ago that the United States Supreme Court formally recognized a woman's constitutional right to an abortion.[2] In the twenty-seven years that have since passed, conflicts have developed between these two fundamental rights which have been the source of much controversy, debate, violence and even murder.[3] "There was a time in this country when the danger from having an abortion was caused by the quality of the procedure itself . . . . [A]s opposition to legalized abortion increased in the 1990s, those seeking an abortion faced a new danger – being injured or killed in a visit to an abortion clinic."[4] Since the more recent decisions regarding abortion rights have reinforced the fundamental right to abortion,[5] the topics for debate have shifted away from

---

* NEW ENGLAND LAW REVIEW Scribes Award Contribution.

1. *See* U.S. CONST. amend. I. "Congress shall make no law . . . abridging the freedom of speech . . . or the right of the people peacably to assemble, and to petition the Government for a redress of grievances." *Id.*

2. *See* Roe v. Wade, 410 U.S. 113 (1973) (holding a Texas law prohibiting abortion unconstitutional as violative of the Fourteenth Amendment right to due process and as an infringement on a mother's right to privacy).

3. Since 1993, at least seven abortion clinic employees have been murdered in the United States, while threats and attempts have been made on many more. *See* Note, *Safety Valve Closed: The Removal of Nonviolent Outlets for Dissent and the Onset of Anti-Abortion Violence*, 113 HARV. L. REV. 1210, 1221 (2000).

4. David G. Pettinari, Hill v. Colorado – *The United States Supreme Court Squares Off with Colorado over the First Amendment Rights of Abortion Protesters*, 77 U. DET. MERCY L. REV. 803, 803 (2000) (discussing the *Hill v. Colorado* abortion clinic buffer zone case as it relates to prior abortion rights case law).

5. *See* Planned Parenthood of Southeastern Pa. v. Casey, 505 U.S. 833 (1992) (upholding the decision in *Roe v. Wade*, reaffirming the fundamental right to abortion, while also allowing for increased restrictions by the states as long as they do not im-

the *Roe*-given "right to choose"[6] toward the right to educate others about the "right to choose" through demonstration and protest outside medical facilities.[7] The United States Supreme Court has recognized that "[t]he First Amendment does not demand that patients at a medical facility undertake Herculean efforts to escape the cacophony of political protests."[8] This concept has been developing and evolving since the *Roe v. Wade* decision; the courts and legislatures taking a more expansive position with each new case and circumstance.[9] In addition to advancements in the medical and procedural aspects of abortion rights inside the clinics, there have also been similar developments regarding demonstration rights occurring outside of these clinics, the constitutionality of which continue to be under scrutiny, but have acquired some acceptance within the First Amendment arena.[10]

This Note will explore the balancing that has transpired, in both the state

---

pose an undue burden on the right to an abortion). *See also* Stenberg v. Carhart, 530 U.S. 914 (2000) (holding a Nebraska state statute criminalizing the performance of partial birth abortions unconstitutional due to its lack of exception for the preservation of the health of the mother and imposing an "undue burden upon a woman's right to make an abortion decision"). *Id.* at 946.

6. *See* Roe v. Wade, 410 U.S. at 122. The Court discussed the "right to choose" in considering

> [t]he principal thrust of appellant's attack on the Texas statutes is that they improperly invade a right, said to be possessed by the pregnant woman, to choose to terminate her pregnancy. Appellant would discover this right in the concept of personal "liberty" embodied in the Fourteenth Amendment's Due Process Clause; or in personal, marital, familial, and sexual privacy said to be protected by the Bill of Rights or its penumbras[.]

Id. at 129.

7. *See infra* note 10 and accompanying text.

8. Madsen v. Women's Health Center, Inc., 512 U.S. 753, 772-73 (1994) (holding that a thirty-six foot buffer zone around an abortion clinic entrance and driveway and some related noise restrictions outside the clinic were constitutional as they "burden[ed] no more speech than necessary to eliminate the unlawful conduct targeted by the state court's injunction," while other provisions were struck down as too burdensome). *Id.* at 776.

9. Since *Roe v. Wade*, courts and state legislatures have decided cases and enacted legislation related to the topics of abortion procedures, governmental funding, requirements for facilities and personnel performing the procedures, and "gag rules" on disbursal of abortion - patient information. *See Insight Roe vs. Wade*, THE COLUMBUS DISPATCH, Jan. 18, 1998, at 4C.

10. *See id.* "Since 1992, the fight over abortion restrictions has shifted from the high court to statehouses and Congress. Most of the court's abortion-related cases have dealt with protests at abortion clinics and the freedom-of-speech rights of demonstrators, not the fundamental right to abortion." *Id.*

and federal court systems and legislatures, between First Amendment freedom of speech rights and abortion-related privacy rights concerning buffer zones around abortion clinics.[11] The scales are tipping in opposite directions depending on how burdensome the provisions are found to be on speech in the various jurisdictions.[12] Massachusetts enacted its own buffer zone law in November 2000, the constitutionality of which is still under debate in the courts.[13] As for the probable outcome in Massachusetts, the forthcoming exploration will show that the United States Supreme Court will uphold the constitutionality of the Massachusetts buffer zone statute and order that its enforcement be upheld indefinitely.[14]

The remainder of this Note will analyze the prior legal hypothesis through the following subsections to ultimately prove its truth. An analysis of the First Amendment to the Constitution will be conducted in Part II.A.[15] Next, Part II.B. will discuss the evolution of abortion jurisprudence through three principal cases: Part II.B.1 will examine *Roe v. Wade*,[16] Part II.B.2 will discuss *Planned Parenthood of Southeastern Pa. v. Casey*,[17] and finally, Part II.B.3 will examine *Stenberg v. Carhart*.[18] Part II.B.4 discusses the newest alternative to surgical abortion procedures, RU-486.[19] In Part II.C., the fundamental cases joining First Amendment constitutionality and abortion clinic protest case law will be examined: Part II.C.1 will *discuss Madsen v. Women's Health Center, Inc.*,[20] Part II.C.2 will examine *Schenck v. Pro-Choice Network of Western New York*,[21] and Part II.C.3 will discuss *Hill v. Colorado*.[22] Part III will examine the facets of the new

---

11. *See Madsen*, 512 U.S. at 765. In *Madsen*, the Court found the standard to be "whether the challenged provisions of the injunction burden no more speech than necessary to serve a significant government interest," thus a balancing between means and ends. *Id.*

12. *See infra* notes 85-258 and accompanying text.

13. *See infra* notes 259-300 and accompanying text.

14. For an analysis comparing the Massachusetts law to a similar law enacted in Colorado, which has been upheld as constitutional by the United States Supreme Court, see *infra* notes 301-52 and accompanying text.

15. *See* U.S. CONST. amend. I. *See also infra* notes 28-48 and accompanying text.

16. 410 U.S. 113 (1973). *See infra* notes 49-54 and accompanying text.

17. 505 U.S. 833 (1992). *See infra* notes 55-66 and accompanying text.

18. 530 U.S. 914 (2000). *See infra* notes 67-75 and accompanying text.

19. *See infra* notes 76-83 and accompanying text.

20. 512 U.S. 753 (1994). *See infra* notes 85-146 and accompanying text.

21. 519 U.S. 357 (1997) (holding the fixed buffer zone and cease and desist provisions of the injunction to be constitutional, but striking down the floating buffer zone provision as a violation of First Amendment free speech rights). *See infra* notes 147-83 and accompanying text.

22. 530 U.S. 703 (2000) (holding that floating buffer zone established through

Massachusetts law: Part III.A. will discuss the factual history of the law,[23] Part III.B. will further discuss the procedural history of the law,[24] and Part III.C.1-5 will analyze the outcome in *Hill v. Colorado* as applied to the Massachusetts law.[25] Section IV will draw a conclusion as to the favorable outcome of the pending litigation and the undoubted future of buffer zones within the Commonwealth of Massachusetts.[26] But the basic inquiry in all of the cases that have come before the courts is: How much buffer is too much, or not enough?[27]

II. BACKGROUND

A. First Amendment to the Constitution of the United States

To answer the question of how much buffer is too much, one must understand that the First Amendment right is not absolute.[28] The government has the right to restrict some speech when specific criteria are met.[29] First

---

a Colorado state statute was constitutional). *See infra* notes 186-258 and accompanying text.

23. *See infra* notes 259-84 and accompanying text.

24. *See infra* notes 285-300 and accompanying text.

25. *See* MASS. GEN. LAWS ch. 266, § 120E½ (2000). *See also infra* notes 301-52 and accompanying text.

26. *See infra* notes 353-57 and accompanying text.

27. What "buffer" means, both figuratively and literally, is "a clear, easily enforced zone of protection" surrounding the medical facilities performing abortion service, including doorways, driveways, public sidewalks, sometimes even the street itself depending upon the circumstances and the jurisdiction, motor vehicles carrying patients, and the actual corporeal beings of the patients themselves. Edwards v. City of Santa Barbara, 150 F.3d 1213, 1216 (9th Cir. 1998) (affirming the injunction enjoining enforcement of the floating buffer zone provision of the ordinance, but upholding the fixed driveway provisions as "a reasonable time, place and manner restriction"). *Id.* at 1215. *See also* Schenck v. Pro-Choice Network of Western New York, 519 U.S. 357 (1997); Madsen v. Women's Health Center, Inc., 512 U.S. 753 (1994). Courts have been more generous in some states than in others as to the amount of buffer deemed constitutional. *See* Ronnie Beth Nadell, Comment, Planned Parenthood Shasta-Diablo, Inc. v. Williams*: Anti-Abortion Protestors, Injunctions, and the Suffocation of Free Speech in California*, 28 SW. U. L. REV. 627, 642 (1999) (discussing differences between the decision in *Madsen*, allowing a thirty-six foot buffer zone, and *Shasta-Diablo II* where the court allowed a buffer zone that stretched from the clinic entrance in the rear of the building to the sidewalk in the front, and further placed the protestors across the street, over a busy four lane highway – totaling a 172-foot buffer zone).

28. *See* Elrod v. Burns, 427 U.S. 347, 360 (1976). "[T]he prohibition on encroachment of First Amendment protections is not an absolute. Restraints are permitted for appropriate reasons." *Id.*

29. *See* Erwin Chemerinsky, *The Fifty-Fifth Cleveland-Marshall Fund Lecture: The First Amendment: When the Government Must Make Content-Based Choices*, 42

Amendment restrictions related to abortion law jurisprudence since the *Roe v. Wade* decision have evolved with the times.[30] As liberties and restrictions ebbed and flowed with respect to a woman's right to choose, restrictions on how protestors can express their views, both affirmatively and negatively, regarding those sensitive issues have also ebbed and flowed.[31]

The basic principle involved in all of the case law and legislation regarding buffer zones outside abortion clinics derives from the fundamental rights of the First Amendment to the Constitution of the United States.[32] The text of the Constitution provides, "Congress shall make no law . . . abridging the freedom of speech . . . or the right of the people peacably to assemble, and to petition the Government for a redress of grievances." [33] "This language has been interpreted to protect 'expressive conduct' or 'symbolic speech,' as well as more traditional means of communicating."[34] Reading the text literally, one would think that no law would be allowed to infringe upon the rights of someone to express ideas through speech.[35]

However, this right is not absolute.[36] "The level of First Amendment protection accorded to speech and other communicative activity may also vary based on the type of speech as well as the type of harm inflicted by the speech."[37] There are some situations where speech or conduct should not have the same level of protection as other situations, and it is up to the government to regulate those cases.[38] "[T]he Supreme Court never has accepted the view that the First Amendment prohibits all government regulation of expression."[39] Limitations by the state or federal government may be appropriate when conflict exists between the right to free speech

---

CLEV. ST. L. REV. 199, 211-13 (1994).

30. *See supra* notes 9-10 and accompanying text.

31. *See infra* notes 85-258 and accompanying text.

32. As the Court in *Hill v. Colorado* stated, "[t]he question is whether the First Amendment rights of the speaker are abridged by the protection the statute provides for the unwilling listener." *Hill*, 530 U.S. at 708. *See also* Schenck v. Pro-Choice Network of Western New York, 519 U.S. 357 (1997); Madsen v. Women's Health Center, Inc., 512 U.S. 753 (1994); Edwards v. City of Santa Barbara, 150 F.3d 1213 (9th Cir. 1998).

33. U.S. CONST. amend. I.

34. Laurence J. Eisenstein & Steven Semeraro, *Abortion Clinic Protest and the First Amendment*, 13 ST. LOUIS U. PUB. L. REV. 221, 226 (1993) (discussing the various levels of protection of First Amendment freedom of speech).

35. *See supra* note 33 and accompanying text.

36. *See* Eisenstein & Semeraro, *supra* note 34, at 227.

37. *Id.*

38. *See id.* "For example, the First Amendment does not protect the right to physically abuse or tortiously harass another." *Id.*

39. ERWIN CHEMERINSKY, CONSTITUTIONAL LAW PRINCIPLES AND POLICIES 750 (1997).

and the contrary rights of others.[40] The balancing act by the courts is an attempt to "accept as 'common ground' that the 'First Amendment does not guarantee the right to communicate one's views at all times and places or in any manner that may be desired.'"[41] When there are legitimate governmental interests to be served, such as the safety and welfare of its citizens, the government has both the right and duty to take action to preserve those interests.[42]

However, the government must use caution in making its decisions.[43] The "traditional bedrock rule of the First Amendment [is]: The government cannot regulate speech based on its content."[44] Just because the government does not agree with what is being said is not an invitation for its regulation; "the government must be both viewpoint neutral and subject matter neutral."[45] "For example, it would be clearly unconstitutional for the government to say that pro-choice demonstrations are allowed in the park but anti-abortion demonstrations are not allowed . . . . Such viewpoint regulation is not allowed."[46] However, there are other situations where the form of expression is so demoralizing or harmful to society, such as obscenity or incitement of violence, that viewpoint and content are irrelevant; these acts fall outside the scope of the First Amendment and receive little constitutional protection, if any at all.[47] Hence, in many different settings,

---

40. *See* Elrod v. Burns, 427 U.S. 347, 363 (1976) (concluding that for the restrictions on public employment to "survive constitutional challenge, [they] must further some vital government end by a means that is least restrictive of freedom of belief and association in achieving that end, and the benefit gained must outweigh the loss of constitutionally protected rights").

41. Tara K. Kelly, *Silencing the Lambs: Restricting the First Amendment Rights of Abortion Clinic Protestors in* Madsen v. Women's Health Center, 68 S. CAL. L. REV. 427, 451 (1995) (quoting Heffron v. Int'l Soc'y of Krishna Consciousness, 452 U.S. 640, 647 (1981)).

42. *See* Amy E. Miller, Note, *The Collapse and Fall of Floating Buffer Zones: The Court Clarifies Analysis for Reviewing Speech-Restrictive Injunctions in* Schenck v. Pro-Choice Network, 32 U. RICH. L. REV. 275, 277 (1998).

43. *See* Chemerinsky, *supra* note 29, at 211. The article discusses how

> First Amendment law is largely focused on limiting the government's ability to regulate speech . . . . Ultimately, the issue is about how to control the government in a situation where the government must act . . . . Substantively, the government must make choices; but procedurally the courts should insist that they are done in the fairest way possible.

Id.

44. *Id.* at 201.

45. *Id.* at 202-03.

46. *Id.* at 203.

47. *See* Miller, *supra* note 42, at 276-77 (discussing the various types of speech that are or are not afforded protection by the First Amendment to the Constitution).

the United States Supreme Court has upheld laws that purport to limit the degree of freedom of speech that can be exercised.[48] The following sections will delineate the constitutional developments since *Roe v. Wade* whose implications on First Amendment rights are based on abortion and privacy rights.

B. Evolution of Abortion Law Jurisprudence

1. Roe v. Wade

Since the 1973 decision in *Roe v. Wade*[49] legalizing abortion, the debate has become more and more heated with each subsequent case decision and legislative enactment.[50] In *Roe v. Wade*, the United States Supreme Court struck down the Texas criminal abortion statutes holding that those statutes which prohibited abortions at any stage of pregnancy, except to save the life of the mother, were unconstitutional.[51] The established right to privacy in place prior to 1973 was expanded to include the right to an abortion.[52] *Roe v. Wade* recognized the right to an abortion free of those state regulations which do not promote state interests.[53] Any restrictions or limitations a state attempts to impose during the first two trimesters of pregnancy will be challenged using strict scrutiny, the highest level of judicial standard of review.[54]

---

48. *See id.* at 277 (stating that "the Supreme Court has . . . recognized that the government generally has a limited right to regulate the freedom of expression when there are important policies served by the regulation").

49. 410 U.S. 113 (1973).

50. *See supra* notes 9-10.

51. *See Roe*, 410 U.S. at 164-66. The Court stated,

> [o]n the merits, the District Court held that the 'fundamental right of single women and married persons to choose whether to have children is protected by the Ninth Amendment, through the Fourteenth Amendment,' and that the Texas criminal abortion statutes were void on their face because they were both unconstitutionally vague and constituted an overbroad infringement of the plaintiffs' Ninth Amendment rights.

Id. at 122.

52. *See id.* at 153.

53. *See id.* at 154.

54. *See Roe*, 410 U.S. at 155. Strict scrutiny dictates that the government interests being served through the regulation be "compelling" in addition to being important and legitimate.

2. Planned Parenthood of Southeastern Pa. v. Casey

Almost twenty years later, the United States Supreme Court decided another abortion case which challenged the validity of the *Roe v. Wade* decision.[55] During the twenty years between the two decisions, courts struck down numerous state laws that infringed upon the privacy rights of wed and unwed couples, and created new requirements related to abortion procedures.[56] In the 1992 decision of *Planned Parenthood of Southeastern Pa. v. Casey*,[57] abortion clinics and their medical staffs challenged the constitutionality of five provisions of the Pennsylvania Abortion Control Act of 1982.[58] The Act (1) required that a woman seeking an abortion give her informed consent prior to the procedure after (2) being provided with certain information at least twenty-four hours before the abortion, (3) mandated that a minor seeking an abortion provide the informed consent of one parent, (4) required a married woman to sign a statement indicating that she has notified her husband that she is seeking an abortion, and (5) dictated certain reporting requirements on facilities providing abortion services.[59] The Court, under the theory of *stare decisis*, held that the core decision in *Roe v. Wade* should be affirmed.[60]

> [T]he State does have an important and legitimate interest in preserving and protecting the health of the pregnant woman, whether she be a resident of the State or a nonresident who seeks medical consultation and treatment there, and that it has still *another* important and legitimate interest in protecting the potentiality of human life. These interests are separate and distinct. Each grows in substantiality as the woman approaches term and, at a point during pregnancy, each becomes 'compelling.'

Id. at 162-63.

55. *See* Planned Parenthood of Southeastern Pa. v. Casey, 505 U.S. 833 (1992) (holding that increased regulation of abortions was constitutional as long as it did not impose an undue burden on the pregnant woman or act as an obstacle to her right to choose to terminate the pregnancy).

56. *See Insight Roe v. Wade*, *supra* note 9, at 4C (discussing the evolution of cases, legislation, and abortion-related events that occurred since *Roe v. Wade*). In the twenty-seven year anthology presented, the article discussed cases invalidating statutes such as a Massachusetts law that prohibited the distribution of contraceptives to unwed people; laws requiring certain locations where abortions could be performed; consent requirements for spouses and minors; instances of violence at clinics including bombings and murders; and state and federal legislation enacted to restrict protest activity outside of abortion clinics. *See id.*

57. 505 U.S. 833 (1992).

58. *See Casey*, 505 U.S. at 844. *See also* 18 PA. CONS. STAT. §§ 3203, 3205, 3206, 3207, 3209, 3210, 3211 (1982).

59. *See Casey*, 505 U.S. at 844.

60. *See id.* at 846. The Court stated, "[a]fter considering the fundamental con-

The Court, however, changed some of the requirements set out in the *Roe v. Wade* case with respect to when a woman's right to choose begins and ends.[61] The result was that a woman has a right to choose an abortion at any time prior to fetal viability, and that the 'undue burden test' should be used in place of the trimester guidelines in assessing when the state is allowed to place restrictions on abortions prior to viability.[62] The Court wrote:

> A finding of an undue burden is a shorthand for the conclusion that a state regulation has the purpose or effect of placing a substantial obstacle in the path of a woman seeking an abortion of a nonviable fetus. A statute with this purpose is invalid because the means chosen by the State to further the interest in potential life must be calculated to inform the woman's free choice, not hinder it.[63]

States are able to enforce restrictions that do not impose an "undue burden" or "substantial obstacle" to abortion, such as the parental notification for minors or the twenty-four hour waiting period.[64] However, if the restriction is too burdensome, such as the requirement that a married woman notify her husband prior to the procedure, that restriction will not pass constitutional muster and will be overturned.[65] *Casey* remains the standard against which future decisions have been and will continue to be measured.[66]

---

stitutional questions resolved by *Roe*, principles of institutional integrity, and the rule of *stare decisis*, we are led to conclude this: the essential holding of *Roe v. Wade* should be retained and once again reaffirmed." *Id.* at 845-46.

61. *See id.* at 869-73.

62. *See id.* at 872-74. "Viability" is defined as: "Capability of living. A term used to denote the power a new-born child possesses of continuing its independent existence. That stage of fetal development when the life of the unborn child may be continued indefinitely outside the womb by natural or artificial life-supportive systems." BLACK'S LAW DICTIONARY 1404 (5th ed. 1979).

63. *Casey*, 505 U.S. at 877.

64. *See id.*

65. *See id.*

66. *See generally, e.g.*, Stenberg v. Carhart, 530 U.S. 914 (2000); Planned Parenthood of Central New Jersey v. Farmer, 220 F.3d 127, 130 (2000) (holding that New Jersey partial birth abortion statute "(1) is void for vagueness; and (2) places an undue burden on a woman's constitutional right to obtain an abortion"); Planned Parenthood of the Blue Ridge v. Camblos, 155 F.3d 352, 355 (1998) (holding that a Virginia act which required minors seeking abortions to notify one or both parents "as legislation that respects the fundamental interests of responsible parents in the rearing and in the educational, moral, and religious development of their children, without unduly burdening the fundamental abortion right, is facially constitutional under the Fourteenth Amendment"); Causeway Medical Suite v. Ieyoub, 109 F.3d 1096, 1099 (1997) (concluding that amendments to Louisiana abortion statute modifying the judicial bypass procedures regarding parental notification for minors seeking abortions do not "com-

3. Stenberg v. Carhart

The most recent development with regard to abortion regulation came in the *Stenberg v. Carhart*[67] ruling in June, 2000. This decision added yet another accelerant to the already heated abortion debate.[68] Dr. Carhart was a Nebraska physician who performed abortions at various times during a woman's pregnancy.[69] He challenged the constitutionality of the Nebraska statute[70] involving the criminalization and prohibition of "partial birth abortions,"[71] arguing that this statute was "so restrictive that 90 percent of the 5,000 abortions performed in the state each year would be threatened."[72] The United States Supreme Court ruled that the Nebraska statute violated the Constitution, and struck it down on two independent grounds: (1) the lack of "any exception 'for the preservation of the . . . health of the mother'"[73] and (2) "it 'imposes an undue burden on a woman's ability' to choose a D&E abortion, thereby unduly burdening the right to choose an abortion itself."[74] Guided by the standards of *Casey*, the Court applied and extended the pre-established limitations on abortion restrictions created by states against women making the difficult decision

---

port with the Due Process Clause of the Fourteenth Amendment" and were "permanently enjoined").

67. 530 U.S. 914, 918 (2000) (holding a Nebraska state statute criminalizing the performance of partial birth abortions unconstitutional due to its lack of exception for the preservation of the health of the mother and imposing an "undue burden upon a woman's right to make an abortion decision").

68. *See id.* at 920-21.

69. *See id.* at 922.

70. *See* NEB. REV. STAT. § 28-328 (Supp. 2000). Section (1) of the statute reads as follows: "No partial birth abortion shall be performed in this state, unless such procedure is necessary to save the life of the mother whose life is endangered by a physical disorder, physical illness, or physical injury, including a life-endangering physical condition caused by or arising from the pregnancy itself." *Id.*

71. A partial birth abortion, also known as a dilution and evacuation (D&E) abortion, is defined in the statute as "an abortion procedure in which the person performing the abortion partially delivers vaginally a living unborn child before killing the unborn child and completing the delivery." *Id.* at § 28-326(9). In more ordinary terminology, a partial birth abortion is "a second-trimester abortion procedure in which a doctor pulls a fetus feet first through the birth canal and then crushes its skull and empties its contents." Mary Leonard, *High Court Voids Nebraska Ban on Type of Abortion Ruling Cites Obstacle to a Woman's Right to End Pregnancy*, THE BOSTON GLOBE, June 29, 2000, at A1.

72. Leonard, *supra* note 71, at A1.

73. *Stenberg*, 530 U.S. at 930 (quoting *Casey*, 505 U.S. at 879).

74. *Id.* (quoting *Casey*, 505 U.S. at 874). *See supra* note 71 for explanation of D&E abortion.

of whether or not to have an abortion.[75]

4. *The FDA Approval of RU-486*

Another extension to the *Roe* and *Casey*-established guidelines regarding abortion procedures came in late September 2000.[76] The Food and Drug Administration (FDA) approved the abortion pill RU-486, ending "a tumultuous decade-long political struggle"[77] and debate over its introduction into American society.[78] As a result of the approval of the pill, patients will no longer have to undergo medical procedures requiring operations. Using the pill requires a few doctor's visits and a prescription for the drug within the forty-ninth day of pregnancy.[79] It gives "American women, for the first time, an abortion option that takes the procedure out of the operating room and into the quiet of the doctor's office . . . [making] abortion a more private affair, with the expulsion of the embryo most likely occurring in a woman's home."[80] However, with the ease and convenience of the pill's utility comes the potential negative impact on abortion clinics.[81] The pill has the likely effect of reducing the number of surgical abortions performed inside clinics, and also gives protesters another reason to express their views outside of them.[82] As a result of the FDA decision, security at abortion clinics and medical facilities has been increased as a safety precaution.[83]

---

75. *See Stenberg*, 530 U.S. at 921.

76. *See* Raja Mishra, *New Option Marks a Turning Point*, THE BOSTON GLOBE, Sept. 29, 2000, at A12.

77. *Id.*

78. RU-486 has been available in many western European nations since 1988, but was not permitted to be imported to the United States due to the Reagan-Bush ban on importation of drugs from abroad which was lifted by President Clinton prior to the FDA approval. *See id.*

79. *See id.*

80. *Id.*

81. *See* Mishra, *supra* note 76, at A12 (discussing how doctors are preparing for a "significant change in their practices" as a result of the pill's introduction).

82. *See* Daniel Barbarisi, *After RU-486 Setback, Abortion Foes Rally*, THE BOSTON GLOBE, Oct. 2, 2000, at B1. The FDA's approval of RU-486 was the main topic for speakers at a "Respect for Life" rally held by the Massachusetts Citizens for Life on October 1, 2000 in Boston. *See id.*

83. *See id.* One example of post-decision retaliation occurred in Illinois where "a Catholic priest drove his car into an Illinois clinic and chopped at the building with an axe until the owner fired off a shotgun to stop him." *Id.*

C. First Amendment & Abortion Clinic Protest Case Law Background and Analysis

The evolution of the constitutionality of buffer zones outside abortion clinics and medical facilities is best explained through three principal cases, which have directed the United States Supreme Court to its current posture regarding the constitutional balance of First Amendment rights and the rights of those seeking medical assistance.[84] These three cases form the foundation upon which future decisions will be made in neighboring jurisdictions.

1. Madsen v. Women's Health Center, Inc.

In *Madsen v. Women's Health Center, Inc.*,[85] a Florida state court in 1992 had issued an injunction permanently enjoining the blocking of, and interference with, public access to an abortion clinic, as well as enjoining the physical abuse of patients seeking services at these clinics by demonstrators outside the clinics.[86] The demonstrators continued their activities, and six months later, the operators of the abortion clinics sought to broaden the injunction to encompass a greater territory with stricter limitations.[87] The reasons for demanding the increased restrictions were due to the intensity of the protests in terms of marches outside the clinic doors and across its driveways, "sidewalk counselors"[88] approaching all vehicles seeking access to the facilities, and the large number of people demonstrating at all noise levels, varying "from singing and chanting to the use of loudspeakers and bullhorns."[89] Demonstrations not only occurred at the

---

84. *See* Madsen v. Women's Health Center, Inc., 512 U.S. 753 (1994); Schenck v. Pro-Choice Network of Western New York, 519 U.S. 357 (1997); Hill v. Colorado, 530 U.S. 703 (2000).

85. 512 U.S. 753 (1994).

86. *See Madsen*, 512 U.S. at 758. An injunction is defined as

> [a] prohibitive, equitable remedy issued or granted by a court at the suit of a party complainant, directed to a party defendant in the action . . . forbidding the latter to do some act . . . which he is threatening or attempting to commit, or restraining him in the continuance thereof, such act being unjust and inequitable, injurious to the plaintiff . . . .

BLACK'S LAW DICTIONARY 705 (5th ed. 1979).

87. *See Madsen*, 512 U.S. at 758.

88. *See id.* The term "sidewalk counselor" refers to those protestors whose purpose is to approach and talk with patients entering the clinic to educate them on alternatives to abortion. *See Schenck*, 519 U.S. at 363.

89. *Madsen*, 512 U.S. at 758. Testimony was presented by doctors, clinic em-

clinics themselves; the medical staff was subjected to picketing outside of their private residences while their neighbors and children were frequently confronted by demonstrators.[90]

The state court concluded "its original injunction had proved insufficient 'to protect the health, safety and rights of women . . . seeking access to [medical and counseling] services.'"[91] Thus the court issued the broader, amended injunction, the constitutionality of which ultimately was challenged before the United States Supreme Court.[92] In addition to the provision prohibiting the blocking of or interference with access to the clinic entrances and parking lots, the amended injunction also created a buffer zone prohibiting "congregating, picketing, patrolling, demonstrating or entering that portion of public right-of-way or private property within [thirty-six] feet of the property line of the Clinic."[93] The injunction also restricted the noise level permitted outside of the clinic during specific hours, prohibiting "singing, chanting, whistling, shouting, yelling, use of bullhorns, auto horns, sound amplification equipment or other sounds or images observable to or within earshot of the patients inside the Clinic."[94] A broader restriction was placed on "physically approaching any person seeking the services of the Clinic unless such person indicates a desire to communicate by approaching" the demonstrators,[95] keeping them to a distance of three hundred feet from the Clinic.[96] The injunction also prohibited "demonstrating or using bullhorns or other sound amplification equipment within [300] feet of the residence[s]" of any of the Clinic employees, staff and owners.[97] Lastly, the injunction prohibited "physically abusing, grabbing, intimidating, harassing . . . or assaulting persons entering or leaving, working at or using the services at the Clinic,"[98] including

---

ployees and patients as to the circumstances of these demonstrations. *See id.* at 758-59. One clinic doctor testified that, "as a result of having to run such a gauntlet to enter the clinic, the patients 'manifested a higher level of anxiety and hypertension causing those patients to need a higher level of sedation to undergo the surgical procedures, thereby increasing the risk associated with such procedures.'" *Id.* at 758. At times, patients also decided to leave without entering the clinics because of the demonstrations, which also increased their health risks by delaying the procedure until a later date. *See id.* at 758-59.

90. *See id.* at 759.

91. *Id.*

92. *See id.*

93. *Id.* (quoting Operation Rescue v. Women's Health Center, Inc., 626 So.2d 664, 679-80 (Fla. 1993)).

94. *Madsen*, 512 U.S. at 760.

95. *Id.*

96. *See id.*

97. *Id.*

98. *Id.* at 760-61.

"any present or former doctor, health care professional, or other staff member, employee or volunteer who assists in providing services at the Clinic."[99]

The Florida Supreme Court upheld the amended injunction as constitutional, stating that the restrictions were content-neutral,[100] a requirement for upholding the constitutionality of an injunction or statute.[101] The Supreme Court ruled that the restrictions met requirements for constitutionality because they were "narrowly tailored to serve a significant government interest, and left open ample channels of communication . . . ."[102]

At the same time that the Florida State Supreme Court was deciding the *Madsen* case, the United States Court of Appeals for the Eleventh Circuit was deciding another case challenging the same injunction, and arrived at the opposite conclusion.[103] Since a conflict existed between the decisions of the Supreme Court of Florida and the United States Appeals Court, the United States Supreme Court granted certiorari on the present case to resolve that inconsistency.[104] Ultimately, the Court held that the "establishment of a 36-foot buffer zone on a public street from which demonstrators are excluded passes muster under the First Amendment, but that several other provisions of the injunction do not."[105]

The United States Supreme Court in *Madsen* abandoned the then-controlling 'time, place and manner test' for both injunctions and statutes that regulate speech.[106] The requirements for this test were set forth in *Ward v. Rock Against Racism.*[107] The standard required that the injunction

---

99. *Id.* at 761.

100. *See Madsen*, 512 U.S. at 761 (citing Operation Rescue v. Women's Health Center, Inc., 626 So.2d 664).

101. *See infra* notes 112-17 and accompanying text. "Content neutrality" refers to the requirement that the government be "both viewpoint neutral and subject matter neutral" meaning that the government cannot "regulate speech based on the ideology of the message . . . [or] on the topic of the speech." Chemerinsky, *supra* note 29, at 202-03.

102. *Madsen*, 512 U.S. at 761.

103. *See id.* (quoting Cheffer v. McGregor, 6 F.3d 705, 711 (11th Cir. 1993) in which the Court of Appeals struck down the injunction, "characterizing the dispute as a clash 'between an actual prohibition of speech and a potential hinderance to the free exercise of abortion rights.'").

104. *See id.* at 762.

105. *Madsen*, 512 U.S. at 757.

106. *See id.* at 764. The *Madsen* Court, in creating a new test for constitutionality, stated, "when evaluating a content-neutral injunction, we think that our standard time, place, and manner analysis is not sufficiently rigorous. We must ask instead whether the challenged provisions of the injunction burden no more speech than necessary to serve a significant government interest." *Id.* at 765.

107. 491 U.S. 781 (1989) (the Court upheld a municipal noise regulation de-

be content neutral,[108] "narrowly tailored to serve a significant government interest"[109] and leave open ample alternative channels of communication.[110] Content neutrality is critical for evaluating whether someone's First Amendment rights have been unconstitutionally restricted or violated.[111] The government cannot restrict speech based on the message contained in the speech.[112] "The government simply cannot attempt to impose an orthodoxy of belief by favoring some ideas over others."[113] The "injunction's restriction of speech must be incidental to the purpose of the injunction."[114] One such example would be an injunction ordered to enjoin the conduct of anti-abortion demonstrators.[115] Since the injunction's purpose in that example is not directed at curbing the message of the protestors, but rather the hostile activities of the protestors, it would be considered to be content neutral under the *Ward* standard.[116]

The *Madsen* court rejected the *Ward* standard for court-ordered injunctions in favor of a new standard of "whether the challenged provisions of the injunction burden no more speech than necessary to serve a significant government interest."[117] The reason for discarding the prevailing standard

---

signed to ensure that musical performances at a public band shell did not disturb residents who lived nearby). The standard the Court used to test the constitutionality of the regulation was a less stringent test; the regulation in question having to meet three criteria to satisfy the test: (1) "whether the government has adopted a regulation of speech because of disagreement with the message it conveys" (whether the provision is content-neutral); (2) whether the regulation is "narrowly tailored to serve a significant governmental interest"; and (3) whether the regulation has "[left] open ample alternative channels of communication." *Id.* at 791-802.

108. *See id.* at 791.

109. *Id.* The Florida Supreme Court had identified a number of legitimate state interests being served by the injunction: "protecting a woman's freedom to seek lawful medical or counseling services in connection with her pregnancy," "ensuring the public safety and order," "promoting the free flow of traffic on public streets and sidewalks," "protecting the property rights of all its citizens," and "the State's strong interest in residential privacy, acknowledged in *Frisby v. Schultz*, 487 U.S. 474 (1988), applied by analogy to medical privacy." *Madsen*, 512 U.S. at 767-68.

110. *See Ward*, 491 U.S. at 791.

111. *See* Alan Phelps, Note, *Picketing and Prayer: Restricting Freedom of Expression Outside Churches*, 85 CORNELL L. REV. 271, 281 (1999) (discussing content neutrality with regard to government restrictions on regulated speech).

112. *See id.*

113. Chemerinsky, *supra* note 29, at 212.

114. Marjorie C. Jones, Recent Development, F. Schenck v. Pro-Choice Network of Western New York: *Floating Buffer Zones Created By Injunction Violate First Amendment Free Speech*, 24 J. CONTEMP. L. 154, 157 (1998).

115. *See id.* at 157-58.

116. *See id.*

117. *Madsen*, 512 U.S. at 765.

was due to the nature of the injunction itself.[118] Injunctions are issued to restrict the behavior of a specific individual or group of individuals, or punish those who have already violated similar decrees, whereas ordinances and statutes are applicable to all persons.[119] The Court determined that the "standard time, place, and manner analysis is not sufficiently rigorous"[120] for injunctions due to the potential risk of their censorship and discriminatory application.[121] The heightened scrutiny for injunctions shifted along the spectrum to rest midway between the strict scrutiny test for content-based regulation of speech and the intermediate time, place, and manner test traditionally used for content-neutral injunctions and ordinances.[122]

In deciding the case, the Court first examined the thirty-six-foot buffer zone outside the clinic.[123] The buffer zone provision prohibited demonstration activity within that zone, requiring protestors to move across the street and away from the clinic's driveway.[124] "The 36-foot buffer zone protecting the entrances to the clinic and the parking lot is a means of protecting unfettered ingress to and egress from the clinic, and ensuring that petitioners do not block traffic . . . ."[125] Balancing public safety and welfare against the restrictions it places on the demonstrators' abilities to still communicate their messages to the patients, the Court held that "the 36-foot buffer zone around the clinic entrances and driveway burdens no more speech than necessary to accomplish the governmental interest at stake."[126]

As for the thirty-six-foot buffer zone on the surrounding private property to the clinic, the "accepted purpose of the buffer zone [was] to protect access to the clinic and facilitate the orderly flow of traffic."[127] Since there was no absolute necessity for patients or medical staff to cross this land to access the clinic, nor was there evidence that petitioners on this property caused obstructions to the clinic entrances from that location, the Court held that this provision "burden[ed] more speech than necessary to protect

---

118. *See id.* at 764.

119. *See id.*

120. *Id.* at 765.

121. *See id.* at 764-65 (quoting from Railway Express Agency, Inc. v. New York, 336 U.S. 106, 112-13 (1949)), stating "[T]here is no more effective practical guaranty against arbitrary and unreasonable government than to require that the principles of law which officials would impose on a minority must be imposed generally."

122. *See* Miller, *supra* note 42, at 283.

123. *See Madsen*, 512 U.S. at 768.

124. *See id.* at 768-69.

125. *Id.* at 769.

126. *Id.* at 770. The governmental interest being "[eliminating] the unlawful conduct targeted by the state court's injunction." *Id.* at 776.

127. *Id.* at 771.

access to the clinic."[128]

The Court's holding with respect to the noise level restrictions and "images observable" prohibition, resulted in opposite outcomes.[129] The Court upheld the limited noise level restrictions as "burden[ing] no more speech than necessary to ensure the health and well-being of the patients at the clinic."[130] The Court stated that "[n]oise control is particularly important around hospitals and medical facilities"[131] and stated:

> Hospitals, after all, are not factories or mines or assembly plants. They are hospitals, where human ailments are treated, where patients and relatives alike often are under emotional strain and worry, where pleasing and comforting patients are principal facets of the day's activity, and where the patient and his family . . . need a restful, uncluttered, relaxing, and helpful atmosphere.[132]

As for the "images observable" provision, since the state court had issued a "blanket ban on all 'images observable,'"[133] the Court found this to "burden more speech than necessary to achieve the purpose of limiting threats to clinic patients or their families"[134] and therefore a violation of the First Amendment.[135] Relating this back to the content-neutrality requirement, the Court said that "[t]he only plausible reason a patient would be bothered by 'images observable' inside the clinic would be if the patient found the expression contained in such images disagreeable."[136]

The Court also struck down the final two provisions of the injunction.[137] The cease-and-desist order based on the prohibition of "physically approaching any person seeking the services of the clinic 'unless such person indicates a desire to communicate' in an area within 300 feet of the

---

128. *Id.* The Court stated that "these provisions sweep more broadly than necessary to accomplish the permissible goals of the injunction." *Id.* at 776.

129. *See Madsen*, 512 U.S. at 772-73. The "images observable" provision concerned images and signs shown outside the abortion clinics which could be seen by patients inside the clinics and might be distressing to those patients. *See id.*

130. *Id.* at 772.

131. *Id.*

132. *Id.* (quoting NLRB v. Baptist Hospital, Inc., 442 U.S. 773, 783-84, n.12 (1979)).

133. *Id.* at 773. The Court said that it was "much easier for the clinic to pull its curtains than for a patient to stop up her ears, and no more is required to avoid seeing placards through the windows of the clinic." *Id.*

134. *Id.*

135. *See id.*

136. *Madsen*, 512 U.S. at 773.

137. *See id.* at 774-75.

clinic"[138] burdened more speech than was necessary.[139] Chief Justice Rehnquist "failed to find enough evidence 'to justify a prohibition on all uninvited approaches of persons seeking the services of the clinic, regardless of how peaceful the contact might be, without burdening more speech than necessary to prevent intimidation and to ensure access to the clinic.'"[140] The Court reasoned that "[a]bsent evidence that the protestors' speech is independently proscribable (i.e., 'fighting words' or threats), or is so infused with violence as to be indistinguishable from a threat of physical harm . . . this provision cannot stand."[141]

The 300-foot buffer zone around the residences of the employees and medical staff was also found to be too burdensome due to the sizeable foot restriction imposed, which was larger than those allowed in previous cases.[142] The Court stated that "a limitation on the time, duration of picketing, and number of pickets outside a smaller zone could have accomplished the desired result."[143]

To summarize, the *Madsen* court allowed the enforcement buffer zones outside of abortion clinics as long as they burdened no more speech than necessary to promote the legitimate government interests, and those which did burden more speech than necessary were ruled unconstitutional.[144] The Court also changed the constitutional standard for evaluating court-ordered injunctions that restrict speech, differentiating them from state statutes and ordinances which retain the old test.[145]

2. Schenck v. Pro-Choice Network of Western New York

In the second principal case, *Schenck v. Pro-Choice Network of Western New York*,[146] the Court distinguished constitutional buffer zones from un-

---

138. *Id.* at 773.

139. *See id.* at 774.

140. Pettinari, *supra* note 4, at 817.

141. *Madsen*, 512 U.S. at 774.

142. *See id.* at 775. *See also* Frisby v. Schultz, 487 U.S. 474, 484 (1988). In *Frisby*, anti-abortion protestors picketed in a residential area outside the home of a doctor who performed abortions. *See id.* at 476. As a result, the town passed a series of ordinances restricting picketing, ultimately making it "unlawful for any person to engage in picketing before or about the residence or dwelling of any individual in the Town of Brookfield." *Id.* at 477. The ordinance was challenged as being unconstitutional on its face, however the challenge failed and the United States Supreme Court upheld the ordinance. *See id.* at 488.

143. *Madsen*, 512 U.S. at 775.

144. *See id.* at 757.

145. *See id.* at 765-66.

146. 519 U.S. 357 (1997).

constitutional buffer zones.[147] One interesting thing about the *Schenck* case is that the United States District Court for the Western District of New York heard the case prior to the *Madsen* decision, while the Court of Appeals and United States Supreme Court proceedings were conducted post-*Madsen*.[148] The outcomes of the *Schenck* proceedings, pre- and post-*Madsen*, resulted in opposite conclusions due to the shifts in the law that *Madsen* effected.[149]

The plaintiffs in *Schenck* were a group of clinics and doctors in upstate New York, who performed abortions at these clinics, that joined with Pro-Choice of Western New York[150] to file a complaint against several pro-life organizations who had planned a blockade of an abortion clinic in 1990.[151] The complaint referenced the past activities of these groups at the clinics and the numerous illegal blockades that had already occurred,[152] and asked

---

147. *See id.* at 361. The distinction between the "fixed buffer zone" or "fixed bubble" and "floating buffer zone" or "floating bubble" is basically that the "fixed" zone is a set area, within which protestors are not allowed to demonstrate or obstruct patients from accessing, whereas "floating" zones travel with the patients themselves: the distance is a set number of feet, but the actual location travels with the patients as they walk from the car to the clinic entrance, for example, and without consent, protestors cannot approach the patient within that surrounding "bubble" at any time during her journey. *See id.* at 371. "A 'floating' or 'bubble' buffer zone does not have a fixed point of reference, but rather follows a moving object or person." Opinion of the Justices to the Senate, 430 Mass. 1205, 1211 n.6 (2000).

148. *See Schenck*, 519 U.S. at 370.

149. *See id.* at 374 (stating that the Court must "apply *Madsen* to the challenged provisions of the injunction and ask whether they burden more speech than necessary to serve a significant governmental interest.").

150. Pro-Choice is "a not-for-profit corporation aimed at maintaining safe and legal access to family planning and abortion services." Deborah A. Ellis & Yolanda S. Wu, *Abortion Rights: Of Buffer Zones and Broken Bones: Balancing Access to Abortion and Antiabortion Protestors' First Amendment Rights in* Schenck v. Pro-Choice Network, 62 BROOK. L. REV. 547, 551 (1996) (discussing the effects of the *Madsen* decision on the *Schenck* case and issues left open from both decisions).

151. *See Schenck*, 519 U.S. at 362.

152. *See id.* The activities that had occurred at the clinics were similar to those in the *Madsen* case. "[P]rotestors would march, stand, kneel, sit or lie in parking lot driveways and in doorways. This conduct blocked or hindered cars from entering clinic parking lots, and patients, doctors, nurses, and other clinic employees from entering the clinics." *Id.* at 362. Some protestors went so far as to "[throw] themselves on top of the hoods of cars" or "used more aggressive techniques, with varying levels of belligerence: getting very close to women entering the clinics and shouting in their faces; surrounding, crowding, and yelling at women entering the clinics; or jostling, grabbing, pushing, and shoving women as they attempted to enter the clinics." *Id.* at 363. There also were "sidewalk counselors" present outside the clinics, targeting women who were headed for the clinic entrances, attempting to persuade the women not to go through with the abortion procedure. *See id.* "[I]f the women continued toward the clinics and

for a temporary restraining order (TRO) to be issued against the protestors to stop a planned blockade for later that month.[153] During the district court proceedings, the court found that the local police had been "unable to respond effectively to the protests"[154] and that "[t]here [had] been substantial uncontradicted evidence that defendants' activities [were] intended, and [did] in fact, prevent and hinder local police from protecting the right of a woman to choose to have an abortion."[155] The day before the large blockade was scheduled to occur, the district court issued the TRO.[156]

The pro-life groups complied with the TRO for a period of time after it was issued, and held peaceful demonstrations in place of the normal block-

---

did not respond positively to the counselors, such peaceful efforts at persuasion often devolved into 'in your face' yelling, and sometimes into pushing, shoving, and grabbing." *Id.*

153. *See Schenck*, 519 U.S. at 362. The complaint also requested that the court issue a permanent injunction against the anti-choice groups and award damages. *See id.* A temporary restraining order is defined as "[a]n emergency remedy of brief duration which may issue only in exceptional circumstances and only until the trial court can hear arguments or evidence . . . on the subject matter of the controversy and otherwise determine what relief is appropriate." BLACK'S LAW DICTIONARY 1313 (5th ed. 1979). The TRO was sought by the plaintiffs due to the strict time deadline surrounding the impending blockade. *See Schenck*, 519 U.S. at 362.

154. *Id.* at 363. The Court referenced numerous reasons such as the constant nature of the protests which overwhelmed the level of police resources available; the fact that the protestors would disperse when the police arrived, but would then return later after the police had left; testimony was difficult due to the private nature of abortions and patients not wanting to be identified in public; and the fact that a lot of the protestors would return to engaging in similar illegal conduct even after prior convictions. *See id.* at 363-64.

155. *Id.* at 364.

156. *See id.* The TRO was intended to stay in effect until a decision was made by the Court on the motion for preliminary injunction by the pro-choice parties. *See id.* The Court stated,

> In pertinent part, the TRO enjoined defendants from physically blockading the clinics, physically abusing or tortiously harassing anyone entering or leaving the clinics, and "demonstrating within 15 feet of any person" entering or leaving the clinics. As an exception to this 15-foot "buffer zone" around people, the TRO allowed two sidewalk counselors to have "a conversation of a nonthreatening nature" with individuals entering or leaving the clinic. If the individuals indicated that they did not want the counseling, however, the counselors had to "cease and desist" from counseling.

Schenck, 519 U.S. at 364. The district court interpreted the cease and desist provision to apply only if "the targeted person or group of persons indicates, either verbally or non-verbally that they do not wish to be counseled." *Id.* at n.1 (quoting Pro-Choice Network of Western N.Y. v. Project Rescue Western N.Y., 799 F. Supp. 1417, 1434 (W.D.N.Y. 1992)).

ades.[157] However, within a short time their conduct reverted back to pre-TRO type behavior, engaging in what the district court termed "constructive blockades"[158] and sidewalk counseling.[159] In the months following the issuance of the TRO, the intensity of the demonstrations increased to the levels of "aggressive sidewalk counseling with occasional shoving and elbowing, trespassing into clinic buildings to continue counseling of patients, and blocking of doorways and driveways."[160]

In February of 1992, the district court issued the permanent injunction that had been requested when the complaint had been filed in 1990, which followed closely the restrictions of the TRO, but also made some significant changes.[161] The injunction expanded upon the language of the TRO, which had "banned 'demonstrating . . . within fifteen feet of any person' entering or leaving the clinics,"[162] by extending the language to: more broadly [ban] "demonstrating within fifteen feet from either side or edge of, or in front of, doorways or doorway entrances, parking lot entrances, driveways and driveway entrances of such facilities" (fixed buffer zones), or "within fifteen feet of any person or vehicle seeking access to or leaving from such facilities" (floating buffer zones).[163]

It was here that the court created a distinction between what is now known as floating and fixed buffer zones.[164] The injunction also clarified the meaning of the "cease and desist" provision, "specifying that once sidewalk counselors who had entered the buffer zones were required to 'cease and desist' their counseling, they had to retreat 15 feet from the people they had been counseling and had to remain outside the boundaries of the buffer zones."[165]

The defendants challenged the injunction as a violation of their First

---

157. *See Schenck*, 519 U.S. at 365.

158. *Id.* "Constructive blockades consisted of 'demonstrating and picketing around the entrances of the clinics, and . . . harassing patients and staff entering and leaving the clinics.'" *Id.* (quoting Pro-Choice Network of Western N.Y. v. Project Rescue Western N.Y., 799 F. Supp. 1417, 1424 (W.D.N.Y. 1992)). However, "[t]he purpose of the constructive blockades was the same as physical blockades: 'to prevent or dissuade patients from entering the clinic.'" *Id.*

159. *See id.* at 363. *See also supra* note 88 for definition of sidewalk counseling.

160. *Schenck*, 519 U.S. at 365 (as evidenced by testimony from clinic volunteer escorts).

161. *See id.* at 366-67.

162. *Id.* at 367.

163. *Id.*

164. *See id. See also supra* note 148 and accompanying text for a more detailed description of the two types of buffer zones.

165. *Id.*

Amendment right to free speech.[166] Since these initial proceedings were prior to the *Madsen* decision, the court analyzed the constitutionality of the injunction based on the then-controlling "time, place, and manner analysis" whose three criteria consisted of: (1) content-neutrality; (2) narrowly tailored to serve a legitimate government interest; and (3) leaves open ample alternative channels of communication.[167] The court held that all three criteria were satisfied, thereby upholding the injunction.[168]

The defendants appealed, but by the time the case was heard by the Court of Appeals for the Second Circuit, the United States Supreme Court had decided *Madsen v. Women's Health Center, Inc.*[169] The test for evaluating injunctions had changed from the "time, place, and manner analysis" used in the district court, to the *Madsen* heightened scrutiny test of "whether the challenged provisions of the injunction burden no more speech than necessary to serve a significant government interest."[170] A panel of the court of appeals first heard the case and reversed the decision of the district court, finding the buffer zone provision and "cease and desist" provision burdened more speech than necessary under the *Madsen* standard.[171] The court of appeals then heard the case en banc limiting consideration to the two provisions struck down by the panel decision, and, again applying *Madsen*, affirmed the district court decision holding that "both the buffer zone provision and the 'cease and desist' provision [were] constitutional."[172] The United States Supreme Court then granted the petitioner's writ of certiorari.[173]

---

166. *See Schenck*, 519 U.S. at 369.

167. *See id.*

168. *See id.* at 369-70. The court found the injunction to be content neutral because it "merely restrict[ed] the volume, location, timing and harassing and intimidating nature of the defendants' expressive speech." *Id.* at 369. The court also held that the injunction was narrowly tailored to serve three legitimate government interests: "public safety, ensuring that abortions are performed safely, and ensuring that a woman's constitutional rights to travel interstate and to choose to have an abortion were not sacrificed in the interest of defendants' First Amendment rights." *Id.* Lastly, the court held that the injunction "left open ample alternative channels for communication, because defendants could still 'picket, carry signs, pray, sing or chant in full view of people going into the clinics.'" *Id.* at 370.

169. *See Schenck*, 519 U.S. at 370.

170. *Id.* at 371 (quoting *Madsen*, 512 U.S. at 765).

171. *See id. See also* Ellis & Wu, *supra* note 151, at 564.

172. Pro-Choice Network of Western N.Y. v. Schenck, 67 F.3d 377, 393 (2d Cir. 1994) (en banc), *aff'd in part, rev'd in part*, 519 U.S. 357 (1997).

173. *See Schenck*, 519 U.S. at 371. "Because *Schenck* was the first court of appeals decision to apply the *Madsen* standard, the Court granted certiorari in the absence of any conflict among the circuits on the application of *Madsen*." Ellis & Wu, *supra* note 151, at 551.

In 1997, seven years after the case was first commenced, the United States Supreme Court issued the final decision in the case, analyzing each of the challenged provisions[174] against the *Madsen* standard of "whether they burden more speech than necessary to serve a significant governmental interest."[175] In refuting petitioner's argument that there were no significant governmental interests being served by the injunction, the Court stated that all of the interests which had previously been found present in *Madsen* could be extended to *Schenck* due to the similarities of the two cases:

> [E]nsuring public safety and order, promoting the free flow of traffic on streets and sidewalks, protecting property rights, and protecting a woman's freedom to seek pregnancy-related services . . . and in combination [were] certainly significant enough to justify an appropriately tailored injunction to secure unimpeded physical access to the clinics.[176]

Although the injunction satisfied the first prong of the *Madsen* test, the Court did not uphold all of the provisions of the injunction being challenged.[177] The Court struck down the floating buffer zone provision stating:

> We strike down the floating buffer zones around people entering and leaving the clinics because they burden more speech than is necessary to serve the relevant governmental interests. The floating buffer zones prevent defendants – except for two sidewalk counselors, while they are tolerated by the target individual – from communicating a message from a normal conversational distance or handing leaflets to people entering or leaving the clinics who are walking on the public sidewalks. This is a broad prohibition, both because of the type of speech that is restricted and the nature of the location.[178]

---

174. *See Schenck*, 519 U.S. at 371. These provisions were:
(i) the floating 15-foot buffer zones around people and vehicles seeking access to the clinics; (ii) the fixed 15-foot buffer zones around the clinic doorways, driveways, and parking lot entrances; and (iii) the 'cease and desist' provision that forces sidewalk counselors who are inside the buffer zones to retreat 15 feet from the person being counseled once the person indicates a desire not to be counseled.

Id.

175. *Id.* at 374.

176. *Id.* at 376.

177. *See id.* at 377-85.

178. *Schenck*, 519 U.S. at 377. The Court expressed that since the buffer zone "floats" with the person (as well as with vehicles), "it would be quite difficult for a protestor who wishes to engage in peaceful expressive activities to know how to remain

As for the fixed buffer zone provision, the Court upheld this provision stating that they "are necessary to ensure that people and vehicles trying to enter or exit the clinic property or clinic parking lots can do so."[179] The Court deferred to the district court with regard to the choice of the distance of fifteen feet for the buffer zone, and stated that "[i]t is clear from the District Court's opinion that its decision to allow two sidewalk counselors inside the buffer zones was an effort to bend over backwards to 'accommodate' defendants' speech rights."[180] Finally, the Court also upheld the "cease and desist" provision that related to the sidewalk counselor exception, holding that it was not contrary to the First Amendment.[181] Therefore, as of the final decision in *Schenck*, floating buffer zones were determined to be unconstitutional according to the United States Supreme Court.[182]

The cases previously discussed, *Madsen* and *Schenck*, reflect the evolution of the balancing between First Amendment rights of abortion clinic demonstrators and a woman's right to an abortion as it relates to court ordered injunctions. Although the "*Madsen* 'burdens' test does not apply to generally applicable ordinances, the Supreme Court in *Schenck* emphasized that *any* floating buffer zone establishes a broad prohibition on speech that affects traditional types of speech . . . ."[183] Until the next case addressing the constitutionality of floating buffer zones was brought before

---

in compliance with the injunction. This lack of certainty leads to a substantial risk that much more speech will be burdened than the injunction by its terms prohibits." *Id.* at 378.

179. *Id.* at 380. The Court agreed with the district court's decision to take the prior conduct of the protestors into account when evaluating the fixed buffer zones. *See id.* "Based on this conduct, the District Court was entitled to conclude that protestors who were allowed close to the entrances would continue right up to the entrance, and that the only way to ensure access was to move *all* protestors away from the doorways." *Id.* at 381. This fixed buffer zone was not nearly as restrictive as one that was imposed in the case of *Planned Parenthood Shasta-Diablo, Inc. v. Williams*, where the injunction against the protestors forced them to remain on the sidewalk across the street from the clinic to keep entry to the clinic accessible. *See* Ronnie Beth Nadell, Comment, *Planned Parenthood Shasta-Diablo, Inc. v. Williams: Anti-Abortion Protestors, Injunctions, and the Suffocation of Free Speech in California*, 28 SW. U. L. REV. 627, 629 (1999). Although the sidewalk across the street also happened to be across a busy four-lane highway, the Court in this case upheld the injunction. *See id.* at 642.

180. *Schenck*, 519 U.S. at 381 n.11.

181. *See id.* at 383. The Court also noted that "[t]hese counselors remain free to espouse their message outside the 15-foot buffer zone, and the condition on their freedom to espouse it within the buffer zone is the result of their own previous harassment and intimidation of patients." *Id.* at 385.

182. *See id.* at 377.

183. Miller, *supra* note 42, at 311.

the United States Supreme Court, the holding in *Schenck* striking down floating buffer zones remained the prevailing rule.[184]

3. Hill v. Colorado

The third principal, and most recent, case that was heard before the United States Supreme Court regarding the constitutionality of buffer zones outside of an abortion clinic was *Hill v. Colorado*.[185] Unlike *Madsen* and *Schenck* which challenged an injunction ordered by the courts, this case involved a state statute in Colorado which allowed for "floating" buffer zones outside abortion clinics, and introduced a new requirement of "knowingly approaches" for violation by demonstrators. [186] There is a convoluted procedural history in this case between the state and federal court systems due to other cases being decided at the same time in other jurisdictions,[187] which had varying degrees of influence on the present case at various times. A clear understanding of this case is important for the analysis of the newly enacted Massachusetts law discussed in the next section of this Note.

As a result of the abortion-related violence that was making its way through the United States in the early 1990s, many state legislatures began enacting stricter regulations on activities outside of abortion clinics.[188] In 1993, the Colorado legislature enacted an ordinance "aimed at protecting employees and patients of abortion clinics."[189] The text of the Colorado statute[190] was a little different than the texts seen in the injunctions in the prior two cases.[191] The statute states:

> [n]o person shall knowingly approach another person within eight feet of

---

184. *See supra* notes 147-83 and accompanying text for detailed discussion of *Schenck v. Pro-Choice Network of Western New York*, 519 U.S. 357 (1997).

185. 503 U.S. 703, 735 (2000) (upholding a Colorado state statute establishing a buffer zone outside of abortion clinics).

186. *See id.* at 707 (quoting COLO. REV. STAT. § 18-9-122(3) (1999)).

187. *See generally* Madsen v. Women's Health Ctr., Inc., 512 U.S. 753 (1994); Schenck v. Pro-Choice Network of W. N.Y., 519 U.S. 357 (1997).

188. *See* Pettinari, *supra* note 4, at 803. Ironically, one of the catalysts for the Colorado statute was the abortion clinic killings and bombings that occurred in Massachusetts in 1993-1994. *See id.* at n.2.

189. *Id.* at 803.

190. *See* COLO. REV. STAT. § 18-9-122(3) (1999).

191. *See id. But cf. supra* notes 91-99, 163-66 and accompanying text.

> such person, unless such other person consents, for the purpose of passing a leaflet or handbill to, displaying a sign to, or engaging in oral protest, education, or counseling with such other person in the public way or sidewalk area within a radius of one hundred feet from any entrance door to a health care facility. Any person who violates this subsection (3) commits a class 3 misdemeanor.[192]

The statute introduced a scienter element, requiring that the protester actually intend to approach a patient for one of the stated purposes, and the statute applied to all health care facilities, not specifically limited to abortion clinics.[193]

The case at hand was brought to the Colorado state court system as a facial challenge to the statute five months after it was enacted.[194] The petitioners were counselors who had engaged in sidewalk counseling[195] outside of abortion clinics prior to the statute being enacted.[196] They challenged the constitutionality of the statute on its face, stating "such activities [related to sidewalk counseling] frequently entailed being within eight feet of other persons and that their fear of prosecution under the new statute caused them 'to be chilled in the exercise of fundamental constitutional rights'"[197] and sought an injunction against its enforcement.[198] The complaint argued that the statute was "vague and overbroad, and that it was a content-based restriction that was not justified by a compelling state interest."[199]

This case involved much of the same type of facts and testimony as seen in the *Madsen* and *Schenck* cases; witnesses describing the need for escorts to accompany patients into and out of the clinics to protect them from the harassment and confrontations by the demonstrators.[200] There was also testimony related to the medical repercussions that these confrontations

---

192. COLO. REV. STAT. § 18-9-122(3) (1999).

193. *See id.*

194. *See Hill*, 530 U.S. at 708.

195. *See supra* note 88 and accompanying text.

196. *See Hill*, 530 U.S. at 708.

197. *Id.* at 708-09.

198. *See id.* at 708.

199. *Id.* at 709.

200. *See id.* at 709-10. One volunteer who acts as a patient escort testified that the protestors "'are flashing their bloody fetus signs. They are yelling, 'you are killing your baby.' They are talking about fetuses and babies being dismembered, arms and legs torn off . . . a mother and her daughter . . . were immediately surrounded and yelled at and screamed at . . . .'" *Id.* at n.7 (quoting Hill v. Thomas, 973 P.2d 1246, 1250-51 (Colo. 1999)).

had on the patients, both physically and mentally.[201]

The start of the complicated procedural history begins with the District Court judge dismissing the complaint, holding that "the statute permissibly imposed content-neutral 'time, place, and manner restrictions' that were narrowly tailored to serve a significant government interest, and left open ample alternative channels of communication."[202] The judge applied the *Ward* factors and "found that the text of the statute 'applies to all viewpoints, rather than certain viewpoints,' and that the legislative history made it clear that the State had not favored one viewpoint over another."[203] He also concluded that the statute's "free zone" satisfied the other two *Ward* criteria in that it was narrowly tailored to the government's significant legitimate interests and left open ample alternative means of communication since within the eight-foot distance, the signs and leaflets of the demonstrators could still be seen, and conversation-level speech could still be heard.[204]

The Court of Appeals for the state of Colorado then affirmed the District Court decision, noting that "even though only seven percent of the patients receiving services at one of the clinics were there to obtain abortion services, all 60,000 of that clinic's patients 'were subjected to the same treatment by protestors.'"[205] The Court of Appeals also reviewed the case in light of the United States Supreme Court decision in *Madsen*, finding that the reasoning supporting the decision in *Madsen* also supported the decision in *Hill*.[206] The court stated,

> section 18-9-122 was content-neutral, because "[l]ike the injunction in Madsen, § 18-9-122(3) does not address only the speech of ant[i]-abortion protestors. The statute would impose the same restriction upon protestors demonstrating in favor of abortion. The law would also apply to protest activity

---

201. *See id.* at 710. For example, one of the witnesses who testified had been through thirty-five surgeries in the preceding eight years and stated:

> Each and every one is tough. And the night before and the morning of any medical procedure that's invasive is the toughest part of all. You don't need additional stressors placed on you while you're trying to do it . . . . We all know about our own personal faith. You don't need somebody standing in your face screaming at you when you are going in for what may be one of the most traumatic experiences of your life anyway. Why make it more traumatic?

Id. at n.8.

202. *Hill*, 530 U.S. at 710.

203. *Id.* at 711.

204. *See id.*

205. *Id.* at 711.

206. *See Hill*, 530 U.S. at 712.

> directed at patients requiring or seeking advice relative to an organ transplant."[207]

When the appeal was taken to the Supreme Court of Colorado, that court denied review and the petitioners then sought a writ of certiorari from the United States Supreme Court.[208] This petition was made during the same time-frame that the United States Supreme Court decided the previously discussed *Schenck* case.[209] Because of this contemporaneous decision, the Court, without issuing a decision, granted certiorari and vacated the Colorado Court of Appeals decision and remanded the case to be considered in light of the *Schenck* decision.[210]

Upon remand to the Colorado Court of Appeals, the court reinstated its judgment holding that "*Schenck* was not controlling and reiterated its earlier position that the controlling authority in *Hill* was *Ward*."[211] The court of appeals basically abandoned the *Madsen* test, which was the controlling standard since *Madsen* and *Schenck*, for deciding the statute's constitutionality and reverted back to prior case law,[212] and the statute was upheld.[213]

The Supreme Court of Colorado next granted certiorari to decide "whether a legislative enactment designed to protect the privacy rights of citizens entering and leaving Colorado health care facilities unduly burdens the First Amendment rights of other citizens."[214] The court concluded it did not, and affirmed the judgment of the Colorado Court of Appeals.[215] The court held that "the First Amendment can accommodate reasonable governmental action intended to effectuate the free exercise of another fundamental right, an individual's right to privacy, here represented by access to medical counseling and treatment."[216]

The court also reiterated that the proper analysis for evaluating the constitutionality of the statute was the "time, place, and manner" test outlined in *Ward*, and not the test enunciated in *Madsen* and *Schenck*.[217] In doing

---

207. Pettinari, *supra* note 4, at 812.

208. *Hill*, 120 S. Ct. at 2487.

209. *See supra* notes 147-83 and accompanying text.

210. *See Hill*, 530 U.S. at 712.

211. Pettinari, *supra* note 4, at 806.

212. Since this was a state statute, the court evaluated its constitutionality based on the *Ward* content-neutrality three-part test. *See supra* note 108 for an explanation of the test factors.

213. *See Hill*, 530 U.S. at 712.

214. Hill v. Thomas, 973 P.2d 1246, 1247-48 (Colo. 1999).

215. *See id.* at 1248.

216. *Id.* at 1253.

217. *See id.*

so, the court identified two distinctions between the instant case and *Schenck*: the differing nature of the regulations and the additional scienter element added to the Colorado statute.[218] First, since the Colorado statute was just that, a legislative-enacted statute, it differed considerably from the injunctive judicial decree in *Schenck*, and the tests justifiably should be different for the two.[219]

Secondly, the "knowingly approaches" prerequisite present in the Colorado statute requires that a protestor intentionally confront a patient with the intention of educating that person in some manner – this requirement allaying fears that unintentional encounters would also be punished alike.[220] The statute "allows a protester to stand still while a person moving towards or away from a health care facility walks past her,"[221] even if that person comes within the eight foot bubble.[222]

The court concluded that the statute represented "a fair legislative balancing of the 'right to protest or counsel against certain medical procedures' while protecting a 'person's right to obtain medical counseling and treatment,'"[223] and through its comprehensive coverage of all health care facilities, was content-neutral, narrowly tailored to further a significant governmental interest, and left open ample alternative channels of communication.[224]

In 1999, the United States Supreme Court granted certiorari "[b]ecause of the importance of the case,"[225] and ultimately affirmed the lower court decision upholding the constitutionality of the Colorado statute.[226] The

---

218. *See Hill*, 530 U.S. at 713.

219. *See Hill*, 973 P.2d at 1255 (discussing the differences between the two government regulations).

> *Schenck* involved a judicially created preliminary injunction drawn solely for the parties before the Court under the circumstances then existing. Here, to the contrary, section 18-9-122(3) is not the creature of our judiciary, but, instead is a statute created by a coordinate branch of government, and is a rule of general application representing the public policy choices of the General Assembly of the State of Colorado.

Id.

220. *See Hill*, 973 P.2d at 1255.

221. *Hill*, 530 U.S. at 713.

222. The court also discussed the two basic requirements of a criminal statute, *actus reus* and *mens rea*, as applied to the "knowingly approach" criteria of the statute. *See Hill*, 973 P.2d at 1257.

223. *Id.* at 1259.

224. *See Hill*, 530 U.S. at 713-14.

225. *Id.*

226. *See id.*

Court again examined the "competing interests at stake,"[227] concluding that "each [had] legitimate and important concerns."[228] The balancing involved the freedom to communicate and "right to persuade" against the right to privacy and "right to be let alone."[229] The Court also found it important to distinguish between "state restrictions on a speaker's right to address a willing audience and those that protect listeners from unwanted communication."[230]

The Court evaluated the different challenges to the constitutionality of the statute: content-neutrality; valid time, place and manner regulation (including the requirements of being narrowly tailored, serving a significant government interest, and leaving open ample alternative channels of communication); overbreadth; vagueness; and a prior restraint.[231] The tantamount issue brought before the Court was "whether the First Amendment rights of the speaker are abridged by the protection the statute provides for the unwilling listener."[232]

The Court concluded that the Colorado statute passed the content-neutrality test on three independent bases:

> First, it is not a 'regulation of speech.' Rather it is a regulation of the places where some speech may occur. Second, it was not adopted 'because of disagreement with the message it conveys.' This conclusion is supported . . . by the State Supreme Court's unequivocal holding that the statute's restrictions apply equally to all demonstrators, regardless of viewpoint . . . . Third, the State's interest in protecting access and privacy, and providing the police with clear guidelines, are unrelated to the content of the demonstrator's speech.[233]

In affirming that the statute was content-neutral, the Court also discussed the situation of a pro-abortion demonstrator within the eight-foot bubble surrounding a patient, stating:

> [T]he relevant First Amendment point is that the statute would prevent both

---

227. *Id.* The Court noted that "both the petitioners' First Amendment interests and the state's interests in protecting access to clinics, safeguarding patient health, and evenhandedly applying the law were legitimate and important." *The Supreme Court 1999 Term Leading Cases: I. Constitutional Law - F. Freedom of Speech and Expression*, 114 HARV. L. REV. 279, 291 (2000).

228. *Hill*, 530 U.S. at 714.

229. *See id.* at 716-17.

230. *Id.* at 715-16.

231. *See Hill*, 530 U.S. at 719-34.

232. *Hill*, 530 U.S. at 708.

233. *Id.* at 719-20.

> speakers, unless welcome, from entering the 8-foot zone. The statute is not limited to those who oppose abortion . . . . It applies to all 'protest,' to all 'counseling,' and to all demonstrators whether or not the demonstration concerns abortion, and whether they oppose or support the woman who has made an abortion decision. That is the level of neutrality that the Constitution demands.[234]

The Court affirmed the conclusion of the Colorado courts that the statute was content-neutral.[235]

In evaluating the *Ward* time, place, and manner criteria, the Court found the regulation was narrowly tailored to serve significant legitimate government interests,[236] and that although "it is not the least restrictive or least intrusive means of serving the statutory goal,"[237] "it leaves open ample alternative channels of communication."[238] The Court discussed the various types of communication regulated by the statute (display of signs and placards, leafleting and oral speech) and concluded that "[i]t is . . . not clear that the statute's restrictions will necessarily impede, rather than assist, the speaker's efforts to communicate their messages."[239] The 8-foot distance still allows the signs and placards to be seen clearly, and a normal level conversation can be had at this distance with any willing listener.[240]

The next challenge by petitioners was that the statute was overly broad[241] for a couple of reasons. First, petitioners argued that "it protects too many people in too many places, rather than just the patients at the facilities where confrontational speech had occurred."[242] The Court actually turned this argument around to support that the statute was content-neutral and did not discriminate since it applied to all health care facilities in the state, not just to abortion clinics.[243] The second reason that petition-

---

234. *Id.* at 725.

235. *See id.*

236. *See id.* The Court discussed the need for tranquil surroundings at hospitals due to the potential of patients in "particularly vulnerable physical and emotional conditions" arriving for treatment. *Id.* at 729.

237. *Hill*, 530 U.S. at 726.

238. *Id.*

239. *Id.* at 727.

240. *See id.* The one item the Court mentioned could be impeded by the distance was the leafleting. However, it noted that "[t]he statute does not . . . prevent a leafletter from simply standing near the path of oncoming pedestrians and proffering his or her material, which the pedestrians can easily accept . . . [or] be free to decline the tender." *Id.*

241. *See Hill*, 530 U.S. at 730.

242. *Id.*

243. *See id.* The Court quoted from *Railway Express Agency, Inc. v. New York*, stating "there is no more effective practical guaranty against arbitrary and unreasonable

ers argued the statute was overbroad was because it "bans virtually the universe of protected expression, including displays of signs, distribution of literature, and mere verbal statements."[244] The Court clarified that the statute does not actually 'ban' any messages, whether they be oral, signs or literature, "[i]t merely regulates the places where communications may occur"[245] and that petitioners' conduct was "within the statute's 'legitimate sweep.'"[246]

As for the vagueness challenge, the petitioners argued that some specific language within the statute lacked clarity and was therefore unconstitutionally vague.[247] The Court explained that "[a] statute can be impermissibly vague for either of two independent reasons. First, if it fails to provide people of ordinary intelligence a reasonable opportunity to understand what conduct it prohibits. Second, if it authorizes or even encourages arbitrary and discriminatory enforcement."[248] In refuting petitioner's claim, the Court discussed the elementary language used in the statute, such as 'knowingly,' 'consent,' 'protest,' 'education,' and 'counseling' stating that "[t]he likelihood that anyone would not understand any of those common words seems quite remote."[249] The Court then stated that it was "unpersuaded by the suggestion that § 18-9-122(3) fails to give adequate guidance to law enforcement authorities . . . [as] one of the section's virtues is the specificity of the definitions of the zones described in the statute."[250]

The final challenge by petitioners concerned the statute imposing "an unconstitutional 'prior restraint' on speech,"[251] an argument which the Court had previously rejected in the *Schenck* and *Madsen* decisions.[252] The Court answered this challenge, stating "[u]nder this statute, absolutely no channel of communication is foreclosed. No speaker is silenced. And no message is prohibited . . . [allowing] every speaker to engage freely in any

---

government than to require that the principles of law which officials would impose upon a minority must be imposed generally." Railway Express Agency, Inc. v. New York, 336 U.S. 106, 112 (1949).

244. *Hill*, 530 U.S. at 730.

245. *Id.* at 731.

246. *Id.* at 732.

247. *See Hill*, 530 U.S. at 732.

248. *Id.*

249. *Id.*

250. *Id.* at 733.

251. *Id.* "A prior restraint is 'an administrative system or judicial order that prevents speech from occurring.'" Pettinari, *supra* note 4, at 831.

252. *See Hill*, 530 U.S. at 733-34. The *Madsen* court held that "the injunction did not prevent protestors from expressing their message, only from expressing their message within the buffer zone." Pettinari, *supra* note 4, at 832. Similarly, in *Schenck*, the Court held that "a prior restraint does not exist if 'alternate channels of communication [are] left open.'" *Id.*

expressive activity communicating all messages and viewpoints subject only to the narrow place requirement imbedded within the 'approach' provision."[253] This provision of the statute "empowers private citizens entering a health care facility with the ability to prevent a speaker, who is within eight feet and advancing, from communicating a message they do not wish to hear."[254]

Though there were strong dissenting opinions from Justice Scalia and Justice Kennedy,[255] the majority of the United States Supreme Court affirmed the judgment of the Colorado Supreme Court,[256] holding for the first time that a floating buffer zone was constitutional, at least under the circumstances and parameters of the *Hill* case.[257] As for the future of floating buffer zone constitutionality, there is yet another jurisdiction wrestling with similar constitutional issues – Massachusetts. The next section of this Note will explore the current status of the newly enacted state law, the factual and procedural history as to how it got where it is today, and where it is expected to go in the foreseeable future.

III. MASSACHUSETTS JOINS IN THE BUFFER ZONE DEBATE

A. Factual History of the Massachusetts Law

At the same time the *Hill* case was escalating through the Colorado court system, the Commonwealth of Massachusetts was debating similar legislation in its own jurisdiction.[258] In 1993 Massachusetts enacted section 120E of Chapter 266 of the Massachusetts General Laws with the purpose of prohibiting the obstruction of access to medical facilities, and to punish those who disobeyed as a criminal offense.[259] This law provided only for remedies for physical obstructions outside of the clinics, but did not solve the problem of emotional obstructions that were caused by speech-related conduct.[260]

---

253. *Hill*, 530 U.S. at 734.

254. *Id.*

255. *See generally Hill*, 530 U.S. at 741-92 (Scalia, J. & Kennedy, J., dissenting).

256. *See id.* at 735.

257. *See id.* at 713. "[E]ven though a 15-foot floating buffer [such as that which was struck down in the *Schenck* case] might preclude protesters from expressing their views from a normal conversational distance, a lesser distance of eight feet was sufficient to protect such speech on a public sidewalk." *Id.* at 712.

258. *See infra* notes 265-84 and accompanying text.

259. *See* MASS. GEN. LAWS ch. 266, § 120E (2000) (effective Nov. 4, 1993).

260. *See id.* The law prohibits only physical obstruction of clinic entrances or imposition to medical procedures inside the facility, but does not address activities which occur outside of the clinics involving the patients entering or leaving the facili-

On December 30, 1994 gunman John Salvi went on a killing spree at two health clinics in Brookline, Massachusetts, killing two abortion clinic staff members and wounding five others.[261] Three years later to the day, Massachusetts filed legislation creating a 25-foot fixed buffer zone around abortion clinics and family-planning health care facilities.[262] The prior clinic access laws "[provided] relief through court injunctions only after a clinic has already been subject to actual incidents or threats and are limited to specific individuals or sites. The buffer zone bill would do more to prevent disturbances in the first place by keeping protesters at a safe distance."[263]

The legislation originally proposed in Senate No. 148 called for a 25-foot fixed buffer zone surrounding any entrance, exit or driveway of a reproductive health care facility, as well as a rectangular corridor zone extending from the boundaries of these locations to the intersection of each with the sidewalk in front of the clinic.[264] Any person who knowingly entered or remained in these buffer-zone areas would be punished criminally and/or civilly.[265] Exempted from enforcement of this provision were patients, clinic employees, law enforcement and other municipal agents acting within the scope of their employment.[266]

The basic purposes expressed for creating this statute were to provide patients with access to reproductive care free from harassment and intimidation, as well as violence, and increase the public safety of patients and clinic staff alike, while still allowing lawful and peaceful demonstrating and leafleting within specified areas.[267] As stated by the Planned Parenthood League of Massachusetts during the hearings on the legislation, "[t]his is not an abortion issue. It is about law, order, public health and

---

ties prior to or after receiving medical assistance. As seen from the testimony of patients and medical staff in the previously discussed cases, the initial decision to go to an abortion clinic is a very emotional decision, and the patients should not be forced to experience increased stress and harassment as a result of their choice.

261. *See* Editorial, *A Buffer Zone For Women's Clinics*, THE BOSTON GLOBE, Dec. 30, 1997, at A14. The two women killed were Shannon Lowney and Lee Ann Nichols. *See id.* Salvi was convicted on two counts of murder and sentenced to life in prison, only to commit suicide in his cell almost two years after the killings, in November 1996. *See* Ed Hayward, *Activists Gear Up To Debate Proposed Clinic Buffer Zones*, THE BOSTON HERALD, July 6, 1998, at 8.

262. *See* Editorial, *supra* note 261, at A14.

263. *Id.* The Boston Globe also stated that "[t]he buffer zone law would provide an additional deterrent to clinic violence." *Id.*

264. *See* Opinion of the Justices to the Senate, 430 Mass. 1205, 1206 (2000).

265. *See id.* at n.2.

266. *See id.*

267. *See* Planned Parenthood League of Massachusetts, *Facts on Buffer Zone Legislation (s. 148)*, *at* http://www.pplm.org./bzfacts.html (visited Feb. 11, 2001).

safety, individual liberties and civil behavior. The Commonwealth has an obligation to protect all of these."[268]

The Massachusetts Senate brought the statute to the state's highest court, the Massachusetts Supreme Judicial Court (SJC), for an advisory opinion as to the constitutionality of the proposed statute prior to its enactment into law.[269] The issue before the SJC was whether "Senate No. 148, by restricting access to buffer zones outside reproductive health care facilities, violate[d] the right of freedom of speech or the right of the people peaceably to assemble as provided by the First Amendment to the Constitution of the United States . . . ."[270]

The SJC ruled that the proposed bill did not violate the First Amendment freedom of speech rights of the demonstrators.[271] The SJC evaluated the statute on the basis of the *Ward* criteria: content-neutrality, narrowly tailored to serve a substantial government interest and leaving open ample alternative means of communication – on each of these issues the SJC found the criteria was satisfied.[272] On the issue of content-neutrality, the SJC determined that since the "buffer zone applies regardless of political viewpoint, Senate No. 148 is a content-neutral statute."[273] The SJC concluded there were four significant, legitimate interests being advanced by the government, to which the statute was narrowly tailored:

> (1) to increase the public safety in and around reproductive health care facilities; (2) to maintain the flow of traffic and prevent congestion around reproductive health care facilities; (3) to enact reasonable time, place, and manner restrictions to reconcile and protect both the First Amendment rights of persons to express their views near reproductive health care facilities and the rights of persons seeking access to those facilities to be free from hindrance, harassment, intimidation, and harm; and (4) to create an environment in and around reproductive health care facilities which is conducive toward providing safe and effective medical services, including surgical procedures, to its patients.[274]

As for the alternative means of communication, the SJC determined that the protestors "may still engage in all forms of protest as they have previously done, but are simply constrained to do so outside the buffer zone."[275]

---

268. *Id.*

269. *See* Opinion of the Justices to the Senate, 430 Mass. at 1205.

270. *Id.* at 1207.

271. *See id.* at 1212. The SJC concluded, "[t]he answer to the question is, 'No.'" *Id.*

272. *See id.* at 1209-12.

273. Opinion of the Justices to the Senate, 430 Mass. at 1209.

274. *Id.* at 1210.

275. *Id.* at 1212.

The final point the SJC discussed was in relation to the fixed buffer zone itself, holding that "[t]he 'fixed' clinic buffer zone at issue in Senate No. 148 is substantially comparable to other 'fixed' clinic buffer zones that either we or the Federal courts have previously upheld under either the *Ward* standard or the heightened standard applicable to injunctions."[276] The SJC referenced the decision in *Schenck v. Pro-Choice Network of W.N.Y.*,[277] reiterating that the Supreme Court stated that "floating buffer zones are problematic because, by their nature, they are difficult for a protester to comply with, and 'this lack of certainty leads to a substantial risk that much more speech will be burdened.' There is no such risk inherent in the fixed buffer zone that would be established"[278] by the Massachusetts law. It is important to realize that the conclusions made by the SJC were based on the language of the statute as originally proposed (as a fixed buffer zone only), and not the language ultimately utilized in the enacted law.[279]

After many years of debate, in November, 2000 Massachusetts finally enacted a statute establishing both a fixed and a floating buffer zone outside of abortion clinics.[280] In the interim period, the bill had been amended

---

276. *Id.* at 1211.

277. 519 U.S. 357 (1997).

278. Opinion of the Justices to the Senate, 430 Mass. at 1211.

279. *See* McGuire v. Reilly, 122 F. Supp.2d 97, 100-01 n.4 (D. Mass. 2000). "[W]hile the Supreme Judicial Court decision informs this Court, it would not be fair to presume that the Supreme Judicial Court has announced its interpretation of this statute in its current form." *Id.*

280. *See* MASS. GEN. LAWS ch. 266, § 120E½ (2000). The pertinent part of the statute stating:

> (a) For the purposes of this section, "reproductive health care facility" means a place, other than a hospital, where abortions are offered or performed.
> (b) No person shall knowingly approach another person or occupied motor vehicle within six feet of such person or vehicle, unless such other person or occupant of the vehicle consents, for the purpose of passing a leaflet or handbill to, displaying a sign to, or engaging in oral protest, education or counseling with such other person in the public way or sidewalk area within a radius of 18 feet from any entrance door or driveway to a reproductive health care facility or within the area within a rectangle not greater than six feet in width created by extending the outside boundaries of any entrance door or driveway to a reproductive health care facility at a right angle and in straight lines to the point where such lines intersect the sideline of the street in front of such entrance door or driveway. This subsection shall not apply to the following:
> (1) persons entering or leaving such facility;
> (2) employees or agents of such facility acting within the scope of their employment;
> (3) law enforcement, ambulance, firefighting, construction, utilities, public works and other municipal agents acting within the scope

from its original form proposed in Senate No. 148 in response to the *Hill v. Colorado* decision.[281] The amended bill, House No. 5401, mirrored the Colorado statute in essentially all aspects, while still maintaining the exemptions originally included in Senate No. 148.[282] Specifically, the language of the enacted statute reduced the fixed buffer zone from twenty-five feet to eighteen feet, the rectangular corridor zone was reduced to a maximum width of six feet, and a new floating buffer zone of six feet around patients and their vehicles was created.[283]

B. Procedural History of the Massachusetts Law

Although the statute was deemed constitutional by the Massachusetts Supreme Judicial Court advisory opinion, prior to its enactment, a lawsuit challenging the statute's constitutionality was brought in federal court shortly after the law went into effect.[284] The facts of this case were quite similar to those in *Hill v. Colorado*, the plaintiffs being three 'sidewalk counselors' bringing a facial challenge to the constitutionality of the statute, seeking both preliminary and permanent injunctions against its enforcement.[285] Plaintiffs claimed that "their fear of criminal prosecution caused them to be chilled in their exercise of fundamental constitutional rights."[286] The issue to be decided by the Federal District Court was four-part: (1) "whether the Act's explicit singling out of reproductive health care facilities indicates that it is a content-based regulation of speech;"[287] (2) "whether the exemption for certain people, particularly employees and agents of the facilities, represents impermissible governmental protection for one side of the abortion debate, while abridging fundamental free

---

of their employment; and
(4) persons using the public sidewalk or street right-of-way adjacent to such facility solely for the purpose of reaching a destination other than such facility.

Id.

281. *See* Brief for Appellant at 7, McGuire v. Reilly, 260 F.3d 36 (1st Cir. 2001) (No. 00-2492).

282. *See id.* at 7.

283. *See* MASS. GEN. LAWS ch. 266, § 120E½ (2000).

284. *See* McGuire v. Reilly, 122 F. Supp.2d 97 (D. Mass. 2000). *See also* Shelley Murphy, *US Judge Rejects Abortion Clinic Law Faults Mass. Rule on Buffer Zones*, THE BOSTON GLOBE, Nov. 21, 2000, at A1.

285. *See McGuire*, 122 F. Supp.2d at 100.

286. *Id.* at 99.

287. *Id.*

speech rights of the other;"[288] (3) "whether the process for physically marking the buffer zones amounts to a discriminatory activation provision;"[289] and (4) "whether the language of the statute, aimed primarily at oral communications, reveals that its true intent is the suppression of speech rather than any of the four stated purposes."[290]

The Federal District Court barred enforcement of the new law, striking it down as being unconstitutional.[291] The district court concluded that the statute was content-based, not content-neutral, due to its application specifically to abortion clinics instead of all health care facilities, and the inclusion of exemptions for clinic employees and staff acting within the scope of their employment.[292] The district court stated that "the statute is a regulation of speech and the content of the only speech regulated is the subject of abortion."[293] Also, the district court discussed how the people who are exempted from the statute are "individuals, [who] because of their personal relationship with the abortion clinic, have a strong financial interest or philosophic incentive to counsel the listener to undergo an abortion and they constitute very zealous advocates for this controversial procedure."[294] The statute was ruled unconstitutional, as being violative of the First Amendment, and a preliminary injunction enjoining its enforcement was ordered.[295]

The order was thereafter lifted by the Court of Appeals for the First Circuit and enforcement of the law, pending appeal proceedings on its constitutionality, was allowed.[296] A press release issued by the Office of Attorney General Tom Reilly on December 22, 2000, indicated that in the latest ruling by the First Circuit, the court concluded that the Attorney General had "demonstrated a strong likelihood of success on the merits of his appeal."[297] The reasoning behind lifting the stay of injunction, accord-

---

288. *Id.*

289. *Id.*

290. *Id.*

291. *See McGuire*, 122 F. Supp.2d at 104. *See also* Murphy, *supra* note 285 at A1. The Boston Globe reported that "the law doesn't pass constitutional muster because it allows employees and agents of the clinics, such as volunteers, to approach patients outside clinics, but prohibits protesters from coming within 6 feet of patients who are within 18 feet of a clinic entrance without their consent." *Id.*

292. *See* Brief for Appellant at 16, McGuire v. Reilly, 260 F.3d 36 (1st Cir. 2001) (No. 00-2492).

293. *McGuire*, 122 F. Supp.2d at 102.

294. *Id.* at 103.

295. *See id.* at 104.

296. *See* Shelley Murphy, *Court Gives Go Ahead on Buffer Zones, State to Enforce Its New Law For Abortion Clinics*, THE BOSTON GLOBE, Dec. 23, 2000, at A1.

297. Ann Donlan, *US Circuit Court of Appeals Lifts Stay of Injunction That Stopped Enforcement of the Buffer Zone Law, at* http://www.ago.state.ma.us/buffer.asp

ing to the circuit court judge, was that "the harm to public safety from barring enforcement of the Buffer Zone law outweighed the 'potential effect' of enforcement on First Amendment rights."[298] The appellate briefs were submitted to the United States Court of Appeals on January 23, 2001, pending the scheduling of appeal proceedings.[299]

C. Application of Hill v. Colorado - Why the 1st Circuit Court of Appeals Should Reverse the District Court Ruling

Since the case has not yet gone to the United States Court of Appeals, or ultimately the United States Supreme Court for final decision, the Massachusetts statute and its case history so far can be analogized to the *Hill* case for a better understanding of what will likely happen once it does reach the higher courts. In examining the language of the two statutes, it is clear that the Massachusetts statute, as amended, is a virtual replica of the Colorado statute and should be afforded the same level of constitutionality.[300]

1. The Statutory Language

The two laws contain the same opening language of "no person shall knowingly approach another person within" a specified number of feet of the person – Colorado inserting eight feet as its specified number of feet, and Massachusetts reducing that to six feet.[301] The 'knowingly approach' provision of these statutes is one thing that differentiates them from the injunctions in *Madsen* and *Schenck* – not only must the protester intentionally approach the patient, he or she must undertake that action with the understanding that he or she is breaking the law and subject to criminal consequences.[302] Massachusetts went one step further than Colorado in

---

(visited Feb. 19, 2001).

298. *Id.*

299. *See* J.M. Lawrence, *Mass. Files Appeal to Uphold Clinic Buffer Zone Law*, THE BOSTON HERALD, Jan. 24, 2001, at 10. The Boston Herald reports that this reinstatement implies that "the statute is likely to win constitutional approval." *Id.*

300. *See* Brief for Appellant at 12-13, McGuire v. Reilly, 260 F.3d 36 (1st Cir. 2001) (No. 00-2492). "The District Court reviewed the legislative history of the Act and specifically concluded that the legislature consciously modeled the Act after the Colorado law." *Id.* at 13 n.3.

301. *Id.* at 13. *See also* COLO. REV. STAT. § 18-9-122 (1999) and MASS. GEN. LAWS ch. 266, § 120E½ (2000).

302. *See supra* notes 221-23 and accompanying text. "[A] petitioner who does not act to approach a person going into or coming out of a health care facility does not

adding "occupied motor vehicle" to the unapproachable zone of a person.[303] Both statutes contain the pre-approach consent requirement, and the consent is required for the same actions, "passing a leaflet or handbill to, displaying a sign to, or engaging in oral protest, education, or counseling with such other person in the public way or sidewalk area."[304]

The radii around the facility differed between the two laws dramatically in the number of feet specified: Colorado allowed for a radius of one hundred feet, while Massachusetts granted only eighteen feet.[305] The locations of the radii differed between the two statutes: Colorado only indicated the radius stemmed from the entranceway to the facility, while Massachusetts also included the facility's driveway.[306] Another distinction between the locations was that the Colorado statute applied to all health care facilities, while the Massachusetts statute narrowed its scope to apply only to reproductive health care facilities.[307] The final distinction between the two is that the Massachusetts statute added exemptions, which the Colorado statute does not provide for, "passers-by, and for clinic workers and others working within the scope of their employment."[308]

2. Content-neutrality

As discussed previously in the analyses of other abortion protest cases, content-neutrality is a threshold question for any court to ask in determining if a statute is constitutional, and under what level of scrutiny it should

---

violate section 18-9-122(3), even if an individual comes within eight feet of him or her, as long as the petitioner does not 'approach' the individual." Hill v. Thomas, 973 P.2d 1246, 1255 (Colo. 1999).

303. Brief for Appellant at 13, McGuire v. Reilly, 260 F.3d 36 (1st Cir. 2001) (No. 00-2492). *See also* COLO. REV. STAT. § 18-9-122 (1999) and MASS. GEN. LAWS ch. 266, § 120E½ (2000).

304. Brief for Appellant at 13, McGuire v. Reilly, 260 F.3d 36 (1st Cir. 2001) (No. 00-2492). *See also* COLO. REV. STAT. § 18-9-122 (1999) and MASS. GEN. LAWS ch. 266, § 120E½ (2000).

305. *See* Brief for Appellant at 13, McGuire v. Reilly, 260 F.3d 36 (1st Cir. 2001) (No. 00-2492). *See also* COLO. REV. STAT. § 18-9-122 (1999) and MASS. GEN. LAWS ch. 266, § 120E½ (2000).

306. *See* Brief for Appellant at 13, McGuire v. Reilly, 260 F.3d 36 (1st Cir. 2001) (No. 00-2492). *See also* COLO. REV. STAT. § 18-9-122 (1999) and MASS. GEN. LAWS ch. 266, § 120E½ (2000).

307. *See* Brief for Appellant at 13, McGuire v. Reilly, 260 F.3d 36 (1st Cir. 2001) (No. 00-2492). *See also* COLO. REV. STAT. § 18-9-122 (1999) and MASS. GEN. LAWS ch. 266, § 120E½ (2000).

308. Brief for Appellant at 14, McGuire v. Reilly, 260 F.3d 36 (1st Cir. 2001) (No. 00-2492). *See also* COLO. REV. STAT. § 18-9-122 (1999) and MASS. GEN. LAWS ch. 266, § 120E½ (2000).

be evaluated.[309] Content-neutrality is based on a showing that the regulation is not motivated by seeking to suppress a specific viewpoint or subject matter of speech.[310] Like the Colorado statute, the Massachusetts law is not a regulation of the type of speech that may be regulated, rather a restriction on the location where this speech may occur.[311] The United States Supreme Court evidenced this notion in its decision in *Frisby v. Schultz*,[312] where it upheld an anti-picketing ordinance which "was obviously enacted in response to the activities of anti-abortion protesters who wanted to protest at the home of a particular doctor to persuade him and others that they viewed his practice of performing abortions to be murder."[313]

The district court in *McGuire* held that the Massachusetts statute was content-based, not content-neutral and founded this conclusion by looking at both the text of the statute and its underlying purposes.[314] Though the textual differences between the two state statutes were slight, with the exception of the few additions in the Massachusetts statute, the court concluded that these were severe enough to convert a content-neutral statute into a content-based one.[315] The court erred in this conclusion; "the Supreme Court's decision in *Hill* mandates the conclusion that the Act's [sic] is content-neutral in its ban on non-consensual approaches within six feet

---

309. Content-neutral and content-based regulations are subject to different levels of scrutiny by the courts. "If a regulation or an order is determined to be content-neutral, it is subjected to review under a judicially created standard of scrutiny . . . [which] is an intermediate level of scrutiny," as seen by the elements introduced in the *Ward* case. Clay Calvert, *Free Speech and Content-Neutrality: Inconsistent Applications of an Increasingly Malleable Doctrine*, 29 MCGEORGE L. REV. 69, 74-75 (1997). Whereas, when a court finds a regulation to be content-based, it "applies 'the most exacting scrutiny to regulations that suppress, disadvantage, or impose differential burdens upon speech because of its content' . . . . This test is often referred to as a strict scrutiny standard of review." *Id.* at 76.

310. *See* Brief for Appellant at 20, McGuire v. Reilly, 260 F.3d 36 (1st Cir. 2001) (No. 00-2492). *See also* discussion *supra* note 101 discussing fundamentals of content-neutrality.

311. *See* Brief for Appellant at 20, McGuire v. Reilly, 260 F.3d 36(1st Cir. 2001) (No. 00-2492). *See also* Hill v. Colorado, 530 U.S. 703, 717-18 (2000) (finding the "statutory language makes no reference to the content of the speech"); City of Erie v. Pap's A.M., 529 U.S. 277, 290 (2000) (finding ordinance banning public nudity was content-neutral where, "[b]y its terms, the ordinance regulates conduct alone").

312. 487 U.S. 474 (1988).

313. *Hill*, 530 U.S. at 725 (refuting the "contention that a statute is 'viewpoint based' simply because its enactment was motivated by the conduct of the partisans on one side of a debate").

314. *See McGuire*, 122 F. Supp. 2d at 99.

315. *See id.* at 101-02 (discussing the differences between the two statutes and the implications of such on the content-neutrality of each).

heart [sic] for the purpose of protest, education and counseling."[316] Because the Massachusetts law applies to specific locations (only abortion clinics, not all medical facilities) and lists exemptions of specific persons acting within the scope of their employment (medical staff or civil service workers), these differences are not of great enough magnitude to convert a content-neutral regulation into a content-based one.[317] In fact, these differences serve to tailor the regulation even more closely to the governmental interests being served.[318]

3. Narrowly Tailored to Significant Government Interests

The court also erroneously concluded that the purposes behind the statute applying only to abortion clinics and exempting those persons with an affiliation to those clinics supported the finding that the statute was content-based, not content-neutral.[319] The court expressed its concern that these provisions were based on and advocated discrimination in favor of the pro-abortion viewpoints.[320] Even if that were an effect caused by the statute's enactment, whether intentional or not, "where . . . a statute is clearly facially content-neutral, the court will rarely, if ever, strike it down based on a discriminatory governmental purpose."[321] The district court, in concluding that the Massachusetts law was intentionally focusing on speech related to abortion protest, rather than the demonstrators' conduct, noted that "[i]t is . . . only speech, and not acts of physical violence, that is regulated by a statute whose purpose is ostensibly public safety."[322] What the District Court did not express, however, is the presence of other laws already in existence that prohibit physical acts of violence outside of the

316. Brief for Appellant at 25, McGuire v. Reilly, 260 F.3d 36 (1st Cir. 2001) (No. 00-2492) (quoting *Hill*, 530 U.S. at 723, stating the identical ban in *Hill* regarding non-consensual approaches within eight feet amounted to only a "minor place restriction on an extremely broad category of communications with unwilling listeners").

317. *See id.* at 25-26.

318. *See infra* notes 320-32 and accompanying text.

319. *See McGuire*, 122 F. Supp. 2d at 103. The court specifically mentioned how "the Colorado statute in *Hill* allowed no exemptions from its restrictions on speech" and stated that "[f]or the Massachusetts statute to pass constitutional muster, this exemption, at the very least, must be stricken from the statute's provisions." *Id.*

320. *See id.*

321. Brief for Appellant at 21 n.3, McGuire v. Reilly, 260 F.3d 36 (1st Cir. 2001) (No. 00-2492) (quoting City of Erie v. Pap's A.M., 529 U.S. at 292 ("[T]his Court will not strike down an otherwise constitutional statute on the basis of an alleged illicit motive.")).

322. Brief for Appellant at 34, McGuire v. Reilly, 260 F.3d 36 (1st Cir. 2001) (No. 00-2492).

clinics, which are not covered by the First Amendment.[323] The need to fill the gap left open by the clinic access laws was the reason that the buffer zone laws were devised; "[t]he Act was a supplement and complement to complete the protection provided by other public safety laws."[324]

The purposes for creating and enacting the Colorado and Massachusetts statutes are significant government interests, which are legitimately served through their enforcement. The interests are basically the same: protection for those who seek medical treatment or counseling services through unobstructed access to health care facilities, both physically and emotionally, while maintaining public safety and the free flow of traffic outside the facilities.[325] Colorado's view was broader and covered all medical facilities and all patients, while Massachusetts focused on reproductive health care facilities and their patients in response to abortion clinic protest and violence.[326]

A statute is considered to be narrowly tailored "if it targets and eliminates no more than the exact source of the 'evil' it seeks to remedy."[327] The fact that Massachusetts specifically tailored its law to apply only to the places where it had incidents requiring the law to be created further supports the significance and legitimacy of its purpose.[328] In addition, the

---

323. *See* Freedom of Access to Clinic Entrances Act (FACE), 18 U.S.C. § 248 (1994), which is a federal law enacted to "[prohibit] anyone from using 'force or threat of force or . . . physical obstruction' as a means of intentionally inflicting injury or intimidation or intentionally interfering with 'any person because that person is or has been . . . obtaining or providing reproductive health services.'" Kathryn D. Piele, *Sabelko v. City of Phoenix: Ninth Circuit Refuses to Burst "Bubble" Protecting Women Entering Health Care Facilities*, 75 OR. L. REV. 1297, 1305 (1996) (discussing the decision in *Sabelko* in light of the *Madsen* and *Schenck* decisions).

324. Brief for Appellant at 35, McGuire v. Reilly, 260 F.3d 36 (1st Cir. 2001) (No. 00-2492).

325. *See* COLO. REV. STAT. § 18-9-122 (2000) and MASS. GEN. LAWS ch. 266, § 120E½ (2000).

326. *See* COLO. REV. STAT. § 18-9-122 (2000) and MASS. GEN. LAWS ch. 266, § 120E½ (2000). In the SJC opinion upholding the constitutionality of the Massachusetts statute, the SJC "found that the explicit reference to reproductive health care facilities did not render the statute a content-based regulation . . . [but the court also] did not address the relevance of exemptions for certain groups of people." *McGuire*, 122 F. Supp. 2d at 101 n.4.

327. Pettinari, *supra* note 4, at 829 (quoting *Frisby*, 487 U.S. at 485).

328. *See, e.g.*, Planned Parenthood League of Mass., Inc., v. Bell, 424 Mass. 573, 574 (1997) (affirming an injunction against an abortion protestor who had dressed up as a clinic employee to gain access to harass patients inside and outside the clinic to remain outside a 50-foot buffer zone). *See also McGuire*, 122 F. Supp. 2d at 101. The court expressed defendants' argument:

> [D]efendants argue [] the Massachusetts legislature crafted the statute to

Colorado statute allowed for a 100-foot fixed buffer zone and eight-foot floating buffer zone, while Massachusetts permitted only an eighteen-foot fixed buffer zone and six-foot floating buffer zone; that in itself is an example of how Massachusetts is even more narrowly tailored than a statute which the U.S. Supreme Court has already upheld as constitutional.[329] Finally, while the Colorado statute is applicable at all times, the Massachusetts statute is only applicable during business hours and only if the buffer zone itself is "clearly marked and posted."[330] "[T]he Act's purposes are significant under the holdings of *Hill* and other Supreme Court cases, and the Act – which is drawn even more narrowly than the Colorado statute upheld in *Hill* – is narrowly tailored to those purposes,"[331] the end result of which should be that the Massachusetts statute, too, is upheld as constitutional.

4. Leaves Open Ample Alternative Channels of Communication

The buffer zone in the Massachusetts statute imposes a smaller area restriction on the speakers than the Colorado statute – six feet versus eight feet.[332] Since the Colorado statute was held by the U.S. Supreme Court to leave open ample alternative channels of communication due to the relatively small floating buffer zone, the Massachusetts statute should be in an even better position for passing constitutional muster.[333] Within a distance

---

be only as restrictive as necessary to serve its purposes. Whereas the Colorado statute applies to all medical facilities, the Massachusetts statute is more narrowly tailored and applies only to reproductive health care facilities. In that way, speech is regulated only where its exercise has given rise to a competing government interest.

Id.

329. *See* Brief for Appellant at 46, McGuire v. Reilly, 260 F.3d 36 (1st Cir. 2001) (No. 00-2492).

330. *Id.* at 46-47. The posting and markings of the buffer zones in Massachusetts were executed by the Boston Police, not the abortion clinic staffs, identifying the eighteen-foot zones in bright orange paint. *See* Kyla Reynolds, *At Abortion Clinics, Friction Resurfaces Buffer Zone Law Hangs in Limbo of the Courts*, BOSTON GLOBE, Dec. 17, 2000, at 1. *See also* Editorial, *Drawing the Line*, BOSTON GLOBE, Nov. 3, 2000, at A22.

331. Brief for Appellant at 16-17, McGuire v. Reilly, 260 F.3d 36 (1st Cir. 2001) (No. 00-2492). For examples of similar government interests from other abortion clinic buffer zone cases, *see Schenck*, 519 U.S. at 376; *Madsen*, 512 U.S. at 767-68; *Edwards*, 150 F.3d at 1216.

332. *See* COLO. REV. STAT. § 18-9-122 (2000) and MASS. GEN. LAWS ch. 266, § 120E½ (2000).

333. *See* Brief for Appellant at 47, McGuire v. Reilly, 260 F.3d 36 (1st Cir. 2001) (No. 00-2492). The buffer zones examined in previous cases were larger than the

of six feet, two people are able to carry on a conversation at normal speaking levels; within a distance of six feet, someone passing by another person holding up a sign or placard can easily make out the words or images on that sign, and probably more easily than if it was up close; within a distance of six feet, a person distributing leaflets can extend his arm to passers-by who may decide to stop and take a leaflet.[334] Outside of the eighteen-foot protected area, demonstrators can conduct themselves as they wish, as their conduct outside the buffer is not covered by the statute.[335] Since the Massachusetts statute burdens even less communication than the Colorado statute, which was upheld as constitutional, the Massachusetts statute should also pass constitutional muster.

5. Support From Surrounding Jurisdictions

As previously discussed, the decision in *Schenck* struck down floating buffer zone provisions in cases involving court-ordered injunctions as unconstitutional and burdening more speech than necessary to further the governmental interests.[336] The decision in *Hill* resolved the issue for legislative ordinances, finding floating buffer zones are constitutional as long as the proper test is satisfied.[337] Since *Schenck*, the courts have gone both one step further and one step back with regard to these two different situations.

Prior to the *Hill* decision, the Ninth Circuit Court of Appeals ruled a floating buffer zone provision of an Arizona statute to be unconstitutional.[338] The appeals court found that although the statute was content-neutral and served legitimate and substantial interests "in protecting a woman's freedom to seek lawful medical or counseling services in connection with her pregnancy,"[339] the statute was not narrowly tailored and therefore did not satisfy the "time, place, and manner" restrictions established in *Ward*.[340] The Colorado statute differed from the Phoenix

---

Colorado buffer zone, and were held to be unconstitutional. *See Schenck*, 519 U.S. at 377; *Sabelko*, 120 F.3d at 162-63; *Edwards*, 150 F.3d at 1215. *See also infra* notes 344-52 and accompanying text.

334. *See* Brief for Appellant at 47, McGuire v. Reilly, 260 F.3d 36 (1st Cir. 2001) (No. 00-2492).

335. *See id.* at 48.

336. *See supra* notes 147-83 and accompanying text.

337. *See supra* notes 186-258 and accompanying text.

338. *See* Sabelko v. City of Phoenix, 120 F.3d 161, 162 (9th Cir. 1997). The U.S. Supreme Court granted certiorari and remanded for consideration in light of the *Schenck* decision, similar to what occurred in *Hill*. *See* Miller, *supra* note 42, at 311.

339. *Sabelko*, 120 F.3d at 164 (quoting *Madsen*, 114 S. Ct. at 1175-76).

340. *See id.* at 165 n.3. The appeals court discussed how "the Phoenix ordinance

ordinance due to the "knowingly approach" requirement in section 18-9-122(3), which was not part of the Phoenix statute, and this was a contributing factor in the difference in decisions.[341] Since the Massachusetts statute follows so closely after the Colorado statute, the decision regarding the constitutionality of the Massachusetts statute should follow in the footsteps of *Hill*, leaving the open question in *Schenck* regarding floating buffer zones for cases involving ordinances finally resolved once and for all.[342]

The open question in *Schenck* regarding the constitutionality of floating buffer zones for cases involving court-ordered injunctions was also resolved at the appellate court level in 1999 in the case *U.S. v. Scott*.[343] This case involved a protestor who demonstrated outside abortion clinics in Connecticut and was sued by the United States and the state of Connecticut for a violation of the Freedom of Access to Clinic Entrances Act.[344] The district court then issued an injunction enjoining the protestor from demonstrating outside the clinics within fourteen feet of the clinic entrances and within five feet of any patients or occupied vehicles, which he repeatedly violated.[345] As a result of the magnitude of Scott's violations, the court then expanded the injunction increasing the fixed buffer zone from fourteen to twenty-eight feet, and the floating buffer zone from five to eight feet around any person or occupied vehicle of a patient at the clinic.[346] The Court of Appeals for the Second Circuit upheld the injunction and the floating buffer zone, distinguishing this case from *Schenck*, and emphasizing "*Schenck's* express refusal to declare all floating buffer zones unconstitutional per se."[347] The appeals court noted that the buffer zone in *Scott* differed from that in *Schenck* in two ways: (1) the floating buffer zone in *Schenck* (fifteen feet) was almost twice as large as the floating buffer zone in *Scott* (eight feet),[348] and (2) the buffer zone in *Scott*

---

suffers the same defects as the injunction in *Schenck*. It contains a broad prohibition on speech with which it is difficult to comply without risking a violation of the ordinance." *Id.*

341. *See* Pettinari, *supra* note 4, at 829-30; *see also Hill*, 973 P.2d at 1255.

342. *See Schenck*, 519 U.S. at 377. The Court stated, "We need not decide whether the governmental interests involved [between individuals] would ever justify some sort of zone of separation between individuals entering the clinics and protesters, measured by the distance between the two." *Id.*

343. 187 F.3d 282 (2d Cir. 1999).

344. *See id.* at 283. *See also* Piele, *supra* note 323 for explanation of FACE.

345. *See Scott*, 187 F.3d at 283.

346. *See id.* at 285-86.

347. Mark R. Kravitz, *Developments in the Second Circuit: 1998-1999*, 32 CONN. L. REV. 949, 961 (2000) (discussing "[t]his year's survey of decisions from the United States Court of Appeals for the Second Circuit").

348. *See id.*

has as its trigger a verbal refusal by a patient for Scott's literature or counseling, absent from the *Schenck* provision.[349] The floating buffer zone at hand in the Massachusetts statute is of comparable size and has a "consent" trigger for its applicability like the injunction in the *Scott* case.[350] Both are "sufficiently distinguishable from that in *Schenck* to survive constitutional challenge."[351] Therefore, the more tailored area of the buffer zone and the scienter element that triggers its operation, both lacking in *Schenck*, are sufficient for the Massachusetts and *Scott* regulations to pass constitutional muster.

IV. CONCLUSION

So what does the future hold? This Note has examined the various pieces of the past that have contributed to the present state and potential future state of abortion clinic buffer zone laws.[352] States have taken it upon themselves to extend their clinic access laws to also include restraints on demonstrations occurring outside their health care facilities so that patients will not have to "undertake Herculean efforts to escape the cacophony of political protests."[353]

As for Massachusetts, due to the continuing conflict between the courts on the issue of the statute's constitutionality, the *McGuire* case will certainly be heard before the United States Supreme Court for resolution; that resolution being that the Massachusetts statute enacted to regulate demonstration activities occurring outside of abortion clinics will pass constitutional muster.[354] Though different in language to some degree from the statute at issue in *Hill v. Colorado*, these distinctions are not grave enough to cause the Massachusetts law to be struck down as unconstitutional.[355] The intent is the same, the restrictions are the same, and therefore the result should also be the same.[356]

---

349. *See id.*

350. *Compare* MASS. GEN. LAWS ch. 266, § 120E½ (2000) *with* text accompanying notes 346-47.

351. Kravitz, *supra* note 347, at 961.

352. *See* discussion *supra* Parts II.A. notes 28-48, II.B.1-4 notes 49-83, II.C.1-3 notes 84-258, III.A. notes 259-84, III.B. notes 285-300, III.C.1-5 notes 301-52.

353. *Madsen*, 512 U.S. at 772-73.

354. *See supra* notes 301-52.

355. *See id.*

356. *See id.*

EPILOGUE

During the time that this article was being prepared for publication, a critical development regarding the future of buffer zones around abortion clinics in Massachusetts arose. On August 13, 2001, the United States Court of Appeals for the First Circuit in Boston upheld the constitutionality of the Massachusetts law creating the buffer zone around abortion clinics.[357] In reversing the lower court decision, the three-member panel of judges held that although the district court found "meaningful distinctions between the Act and the Colorado statute . . . [which the lower court judge ruled] undermined the constitutionality of the Act, and preliminarily enjoined the Act's enforcement,"[358] these distinctions "do not make a dispositive difference. *Hill* controls, and the Act, on its face, lawfully regulates the time, place, and manner of speech without discriminating based on content or viewpoint."[359]

The court first examined the First Amendment challenge on the basis of content-neutrality and disagreed with the district court's opinion that the Act "discriminates against abortion-related speech, and that the employee exemption compounds this evil by facilitating the airing of pro-choice sentiments while simultaneously restricting the expressive activities of pro-life partisans."[360] In examining the speech that is restricted by the legislature, the judges opined that "[a]lthough the Act clearly affects anti-abortion protesters more than other groups, there is no principled basis for assuming that this differential treatment results from a fundamental disagreement with the content of their expression."[361] The lower court, in emphasizing that the legislation applied "exclusively to speech communicated at abortion clinics and not . . . to *all health care facilities*,"[362] focused on the effects of the Act as opposed to the underlying purposes of the Act, which the Appeals Court determined to be a misinterpreted application of First Amendment law.[363] Based on the legislative history supporting the need for a solution to the serious public safety problems that had previously occurred at reproductive health care facilities (RHCFs), the statute's purposes of "promoting public health, preserving personal security, and

357. *See* McGuire v. Reilly, 260 F.3d 36 (1st Cir. 2001).

358. *Id.* at 39.

359. *Id. See also* J.M. Lawrence, *Appeals Court OKs Abortion Clinic Buffer Zones*, BOSTON HERALD, August 14, 2001, at 1.

360. *McGuire*, 260 F.3d at 44 (citation omitted).

361. *Id.*

362. *Id.* (emphasis in original).

363. *See id.*

affording safe access to medical services"[364] were justifiable for targeting RHCFs. In summary, "the legislature was making every effort to restrict as little speech as possible while combating the deleterious secondary effects of anti-abortion protests."[365]

The Appeals Court found numerous defects in the lower court's findings on the employee exception as a basis for defeating the statute's constitutionality. The first defect was in the district court's determination that the "statutory exemption for clinic agents and employees constitutes impermissible viewpoint-based discrimination."[366] The district court based its determination on the theory that, "by allowing clinic employees to enter the floating buffer zones without constraint, the Act permits free-ranging expression of pro-choice views while suppressing pro-life messages."[367] The Appeals Court wrote that this determination "[rested] on an unsubstantiated factual foundation"[368] and that, when unsupported by probative evidence, "[a] judge's intuition cannot take the place of proof."[369]

The second defect, relating to the district court's treatment of the statutory employee exemption, was found in the court's neglect to consider the secondary effects of the exception.[370] The Appeals Court stated that there was no evidence to support a finding that RHCF employees were responsible for past problems occurring at RHCFs, therefore, "excluding those individuals [through the employee exemption] does not undermine the legitimacy of the Act as a vehicle to curb the secondary effects of particular conduct and thereby achieve the legislature's announced purposes."[371] Even where a RHCF employee has had to approach a patient to protect her from a protester or physical altercation, those services are within the clinic personnel's scope of employment and do not constitute anything more than

---

364. *Id.*

365. *Id.*

366. *McGuire*, 260 F.3d at 45.

367. *Id.* *See also* Lawrence, *supra* note 360, at 1. According to Ray Neary, interim executive director of Massachusetts Citizens for Life, in a statement regarding the conduct of the abortion clinic escort employees, "I've seen them physically take hold of somebody going into a clinic and pull them away from even stopping to listen to the message of the pro-life protester." *Id.* Conversely, Dianne Luby, executive director of Planned Parenthood of Massachusetts, stated that the "escorts 'are simply trying to assist women' entering clinics" and "called the claims against escorts 'one more strategy' of abortion foes." *Id.*

368. *McGuire*, 260 F.3d 36, 45 (1st Cir. 2001).

369. *Id.*

370. *See id.* at 46 (the secondary effects being "securing public safety in and around RHCFs, preventing traffic congestion, and balancing free speech with the need to maintain a salutary atmosphere for those seeking access to medical services").

371. *Id.*

providing safe access to the facility.[372] The Appeals Court concluded this discussion in stating that "[b]ecause we can envision at least one legitimate reason for including the employee exemption in the Act, it would be premature to declare the Act unconstitutional for all purposes and in all applications."[373]

The court then proceeded to briefly examine the bases for the intermediate level scrutiny standard of review, which attaches to content-neutral time, place, and manner restrictions: "narrowly tailored to serve significant state interests while leaving open ample alternative channels of communication."[374] In holding that "[t]he Act passes this test,"[375] the court navigated through the components, proving each element along the way.[376] The court restated the purposes intended by the legislature, and stated how the "interests that underlie these purposes are firmly rooted in the states traditional police powers, and these are precisely the sort of interests that justify some incidental burdening of First Amendment rights."[377] As for narrow tailoring, the court establishes that this requirement is satisfied. The legislative history shows that the existing law at the time the legislation was enacted was inadequate to address the "public safety, personal security, traffic, and health care concerns created by persistent demonstrations outside RHCFs."[378] The legislature, in weighing the conclusions of the *Hill* decision, created a bill intended to pick up where the insufficient existing statutory provisions left off.[379] In utilizing the *Hill* statute to formulate the new Massachusetts legislation, the channels of communication in the Massachusetts law are actually more open than those in *Hill* due to the "Act's less commodious six-foot floating buffer zone."[380]

The next steps for counsel on each side of the issue have yet to be determined. Dwight G. Duncan, attorney for plaintiff Mary Anne McGuire, stated that "he and other lawyers would study the ruling before deciding

---

372. *See id. See also* Lawrence, *supra* note 360, at 1. Per Melissa Kogut, director of the National Abortion and Reproduction Rights Action League, "claims of clinic escorts advocating abortion are 'outrageous'" since "[e]scorts are not there to provide medical care or counseling. They're there to facilitate public safety." *Id.*

373. *McGuire*, 260 F.3d at 47.

374. *Id.* at 48. This standard and its components have been discussed thoroughly in the previous sections of this article.

375. *Id.*

376. *See id.*

377. *Id.* The court also referred back to similar legitimate interests in the aforementioned *Hill*, *Schenck*, and *Madsen* cases.

378. *Id.* at 49.

379. *See McGuire*, 260 F.3d at 49.

380. *Id.*

whether to pursue an appeal to the U.S. Supreme Court."[381] Massachusetts Attorney General Thomas Reilly, alternatively, stated that he "hopes anti-abortion protesters will reconsider further legal wrangling."[382] Only time will tell as to what the fate of the Massachusetts buffer zone law will be, but for today and the foreseeable future, floating buffer zones created to protect patients who are seeking medical services at reproductive health care facilities in Massachusetts from unwanted demonstrations are constitutional. If further legal action is taken, the U.S. Supreme Court has already favorably ruled on a practically identical statute from the state of Colorado, so it makes logical sense that the Court would also uphold the Massachusetts statute as passing constitutional muster.

Jessica Wainwright

381. Joe Spurr, US *Court Upholds Clinic Buffer Zone*, BOSTON GLOBE, August 14, 2001, at B2.

382. Lawrence, *supra* note 360, at 1.

Mandatory Rent Escrow: A Bill Relative to Rent Withholding

I. INTRODUCTION

The focus of this Note is the proposed amendment to Chapter 239, section 8A of the Massachusetts General Laws [hereinafter Chapter 239, 8A].[1] Chapter 239, 8A, as it currently reads, allows a tenant to unilaterally withhold rent for breach of the implied warranty of habitability (hereinafter the IWH).[2] The breach can then be raised as a defense to a summary process action for nonpayment of rent, provided that the landlord was put on notice of the violations before the tenant was in arrears with rent.[3] Landlord groups, mainly those purporting to represent small property owners, are aggressively lobbying the Massachusetts Legislature to amend the law to require that tenants deposit withheld rent into an escrow account before

---

1.*See* MASS. GEN. LAWS ch. 239, § 8A (2000). Chapter 239 is the chapter in the Massachusetts General Laws dealing with summary process for possession of land. In Massachusetts, the term "summary process" is used to describe the action for eviction of a tenant.

2.*See id. See also* ROBERT S. SCHOSHINSKI, AMERICAN LAW OF LANDLORD AND TENANT 123-24 (1980). Schoshinski defined the implied warranty of habitability [hereinafter the IWH] as such:

> Stripped to its essentials, the doctrine rejects the premise that a lease is only a transfer of land to be governed by special principles of real property law. Instead, the implied warranty of habitability theory treats the transfer of the demised premises as analogous to a sale of goods, and protects the expectations of the landlord and tenant by applying ordinary notions of contract law including mutual dependency of lease covenants.

Id.

3."Violations" refers to conditions of disrepair including those that violate article III of the State Sanitary Code, such as inadequate heat during certain months, infestations of insects or rodents, no water, no hot water, defective drains, sewage system or toilet facilities, frequent lack of sufficient hot water and roof or wall leaks. *See* MASS. REGS. CODE tit. 105, §§ 410.001 – 410.960 (2001) (State Sanitary Code – General Procedures and Minimum Standards of Fitness for Human Habitation). *See e.g.*, Boston Housing Auth. v. Hemingway, 293 N.E.2d 831 (Mass. 1973); Berman & Sons, Inc. v. Jefferson, 396 N.E.2d 981 (Mass. 1979); King v. G & M Realty Corp., 370 N.E.2d 413 (Mass. 1977); McKenna v. Begin, 325 N.E.2d 587 (Mass. App. Ct. 1975).

being allowed to raise breach of warranty as a defense to a summary process action for nonpayment of rent.[4] These groups feel that this proposed amendment will separate those tenants who are legitimately withholding rent from those who are "manipulating the system,"[5] which may ultimately result in substantial financial loss to the property owner.[6]

This Note will explore the history of the IWH in the United States, with specific focus on Massachusetts. This Note will also examine an empirical study performed on a pilot program instituted in California[7] regarding mandatory rent escrowing, and further, will discuss the ramifications to both the low-income community and the leasing population as a whole, in the event that the proposed legislation will become law. Finally, this Note will recommend why Massachusetts should not amend Chapter 239, 8A. The IWH operates as an effective tool in enforcing the State Sanitary Code; the proposed amendments to Chapter 239, 8A will further disadvantage low-income tenants in their struggle to level the proverbial playing field.[8]

II. HISTORICAL DEVELOPMENTS

A. Implied Warranty of Habitability

The doctrine of IWH first emerged in a Wisconsin case in 1961.[9] "[T]he

---

4. Telephone Interview with Lenore Schloming, President of Small Property Owners Association (SPOA) (Oct. 3, 2000). SPOA is a Massachusetts organization with approximately thirty-five members in and around Cambridge and Boston. *See also infra* Part III and accompanying notes.

5. Supporters of the bill claim that many tenants create the violations themselves in order to stop paying rent. They further claim that many tenants will then call the Inspectional Services Department [hereinafter ISD], or their local board of health, and report these violations. The "trick" is then to obstruct the landlord's efforts to make repairs so that the tenant can continue to live in the apartment without paying rent. *See Small Property Owners of America, at* http://www.spoa.com/pages/evictions.html (last visited Jan. 25, 2002).

6. *See id.*

7. *See infra* Part III. C. This Note will not examine other states' laws addressing rent withholding as that would be the subject of another Note altogether. The California Project is most similar to the issue of mandatory rent escrow and therefore will be the only other state this Note will discuss.

8. *See infra* note 21 and accompanying text; *see also infra* Part III.

9. *See* Pines v. Perssion, 111 N.W.2d 409, 412-13 (Wis. 1961). In discussing the *Pines* decision one commentator noted:

> In dictum the court developed a theory beyond the furnished home exception and predicated the landlord's obligation to convey habitable premises on municipal housing code and general public policy. Moreo-

court determined that the tenants were not liable for rental payments when their furnished home proved uninhabitable."[10] The dictum in Pines set the stage for the holding of *Javins v. First Nat'l Realty Corp.*,[11] the leading decision of the doctrine of IWH.[12] In 1970 the United States District Court of Appeals for the District of Columbia, in *Javins*, stated that "[w]hen American city dwellers, both rich and poor, seek 'shelter' today, they seek a well known package of goods and services – a package which includes not merely walls and ceilings, but also adequate heat, light and ventilation, serviceable plumbing facilities, secure windows and doors, proper sanitation, and proper maintenance."[13] The court recognized that the traditional landlord-tenant law, derived from feudal property law, was simply not reasonable in an urban setting.[14] The court also recognized that it had a "duty to reappraise old doctrines in the light of the facts and values of contemporary life – particularly old common law doctrines which the courts themselves created and developed."[15] The rationale of the *Javins* holding led many courts to adopt the IWH, and has been "an impetus toward statutory changes in the common law in many . . . states."[16]

At the time that the *Javins* case was decided, the problems that landlords and tenants in urban settings were experiencing were increasingly being handled by provisions in the lease – that is, specific covenants in the lease that set out exactly each party's responsibilities. This ultimately reintroduced a contractual ingredient into property law.[17] The court considered this trend toward treating and interpreting leases like any other contract to be prudent,[18] and ruled that "the old no-repair rule cannot coexist with the

---

ver, the court suggested that the tenant's obligation to pay rent should be considered dependent upon, rather than independent of, the landlord's duty to convey habitable premises.

SCHOSHINSKI, *supra* note 2, at 124.

10. SCHOSHINSKI, *supra* note 2, at 124.

11. 428 F.2d 1071 (D.C. 1970).

12. *See* Boston Housing Auth. v. Hemingway, 293 N.E.2d 831, 842 n.11 (Mass. 1973). *See also* SCHOSHINSKI, *supra* note 2, at 126-27; Erica L. Fox, *Alone in the Hallway: Challenges to Effective Self-Representation in Negotiation*, 1 HARV. NEGOT. L.REV. 85, 90 (1996).

13. *Javins*, 428 F.2d at 1074 (footnote omitted).

14. *See id.* at 1078. "In 1968 more than two thirds of America's people lived in the 228 largest metropolitan areas. Only 5.2% lived on farms." *Id.* at 1078 n.39 (citing THE WORLD ALMANAC 1970, 251 (L. Long ed.)).

15. *Javins*, 428 F.2d at 1074.

16. SCHOSHINSKI, *supra* note 2, at 126-27.

17. *See Javins*, 428 F.2d at 1074 (citing 2 R. POWELL, REAL PROPERTY ¶ 221[1] at 179 (1967)).

18. *See id.* at 1075.

obligations imposed on the landlord by a typical modern housing code, and must be abandoned in favor of an implied warranty of habitability."[19]

This was the beginning of the IWH theory, which imputes a contractual form of consumer protection into landlord-tenant law.[20] The *Javins* court further stated that in regard to products liability, "the relationship of landlord and tenant suggests further compelling reasons for the law's protection of the tenants' legitimate expectation of quality."[21]

B. Covenant of Quiet Enjoyment

The idea of imposing a duty upon the landlord to remedy conditions of disrepair, especially in low-income areas, is not a new one. In 1922, Justice Cardozo wrote:

> We may be sure that the framers . . . when regulating tenement life, had uppermost in thought the care of those who are unable to care for themselves. The Legislature must have known that unless repairs in the rooms of the poor

---

19. *Id.* at 1076-77. The "no-repair rule" refers to the traditional common law rule. Before *Javins* was decided, many courts were generally unwilling to imply warranties of quality or habitability into leases of apartments, holding that the landlord was under no obligation to remedy a condition of disrepair unless he had covenanted to do so in the lease. At common law, the tenant took the property as it was, and the duty was imposed on the tenant to keep the premises in repair:

> Unless the lease contained an express condition to the contrary, the residential tenant had no right to a habitable dwelling. In fact, the landlord made no implied promise that leased premises were fit for any particular purpose. No matter what the purpose of the lease, the tenant's obligation to pay rent was considered unconditional.

SCHOSHINSKI, *supra* note 2, at 109. "The linchpin of this doctrine is the idea that the tenant has possession and exclusive 'control' of the premises." *Id.* at 112.

20. *See* MASS. GEN. LAWS ch. 93A §§ 1–11 (2000). *See also* MASS. REGS. CODE tit. 940, §§ 3.01, 3.17 (2001); Brown v. LeClair, 482 N.E.2d 870 (1985). The *Javins* Court also addressed this issue: "[o]ur approach to the common law of landlord and tenant ought to be aided by principles derived from the consumer protection cases . . . ." *Javins*, 428 F.2d at 1079.

21. *Id.* The court also discussed the unequal bargaining power between the landlord and the tenant, noting that

> the increasingly severe shortage of adequate housing further increases the landlord's bargaining power and escalates the need for maintaining and improving the existing stock . . . , [also] the findings by various studies of the social impact of bad housing has led to the realization that poor housing is detrimental to the whole society, not merely to the unlucky ones who must suffer the daily indignity of living in a slum.

Id. at 1079-80.

> were made by the landlord, they would not be made by anyone. The duty imposed became commensurate with the need.[22]

This same duty was also recognized in what courts termed the "implied covenant of quiet enjoyment."[23] Almost every jurisdiction, in one form or another, recognizes that the landlord impliedly covenants that the tenant shall have the "quiet and peaceable possession and enjoyment of the premises."[24] If the tenant's quiet enjoyment was breached, then the tenant was considered to have been "constructively evicted."[25] A constructive eviction occurred only if a tenant was substantially and materially deprived of the use and enjoyment of the premises, and thereafter abandoned the premises.[26] The theory has been criticized for requiring the tenant to vacate the premises before being able to raise the defense.[27] The theory of quiet enjoyment as a tenant's sole means of recourse against her landlord has been overtaken by the IWH, mainly due to the establishment and enforcement of sanitary codes.[28] A breach of quiet enjoyment counterclaim, however, allows the court to award attorneys' fees. [29]

---

22. Altz v. Leiberson, 134 N.E. 703, 704 (N.Y. 1922).

23. SCHOSHINSKI, *supra* note 2, at 94.

24. *Id.* The covenant of quiet enjoyment was "[i]nitially developed as a means of protecting the tenant against ouster by a superior title holder . . . ." *Id.*

25. *See* HERBERT THORNDIKE TIFFANY, 2 LANDLORD AND TENANT § 185, at 1259 (1910). Courts usually referred to particular acts or a series of acts on the part of the landlord as constituting a constructive eviction. The tenant was considered to be "constructively evicted" if he left the premises

> as a result of the landlord's interference with his enjoyment. In order that an eviction may take place as a result of acts on the part of the landlord . . . it is necessary that they be such as to indicate an intention on the landlord's part to deprive the tenant of the possession.

Id.

26. *See* Bartlett v. Farrington, 120 Mass. 284 (1876).

27. *See* SCHOSHINSKI, *supra* note 2, at 100-01.

28. *See* Lemle v. Breeden, 462 P.2d 470, 475 (Haw. 1969). "The doctrine of constructive eviction, as an admitted judicial fiction . . . no longer serves its purpose when the more flexible concept of implied warranty of habitability [and fitness] is legally available." *Id.*

29. *See* MASS. GEN. LAWS ch. 186, § 14 (2000). Massachusetts awards damages for breach of quiet enjoyment in the amount of actual damages, or three months rent, whichever is higher. The use of the implied covenant of quiet enjoyment is usually raised in Massachusetts as an alternative to other defenses and counterclaims (mainly Chapter 239, 8A) due to the fact that if awarded damages under this section you cannot be awarded damages under another theory for the same injury:

C. Development of IWH in Massachusetts

Massachusetts first recognized the IWH as a defense to a summary process action for nonpayment of rent in *Boston Housing Authority v. Hemingway*.[30] The tenants in this case repeatedly complained to their landlord of existing conditions of disrepair. The landlord refused to make repairs so the tenants called the Inspectional Services Department [hereinafter ISD].[31] The tenants only began to withhold rent after ISD served the landlord with a report citing serious code violations.[32] The Boston Housing Authority [hereinafter the BHA] then brought an action for nonpayment of rent against the tenants.[33] The tenants petitioned the court to find that "the obligations of the Housing Authority to supply and maintain premises in compliance with the Housing Regulations and the obligations of the tenant to pay rent under a rental agreement are dependent."[34] The trial court ruled in favor of the BHA and refused to allow the tenants to raise the alleged

---

> We see no reason . . . for the plaintiff to recover cumulatively for a breach of the implied warranty of habitability and for interference with her quiet enjoyment of the premises. In the absence of any statute authorizing recovery beyond her actual loss, she may not recover for the same wrong under each theory.

Darmetko v. Boston Hous. Auth., *393 N.E.2d 395, 398* (Mass. 1979).

30. *See* 293 N.E.2d 831 (Mass. 1973). Chapter 239, 8A was first enacted in 1965, and in 1967 read in part:

> There shall be no recovery under this chapter, pursuant to a notice to quit for nonpayment of rent . . . of any tenement rented or leased for dwelling purposes if such premises are in violation of the standards of fitness for human habitation established under state sanitary code . . . and if such violation may endanger or materially impair the health or safety of persons occupying the premises; provided, however (1) that the person occupying the premises, while not in arrears in his rent, gave notice in writing to the person to whom he customarily paid his rent (a) that he would, because of such violation, withhold all rent thereafter becoming due until the conditions constituting such violations were remedied and (b) that a report of an inspection of such premises has been issued by the board of health . . . which report states that such violation exists and that it may endanger or materially impair the health or safety of persons occupying said premises.

Hemingway, 293 N.E.2d at 836 n.4 (emphasis added). Chapter 239, 8A was amended in 1969 to provide that written notice from the Boston Housing Inspection Department, or any other appropriate local agency, was sufficient to satisfy the obligation the tenant had to notify the landlord in writing of the violations. *See id.* at 836 n.5.

31. *See id.* at 835.

32. *See id.*

33. *See id.*

34. *Id.* at 836.

breach of an implied warranty of habitability as a defense to the summary process action.[35]

The trial court was correct in applying the common law, as it then existed in Massachusetts.[36] The time had come, however, to reevaluate the common law. On appeal, the Massachusetts Supreme Judicial Court [hereinafter the SJC] conducted a thorough analysis of the history of the common law, and of a series of cases upon which the BHA was relying, "which established and followed the doctrine of caveat emptor[37] and independent covenants between landlord and tenant."[38] "Basically, under this ancient view, the axiom was 'there is no law against letting a tumbledown house.'"[39] These cases all held that if the landlord was not expressly bound by the covenant to repair conditions of disrepair (including what would now be considered sanitary code violations) he was under no duty to do so, and the tenant had no right to quit the premise or to refuse to pay rent; his only recourse would be to sue for damages after being evicted.[40] "Given the rural agrarian context in which these [judicially formulated rules existed,] the independent covenants made sense."[41]

To combat the harsh results that did occur, the Massachusetts courts de-

---

35. *See Hemingway*, 293 N.E.2d at 836-37.

36. *See id.* The "common law" refers to the independent covenants rule, which was still in effect in Massachusetts at the time. *See* SCHOSHINSKI, *supra* note 9 and accompanying text.

37. This maxim basically stands for the proposition that a purchaser must examine, judge, and test for himself before purchasing. At common law, a lease was considered to be a sale of a premises for a certain period of time, and unless the lease contained an express provision to the contrary, the lessee had no right to a habitable dwelling, and thus took the property as it was; therefore the maxim of caveat emptor applied to residential tenants the same as it applied to any other purchaser of goods. *See* SCHOSHINSKI, *supra* note 2, at 109.

38. *Hemingway*, 293 N.E.2d at 837.

39. Young v. Garwacki, 402 N.E.2d 1045, 1047 (Mass. 1980) (citation omitted).

40. *See Hemingway*, 293 N.E.2d at 837. "Thus, originally at common law, the tenant could not even escape his rental obligation when the demised premises were destroyed because of the law's view that the land and not the premises was the essential part of the transaction." *Id.*

41. *Id.* At common law, the tenant had a duty to make minor repairs which were necessary to preserve the leased premises. *See* SCHOSHINSKI, *supra* note 2, at 269. In an agrarian economy it was expected that the tenant would make repairs since he likely possessed the requisite skills and materials to do so. *See id.* This may still be the expectation of parties to agricultural and long-term tenancies. Urban tenants, however, are much less likely to possess the skills or financial resources to make such repairs. *See id.* "There is little doubt that the rule fails to conform with the expectations of the modern residential tenant seeking a habitable dwelling and attendant services." *Id.*

veloped the constructive eviction defense.[42] This defense, however, was ineffective for tenants who were more interested in securing a habitable dwelling than a plot of land.[43] Judicial reform over the years gradually eroded the independent covenants rule, and enactment of the state sanitary code effectively replaced the constructive eviction theory as a means of redress for the uninhabitable premises.[44]

Chapter 111, §§ 127A – 127F of the Massachusetts General Laws lays out the adoption and enforcement of the state sanitary code and provides tenants with a private remedy of enforcement.[45] "The common purpose

---

42. *See* Royce v. Guggenheim, 106 Mass. 201, 202 (1870). "Any act of a permanent character, done by the landlord, or by his procurement, with the intention and effect of depriving the tenant of the enjoyment of the premises demised, or of a part thereof, to which he yields and abandons possession, may be treated as an eviction." *Id. See also supra* Part II. B.

43. *See Hemingway*, 293 N.E.2d at 838.

44. *See id.* at 838-39.

45. Section 127A reads in part:

> The *code shall deal with matters affecting the health and well-being* of the public in the commonwealth in subjects over which the department takes cognizance and responsibility, including, but not limited to, *standards of fitness for human habitation, housing and sanitation standards* . . . . Said code may provide for the demolition, removal, repair or cleaning by local boards of health and, in the cities of Boston, Worcester and Cambridge, by the commissioner of housing inspection, of any structure which so fails to comply with the standards of fitness for human habitation or other regulations in said code, as to endanger or materially impair the health or well-being of the public . . . . This designation shall not be construed as prohibiting an inspector or other authorized person from certifying that any other violation or combination or series of violations of said code or other applicable laws, ordinances, by-laws, rules or regulations may endanger or materially impair the health or safety of said persons when such certification is otherwise appropriate.

MASS. GEN. LAWS ch. 111, § 127A (1996) (emphasis added). Section 127B addresses dwellings that are deemed to be unfit for human habitation, and reads in part:

> Upon a *determination by the board of health*, or, in the cities of Boston, Worcester and Cambridge, by the commissioner of housing inspection, *after examination as provided in said code*, that a building, tenement, room, cellar, mobile dwelling place or any other structure (a) *is unfit for human habitation, (b) is or may become a nuisance, or (c) is or may be a cause of sickness or home accident to the occupants or to the public, it may issue a written order to the owner or occupant or any of them thereof, requiring the owner or occupant to vacate, to put the premises in a clean condition*, or to comply with the regulations set forth in said code which are not being complied with or to comply with the rules and regulations adopted by the board of health as being necessary for the particular locality . . . . If the *owner or occupant refuses* to comply with such order, the *board of health may cause the premises to be properly cleaned at the expense of the owner or occupant*, remove the occupant

> forcibly and close up the premises, or it may issue a written notice to the owner of such building, as appearing in the current records of the assessors of such town, setting forth the particulars of such unfitness and requiring that the conditions be remedied. If the person so notified fails within a reasonable time to remedy the conditions thus set forth, the superior court on a petition in equity brought by the board of health, shall have jurisdiction, to enforce the requirements of the board of health, by appointment of a receiver in accordance with the provisions of section one hundred and twenty-seven I, or by providing other equitable or legal relief.

MASS. GEN. LAWS ch. 111, § 127B (1996) (emphasis added). Section 127C states:

> If the condition of any building or any part thereof used for residential purposes is in violation of the standards of fitness for human habitation established under the state sanitary code, as provided by section one hundred and twenty-seven A, or in violation of any board of health standards, a petition may be filed by any affected tenant or by the board of health, or, in the cities of Boston, Worcester and Cambridge, by the commissioner of housing inspection, in a district court, housing court, or superior court. The petition shall state that the premises have been inspected by the board of health or other appropriate inspection agency and that the condition of said premises has been found to be in violation of the state sanitary code; that such condition may endanger or materially impair the health or well-being of any tenant therein; and that said condition was not substantially caused by the tenant or any other person acting under his control. An affected tenant may also file a petition in such court without stating that such premises have been found by an appropriate inspection agency to be in violation of any such standards, provided that he states (a) facts sufficient to demonstrate a likely violation of the state sanitary code which may endanger or materially impair the health or well-being of any tenant in the building or premises occupied by such tenant, (b) that the conditions in question were not substantially caused by such tenant or any person acting under his control and (c) that a request for inspection of such premises was made to the appropriate inspecting agency at least twenty-four hours prior to the filing of the petition and that there has been no inspection.

MASS. GEN. LAWS ch. 111, § 127C (1996) (emphasis added). Section 127F, Court Order Regarding Rental Payments; Disbursement of Funds; Receiver:

> If the court finds after hearing that the facts are as alleged in said petition, it may by written order authorize the petitioner or any other affected tenant to pay to the clerk of the court the fair value of the use and occupation of the premises, or such installments thereof from time to time as the court may direct; provided the court finds that any such violation may endanger or materially impair the health, safety or well-being of such tenant and that such payments are necessary to remedy the condition constituting the violation; and that said tenant is not in arrears in his rent, taking into account any abatement of rent due to the existence of such violations, or if he is in arrears is willing to pay any arrearage into the court as ordered. A person occupying the premises shall not be considered to be in arrears in his rent when the amount of rent the landlord alleges in good faith to be due is equal to or less than the amount of any counterclaim that said person may bring in good faith against the landlord, including any damages owed because of a breach of warranty or a violation of any other law. In determining said fair value, the court may be guided by any findings issued by the department or any other

underlying these new statutes was to create a private remedy for these public violations by giving the tenant the power to initiate the Code enforcement process . . . by bringing before the courts recalcitrant landlords who refused to remedy conditions in the tenants' dwelling units . . . ."[46] The state recognized, however, that "many tenants, especially poor tenants, [may] not avail themselves of a remedy which required them to sue their landlords," and therefore, the legislature created the remedy set forth in Chapter 239, 8A.[47] Both Chapters 239, 8A, and 111, 127A – 127F were designed to "promote repairs by giving the landlord the incentive of getting his rent if he complies with the Code regulations."[48] The *Hemingway* decision was a very important decision because it allowed, and continues to allow, tenants protections and defenses to summary process.

D. History of Rent Escrowing in Massachusetts

The provisions discussed above were essentially designed to protect tenants against their more powerful landlords. The protections of the property owners were addressed by the Boston Housing Court [hereinafter the BHC] when it first ruled on the issue of rent escrowing in 1974.[49] In *Slack*, the BHC held that Chapter 239, 8A, as it was then drafted, allowed the court the power to make a "pre-trial order."[50] The plaintiff in that case

---

> agency concerned with the condition of residential premises and any other evidence relative to the effect of violations of such standards upon the use and occupation of residential premises. Any order of the court shall be forthwith sent by the clerk of court to any and all mortgagees and lienors of record as the court deems proper under the circumstances.

MASS. GEN. LAWS ch. 111, § 127F (1996) (emphasis added).

46. *Hemingway*, 293 N.E.2d at 839.

47. *See id.* at 839. Chapter 239, 8A,

> authorizes tenants to withhold rent where a housing code inspection report citing sanitary code violations has been issued to the landlord. The tenant can take this action without having to sue the landlord and without fear of being evicted for nonpayment of rent. Chapter 239, § 8A, amended summary process so that the landlord's violation of State Sanitary Code standards of fitness for human habitation, if properly raised, would constitute a defence for the tenant in that proceeding.

Id.

48. *Id.* at 840.

49. *See* Copley Mgmt. & Dev. Corp. v. Slack, 74-SP-02335 (Mass. Dist. Ct. 1974) (the tenant had been withholding rent due to conditions of disrepair in the subject premises).

50. *See id.* at 6. A pre-trial order refers to the amount of withheld rent money that the judge could order the tenant to pay into the court before the start of a trial on the merits. *See id.* This was the first time that Chapter 239, 8A had been read to allow a pre-trial order as opposed to ordering money into court after a full

moved to require the defendant to deposit withheld rent with the court, as well as any future rent which would become due during the summary process proceeding.[51] The defendant countered that a full evidentiary hearing before a jury would be required in order to determine exactly what amount would be deposited if the court were to grant such a motion.[52] The amount to be deposited would have to be based on the fair rental value of the premises with the conditions of disrepair, and not the contract rent.[53] Because Judge Daher's opinion in *Slack* was in direct contrast to other BHC judges, he reported the matter, *sua sponte*, to the SJC.[54]

The defendant's position in *Slack* was that the task of having a full evidentiary hearing would be so burdensome to the court that no determination as to the amount should be made until after a trial on the merits.[55] The issue of a pre-trial hearing involves constitutional considerations of due process.[56] At the time that *Slack* was decided, Chapter 239, 8A read, in pertinent part:

> The Court may require the person claiming a defense under this section to pay all or any portion of the rent due or to become due into court, or make a deposit in court of the entire amount, or such installments thereof from time to time, as the court may direct, for the occupation of the premises.[57]

Judge Daher interpreted this to mean that the power to order money into

---

trial on the merits. *See* Telephone Interview with Harvey J. Chopp, Esq., Court Administrator, The Commonwealth of Massachusetts, Housing Court Department (Oct. 6, 2000).

51. *See id.* at 5.

52. *See id.*

53. *See id.*

54. *See id.* at 14.

55. *See Slack*, 74-SP-02335 at 6.

56. *See* Fuentes v. Shevin, 407 U.S. 67 (1972) (holding that due process requires "an opportunity for an adversary type hearing before a person can be even *temporarily* deprived of any possessory interest in personalty"). *See also* Lucky Ned Pepper's Ltd. v. Columbia Park and Recreation Assoc., 494 A.2d 947 (Md. App. 1985) (holding that the requirement of payment of past due rent into escrow as a precursor to a jury trial is an unconstitutional infringement of one's right to a civil jury trial because it places a premium on the exercise of that right); Amanuensis Ltd. v. Brown, 65 Misc.2d 15 (N.Y.S. 1971) (holding unconstitutional statutory provisions which required that tenants deposit the amount of rent sought to be recovered with the court as a precondition to asserting conditions defenses); Sniadach v. Family Fin. Corp., 395 U.S. 337 (1969) (striking down as unconstitutional a Wisconsin Statute which permitted the prejudgment garnishment of wages).

57. *Slack*, 74-SP-02335 at 6.

the judicial system was in the hands of the courts.[58]

The court first resolved the issue of whether any type of hearing was required before a pre-trial order could be made,[59] and found that "courts cannot know what money is to be deposited, until evidence is received as to the worth of the apartment."[60] Therefore, the amount to be deposited with the court was to be based on the fair rental value of the apartment, not the contract rent.[61] Judge Daher went on to find that when a tenant raises a Chapter 239, 8A defense, the court should immediately allow the Plaintiff to bring a motion asking for money payments into court and then have a hearing on the matter.[62] Believing that *Boston Housing Authority v. Hemingway*[63] also interpreted Chapter 239, 8A to require a pre-trial hearing, Judge Daher quoted the opinion written by Judge Tauro:[64]

> Unlike the more cumbersome procedure established in C. 111, §§ 127A – 127H[65] where the landlord can use delaying tactics by asking for repeated continuances, C. 239, § 8A's provisions force him to take the initiative by making repairs or contesting the tenant's actions in court in order to recover the withheld rent.[66]

Judge Daher stated that the last part of the quote, which specifically deals with recovering withheld rent, makes it very clear that *Hemingway*

---

58. *See id.*

59. Judge Daher turned to Fuentes v. Shevin, 407 U.S. 67, 87 n.18 (1972) in which the Supreme Court held: "'The issues decisive of the ultimate right to continued possession, of course, may be quite simple. The simplicity of the issues might be relevant to the formality or scheduling of a prior hearing. *See Lindsay v. Normet*, 405 U.S. 56, at 65. *But it certainly cannot undercut the right to a prior hearing of some kind.*'" (emphasis added by Judge Daher). *Slack*, 74-SP-02335 at 7.

60. *Slack*, 74-SP-02335 at 7.

61. *See id.* The court reasoned that:

> [i]f the Court limits itself solely to the contract rent, that money which is in excess of the apartment's worth, would be considered confiscatory and unconstitutional . . . . Mass. Gen. Laws, Ch. 239, § 8[A] has spoken, in that the statute refers to 'all or any portion of the rent.' If money is to be paid into court, clearly it must be that which relates to the worth of the apartment.

Id. at 7-8.

62. *See id.* at 8.

63. 293 N.E.2d 831 (Mass. 1973).

64. *See id.* at 835.

65. *See supra* note 45 and accompanying text.

66. *Hemingway*, 293 N.E.2d at 839.

stands for the proposition that a tenant must withhold rent in order to claim the defense in Chapter 239, 8A. Judge Daher went on to find that if the tenant is withholding rent, "[w]hat prejudice is there, if the Court decides that it, and not the tenant, should be the one to hold the rent due?"[67]

The next issue of concern was the delay of the trial, and the proceeding as a whole, which would ultimately result if a pre-trial hearing were required before the court would order a deposit. Judge Daher stated that neither party should complain about the delays,[68] and must realize that although the requirement of a pre-trial hearing may involve a duplication of the trial itself, it is necessary in order to achieve the main objective of Chapter 239, 8A.[69] Judge Daher further commented on the "radical change in landlord-tenant law," which Massachusetts was experiencing, and that *Hemingway* "merely confirmed that which had already occurred."[70] Judge Daher determined that the SJC in *Hemingway* had explored the intent of the legislature in enacting Chapter 239, 8A, and went even further by stating that "[t]he statute in essence is *Hemingway*."[71]

Hemingway was decided on March 5, 1973.[72] Between that time and the time the *Copley Management & Development Corp. v. Clair Slack*[73] case was decided, the Legislature amended Chapter 239, 8A several times.[74] Judge Daher did not make a pre-trial order in this case due to the fact that there was disagreement in the BHC over the following three questions:

1. What rules are applicable to summary process proceedings?

---

67. *Slack*, 74-SP-02335 at 9.

68. *See id.* at 10. Judge Daher held that the landlord should not complain because it would be his motion that initiated the delay, and the tenant should not complain because it is he who desires that the apartment be habitable. *See id.*

69. *See id.* In order "to improve living conditions in our commonwealth, obtaining compliance with the State Sanitary Code, Minimum Standards of Fitness for Human Habitation, and the Building Code of the City of Boston, . . . it is imperative that the withheld rent be deposited." *Id.*

70. *Id.* at 11 (specifically referring to Chapter 239, 8A).

71. *Id.* at 12. "It is a well settled rule of statutory interpretation that, when a statute after having been construed by the courts is re-enacted without material change by the Legislature, it is presumed to have adopted the judicial construction put upon it." *Id.* at 12-13 (quoting Nichols v. Vaughan, 105 N.E. 376, 378 (Mass. 1914)).

72. *See* Boston Housing Authority v. Hemingway, 293 N.E.2d 831 (1973).

73. *See Slack*, 74-SP-02335 (Mass. Dist. Ct. 1974).

74. *See id.* "St. 1974, C. 269, St. 1973, C. 471 – June 29, 1973; St. 1973, C. 1007 – November 8, 1973 . . . Th[e] section of 8(a), involving rent withholding, was left untouched; the present statute, with respect to paying money into court remained intact, and exactly the same as when the Supreme Judicial Court rendered its *Hemingway* decision." *Id.*

> 2. Does the Court have the power to make a pre-trial order, under Mass. Gen. Laws. C. 239, § 8(a), and if it does, is some type of hearing required?
>
> 3. If money is to be paid into court, can the Court consider the actual value of the premises, or is the Court limited to the contract rent?[75]

These issues were never resolved by the SJC.[76] Rather, the Legislature enacted St. 1975, c. 467, § 3 on July 11, 1975, "which dramatically rewrote the portion of the statute interpreted by Judge Daher, and thus overruled the *Slack* decision.[77] Thereafter, the Legislature amended the statute several more times which further altered the nature of the statute."[78] Although the courts today still have the power to order a pre-trial deposit, the amendments which have been made to Chapter 239, 8A now require the court to consider all counter-claims raised by the tenant. This makes the pre-trial hearing an exact duplicate of the trial on the merits, which effectively discourages judges from making the order.[79]

Chapter 239, 8A has been amended several times since 1975,[80] and the Massachusetts courts have since liberally interpreted the IWH. For example, in 1979 the SJC, in *Berman v. Jefferson*[81] was called upon to decide:

> (1) whether a tenant must pay full rent without abatement when the landlord, acting without fault or bad faith, fails to maintain a dwelling in habitable condition and (2) whether the tenant's obligation to pay full rent persists until the landlord has had a reasonable time to repair the defect.[82]

The court cited to *Boston Housing Authority v. Hemingway* and acknowledged the impact that social changes have had on the landlord-tenant relationship.[83] Recognizing that legislative changes had effected landlord obligations and tenant remedies, the court made reference to the fact that a lease is a contract with mutually dependent covenants – the tenant cove-

---

75. *Id.*

76. Chopp Interview, *supra* note 50.

77. *Id.*

78. *Id.*

79. *See id.* The court recognizes the statutory right of summary process defendants to assert counter-claims. *See* Rule 5 of the MASSACHUSETTS UNIFORM SUMMARY PROCESS RULES.

80. *See* MASS. GEN. LAWS ch. 239, § 8A (2000).

81. 396 N.E.2d 981 (Mass. 1979).

82. *Id.* at 983. The tenant in this case withheld rent because she was without heat for several weeks. The landlord brought a summary process action for non-payment of rent, and the tenant, in her answer, raised the counter-claim of breach of the IWH. *See id.*

83. *See id.* at 983-84.

nants to pay rent, which is dependent on the landlord's covenant to keep the premises in a habitable condition.[84]

The landlord's main argument in *Jefferson*, and subsequent appeal, revolved around the issue of his lack of fault.[85] He argued that "to impose strict liability would penalize the landlord who is acting reasonably and would impose a duty impossible to meet."[86] The duty to which he is referring, however, is the essence of Chapter 239, 8A, and is a right protected by law.[87] The problem with the landlord's argument is that there is no fault requirement or standard in the IWH[88] - basically, the "innocence" of the landlord is irrelevant.[89] The landlord's argument is essentially one of economic burden.[90] The SJC responded with empirical evidence which, "tends to show a very low rate of tenant enforcement,"[91] further stating that the economic burden a landlord may suffer as a result of enforcing the IWH does not compare to the loss of shelter a tenant living in an unhabitable building suffers:

> The landlord would have us avert our eyes from the clear teaching of *Hemingway*: '[T]he essential objective of the leasing transaction is to provide a dwelling suitable for habitation.' *Boston Hous. Auth. v. Hemingway* [citation omitted]. A dwelling afflicted with a substantial Sanitary Code violation is not habitable. The essential objective of the warranty is to make sure that the tenant receives what she is paying for. The tenant may not excuse her obligation with mere reasonable efforts to pay rent. Nor may the landlord avoid

---

84. *See id.* at 984.

85. *See id.* The landlord made reference to the fact that no tenant expects a perfect apartment, and that upkeep takes time – "systems breakdown; maintenance takes time; the law should reflect this reality." *Id.*

86. *Id.*

87. *See* Berman v. Jefferson, 396 N.E.2d 981, 984 n.6 (Mass. 1979) ("A tenant has a right to expect that the landlord will comply with the law.").

88. *See id.* at 984-85.

89. *See* MASS. GEN. LAWS ch. 239, § 8A. The 1977 amendment to 239, 8A struck the provision allowing a tenant to withhold rent for code violations only if the "owner . . . had not taken reasonable steps to remedy such conditions" thereby allowing the tenant to withhold rent without considering whether the landlord is at fault or is taking reasonable steps to repair. *Id.*

90. *See* Berman v. Jefferson, 396 N.E.2d at 985. The Greater Boston Real Estate Board submitted an amicus brief and argued "that to make the landlord the insurer of apartment maintenance will have deleterious economic effects on landlords and on the housing market generally. Increased costs derived from the warranty will result in increased rents . . . . Thus, imposing full warranty protection will ultimately reduce the stock of low income housing." *Id.* at 984 n.8.

91. *Id.* at 985 n.8 (citing Samuel Bassett Abbott, *Housing Policy, Housing Codes and Tenant Remedies: An Integration*, 56 B.U. L. REV. 1, 63 (1976)).

his duty with mere reasonable efforts to provide a habitable dwelling.[92]

Other significant holdings in reference to the IWH in Massachusetts include the Appeals Court decision in 1977 ruling that damages for breach of the IWH were to be measured by reducing the tenant's rent by a percentage which would reflect the diminution in value of the premises due to the existence of the defects which gave rise to the breach;[93] and the SJC's holding in 1983 that a landlord is even liable to repair violations of the sanitary code in an apartment occupied by a tenant at sufferance who continued to pay rent.[94]

E. Housing Codes

Housing Codes regulate and govern dwelling units, usually permanent residences such as apartments or houses as opposed to hotels or motels.[95] The Codes regulate four different areas of housing: "structural elements, facilities, services and the number of occupants."[96] Housing Codes differ

---

92. *Id.* at 985.

93. *See* McKenna v. Begin II, 362 N.E.2d 548 (Mass. App. Ct. 1977). The SJC also held in 1982 that a tenant who justifiably withholds rent due to conditions of disrepair is allowed to recover damages under the Massachusetts Consumer Protection Law, MASS. GEN. LAWS ch. 93A. *See* Wolfberg v. Hunter, 432 N.E.2d 467 (Mass. 1982). The proper measure of damages under 93A is to calculate the rental value of the apartment as warranted (agreed rent) minus the value of the apartment in defective condition, plus reasonable expenses incurred as a result of the condition. *See id.* at 473-74. The judge is then allowed to double or treble that amount if he finds that the breach of the IWH was a willful or knowing violation of chapter 93A, or if he finds that the landlord made a bad faith refusal to settle. *See id.* Lastly, the total amount of withheld rent will be subtracted in order to prevent excessive recovery. *See id.* In some instances, a landlord can even be held liable in tort for intentional infliction of emotional distress. *See* Haddad v. Gonzalez, 576 N.E.2d 658, 667 (Mass. 1991).

94. *See* Brown v. Guerrier, 457 N.E.2d 630, 632 (Mass. 1983).

95. *See* Samuel Bassett Abbott, *Housing Policy, Housing Codes and Tenant Remedies: An Integration*, 56 B.U. L. REV. 1, 40 (1976).

96. *Id.*

> Structural elements . . . include walls, ceilings, floors, roofs, windows, and staircases. Facilities include fixtures . . . such as toilets, bathtubs, sinks, radiators, window screens, stoves, door and window locks, electrical outlets and air conditioning units. Services include heat, hot and cold water, electricity, sanitary sewage disposal, and maintenance and repair supplied to the dwelling unit. Occupancy standards set limits on the number of occupants . . . or the occupancy of rooms used for sleeping.

Id. at 40-41.

from other building and zoning codes in that they only cover habitations, and are only concerned with ongoing repairs, maintenance and other services.[97]

Housing codes were first enacted in the mid to late nineteenth century.[98] New York's tenement house law of 1867 was mainly enacted in response to the large concentration of immigrant poor who were living in deplorable conditions, and in response to the fear that these conditions were responsible for the widespread cholera epidemics which effected the poor and rich.[99] The public health purpose was clear due to the creation of the Metropolitan Board of Health one year earlier as an enforcement agency.[100]

Massachusetts enacted a statute similar to the New York statute in 1868.[101] These statutes were enacted in an effort to eradicate a "housing evil":

> Housing evils as we know them today are to be found in dangerous and disease-breeding privy vaults, in lack of water supply, in dark rooms, in filthy and foul alleys, in damp cellars, in basement living rooms, in conditions of filth, in inadequate methods of disposal of waste, in fly-born disease, in cramped and crowded quarters, in promiscuity, in lack of privacy, in buildings of undue height, in inadequate fire protection, in the crowding of buildings too close together, in the too intensive use of land.[102]

The migration of lower income African-Americans and Whites to the urban centers of the United States after World War II was a principal impetus in the enactment of the Housing Act of 1949.[103] "Implicit in this legislati[on] . . . was the belief that a housing code, properly drafted and effectively enforced, would be a significant weapon in the fight against

---

97. *See id.* at 41.

98. *See id.* "The earliest example of a housing law as such is the New York tenement house law of 1867." *Id.* (citation omitted).

99. *See id.* at 42.

100. *See id.*

101. *See* Abbott, *supra* note 95, at 42.

102. LAWRENCE VEILLER, A MODEL HOUSING LAW 5 (1920). Veiller stated that the only way to remedy these evils is through legislation which would compel repair. *See id.*

103. *See* Housing Act of 1949, ch. 338, §§ 101-10, 63 Stat. 413 (*codified as amended*, 42 U.S.C. §§ 1450-90 (1970)). The Housing Act provided federal grants and loans for slum clearance and urban redevelopment. *See* Abbott, *supra* note 95, at 43. In the mid-1950s the concept was changed from redevelopment to renewal. *See id.* "The change signaled the belated recognition that conserving and upgrading existing structures and neighborhoods might, in some cases, be preferable to bulldozing away whole sections of cities." *Id.*

spreading 'blight' – the phenomenon of undermaintenance . . . ."[104] As a justification for the enactment and need for housing codes, the associate general counsel for the Housing and Home Finance Agency[105] stated:

> Slums and blight are brought about by owners of property who are unable or unwilling to maintain or improve their property at decent levels, by unconscionable, profiteering landlords squeezing bootleg profits out of wretched housing, and occasionally by tenants who are indifferent to their squalid environments. But . . . the ultimate causation factor is the local government itself [that fails] to enforce effectively . . . adequate police power measures to control bad housing, improper environments and overcrowding.[106]

As a result of the Federal Housing Code requirement, communities across the United States began enacting their own Housing Codes.[107]

No matter how well-thought-out and properly drafted Housing Codes are, they are of little significance if they are not enforced. Housing Codes were initially enforced administratively by the boards of health, which were empowered to order the tenants to vacate the premises in the event of substantial Code violations.[108] The problem with this method was that it was very drastic and did not solve the problem, in fact, it often resulted in removing "dwelling units from the housing stock, temporarily or permanently, and leav[ing] tenants homeless."[109] The other "traditional" method of Code enforcement was the criminal proceeding, which usually led to a fine or imprisonment, but did not require the owner to remedy the violations.[110]

Tenant remedies arose mostly due to frustration with administrative enforcement.[111] Of the major remedies that have evolved over the years, the only remedy which will be discussed in this Note is rent withholding.[112]

---

104. *Id.*

105. The Housing and Home Finance Agency was the agency that administered the urban renewal program. *See id.* at 44.

106. *Id.* at 44 (*citing* Guandolo, *Housing Codes in Urban Renewal*, 25 GEO. WASH. L. REV. 1, 1-14 (1956)).

107. *See id.* The National Commission on Urban Problems reported that by 1968 almost 5,000 communities had enacted housing codes. *See id. See also* Judah Gribetz & Frank P. Grad, *Housing Code Enforcement: Sanctions and Remedies*, 66 COLUM. L. REV. 1254, 1260 n.19 (1966).

108. *See* Abbott, *supra* note 95, at 49.

109. *Id.* at 49.

110. *See id.* at 50-51.

111. *See id.* at 56.

112. The major remedies for tenants are: rent application, rent withholding, rent abatement and mandatory injunctions. *See* GOLDSTEIN, *infra* note 115 and accompanying text (discussing traditional tenant remedies).

The IWH, in large part, was created to help tenants, predominantly low-income, enforce the housing code.[113] The SJC made it abundantly clear in *Hemingway* that the purpose of the IWH was to provide disadvantaged tenants a remedy against the system.[114] Allowing tenants to withhold rent due to code violations, and thereafter allowing them to raise the breach of IWH as a defense to a nonpayment summary process action, has been the most effective (of all available remedies) in forcing landlords to remedy code violations.[115] The majority of tenants who are sued for nonpayment of rent "are almost always poor and living near subsistence. In contrast, landlords generally have more money, education, and experience with navigating the court systems."[116]

The "purposes" of the Massachusetts State Sanitary Code "are to protect the health, safety and well-being of the occupants of housing and of the general public . . . ."[117] The Scope section of the Code states that "no per-

---

113. *See* Abbott, *supra* note 95, at 3, 25. Abbott concluded in his article "that housing code enactment is the real basis for the adoption of the warranty and that more effective code enforcement is its purpose." *Id.* at 3.

114. *See supra* notes 46-48 and accompanying text.

115. The doctrine of constructive eviction was a remedy to the tenant only in the sense that it allowed the tenant to break the lease and leave the premises – it did not require the landlord to repair the premises or allow the tenant to remain in possession. *See supra* Section B and accompanying text. Other traditional remedies available to tenants have included: (1) abatement of the rent due to the conditions of disrepair; (2) repair and deduct – the tenant would make his own repairs and deduct the cost from the next month's rent; (3) vacate and rescind – this is the same principle as constructive eviction, whereby the tenant is allowed to break the lease and leave the premises; (4) escrow and receivership – tenant continues to pay rent into escrow to be used for repairs, or until the repairs are made (essentially what the proposed Bill Relative to Rent Withholding seeks to mandate); and (5) the tenant can stay in the premises and bring an affirmative case against the landlord for damages as a result of the conditions of disrepair. *See* PAUL GOLDSTEIN, REAL PROPERTY 956-58 (University Casebook Series, 1984).

116. Joel Kurtzberg, *Freeing the Parties From the Law: Designing an Interest and Rights Focused Model of Landlord/Tenant Mediation*, 1997 J. DISP. RESOL. 53, 60 (1997). *See also* Barbara Bezdek, *Silence in the Court: Participation and Subordination of Poor Tenants' Voices in Legal Process*, 20 HOFSTRA L. REV. 533, 557 (1992) (noting that rent court "is a theater of class conflict in which businesses and their hirelings constitute a class of professional claimants exercising significant advantages over the individual defendants whom they bring before the court, who are poor and poorly situated with respect to the attributes that garner respectful hearing in court rooms"). Landlords are represented by legal counsel in summary process actions over 50% of the time, whereas tenants are represented by counsel only 7% of the time; landlords are awarded possession of the premises in close to 90% of the cases. *See* Massachusetts Law Reform Institute, *Summary Process Survey*, Number 3, at 3 (Sept. 1999) [hereinafter Survey].

117. MASS. REGS. CODE tit. 105, § 410.001 (1986).

son shall occupy as owner-occupant or let to another for occupancy any dwelling . . . for the purpose of living, sleeping, cooking or eating therein, which does not comply with the requirements of" the State Sanitary Code.[118] Pursuant to chapter 111, section 127A,[119] "the Department of Public Health has promulgated the State Sanitary Code to assure, among other things, that any premises rented for dwelling purposes are fit for human habitation. Substantial violations of the State Sanitary Code generally make a dwelling uninhabitable."[120]

The State Sanitary Code comes into practical play when a tenant, suspecting code violations in the apartment, calls ISD, or the local Board of Health and requests an inspection.[121] The tenant must usually request a general inspection if they want the inspector to examine the entire apartment. The inspection is most effective when the tenant is present with a list of conditions they want specifically inspected. If the tenant is without heat, light or sanitary drainage system, the inspection should be performed within twenty-four hours.[122] The inspector is required to inspect the premises for these conditions, and is also allowed to view as much of the building as is necessary to complete the inspection.[123] The inspector will then issue a written report of his findings – one copy is delivered to the tenant and another copy is delivered to the owner of the building.[124] The inspection report must be certified by the ISD or the Board of Health in order to be admissible as evidence in court.[125] The inspection report, if properly certified, is a strong piece of evidence to support the tenant's contention that the landlord had notice of the conditions of disrepair.[126]

The State Sanitary Code plays a vital role in ensuring the quality of housing every Massachusetts resident is entitled to enjoy. Chapter 239,

---

118. *See* tit. 105, § 410.010.

119. *See supra* note 45 and accompanying text.

120. Poncz v. Loftin, 607 N.E.2d 765, 766 (Mass. App. Ct. 1993).

121. *See* MASS. REGS. CODE tit. 105, § 410.820 (2001) (Inspection Upon Request).

122. *See* tit. 105, § 410.750. These violations, along with other violations such as no safe water supply, inadequate exits, inadequate locks, accumulation of garbage or filth which may provide food or shelter for pests or may contribute to accident or disease, structural defects, and the presence of lead paint if a child under the age of six resides in the apartment are considered to be "emergency conditions" because they may endanger or impair the health, or safety of those occupying the premises, and are required to be repaired within five days of the inspection. *See id.*

123. *See* MASS. REGS. CODE tit. 105, § 400.100 (2001).

124. *See* tit. 105, § 400.400. The inspector can cite either the landlord or the tenant or both for violations of the Sanitary Code. *See id.*

125. *See* MASS. GEN. LAWS ch. 239, § 8(A) (2000).

126. *See id.*

8A, as it is currently drafted, provides residential tenants with the most effective means of enforcing the Code.[127]

III. PROPOSED LEGISLATION REQUIRING MANDATORY RENT ESCROWING

The past several years have seen the introduction to the Legislature of the Commonwealth of Massachusetts several different, albeit similar, bills "relative to rent withholding."[128] These bills generally seek to strike the second, third and fourth paragraphs of Chapter 239, 8A, in whole or in relevant part[129] and insert in their place the mandatory requirement of rent

---

127. *See supra* note 47 and accompanying text.

128. House Bill No. 5185 was entitled "Bill Relative to Rent Withholding;" Senate Bill No. 541 was entitled "An Act to Clarify the Law Relative to Rent Withholding;" and House Bill No. 3906 was entitled "An Act Relative to Rent Withholding."

129. *See* MASS. GEN. LAWS ch. 239, § 8(A)'s second paragraph:

> Whenever any counterclaim or claim of defense under this section is based on any allegation concerning the condition of the premises or the services or equipment provided therein, the tenant or occupant shall not be entitled to relief under this section unless: (1) the owner or his agents, servants, or employees, or the person to whom the tenant or occupant customarily paid his rent knew of such conditions before the tenant or occupant was in arrears in his rent; (2) the plaintiff does not show that such conditions were caused by the tenant or occupant or any other person acting under his control; except that the defendant shall have the burden of proving that any violation appearing solely within that portion of the premises under his control and not by its nature reasonably attributable to any action or failure to act of the plaintiff was not so caused; (3) the premises are not situated in a hotel or motel, nor in a lodging house or rooming house wherein the occupant has maintained such occupancy for less than three consecutive months; and (4) the plaintiff does not show that the conditions complained of cannot be remedied without the premises being vacated; provided, however, that nothing in this clause shall be construed to deprive the tenant or occupant of relief under this section when the premises are temporarily vacated for purposes of removal or covering of paint, plaster, soil or other accessible materials containing dangerous levels of lead pursuant to section one hundred and ninety-seven of chapter one hundred and eleven.

Third paragraph:

> Proof that the premises are in violation of the standard of fitness for human habitation established under the state sanitary code, the state building code, or any other ordinance, by-law, rule or regulation establishing such standards and that such conditions may endanger or materially impair the health, safety or well-being of a person occupying the premises shall create a presumption that conditions existed in the premises entitling the tenant or occupant to a counterclaim or defense under this section. Proof of written notice to the owner or his agents, servants,

escrow before a tenant will be allowed to raise a defense of Sanitary Code violations to a summary process action for nonpayment of rent.[130] A recent bill, House Bill 3906, entitled "An Act Relative to Rent Withholding" was submitted by then-Governor of Massachusetts, Argeo Paul Cellucci on February 6, 2001.[131]

> or employees, or to the person to whom the tenant or occupant customarily paid his rent, of an inspection of the premises, issued by the board of health, or in the city of Boston by the commissioner of housing inspection, or by any other agency having like powers of inspection relative to the condition of residential premises, shall create a presumption that on the date such notice was received, such person knew of the conditions revealed by such inspection and mentioned in such notice. A copy of an inspection report issued by any such agency, certified under the penalties of perjury by the official who inspected the premises, shall be admissible in evidence and shall be prima facie evidence of the facts stated therein.

Fourth paragraph:

> There shall be no recovery of possession pursuant to this chapter pending final disposition of the plaintiff's action if the court finds that the requirements of the second paragraph have been met. The court after hearing the case may require the tenant or occupant claiming under this section to pay to the clerk of the court the fair value of the use and occupation of the premises less the amount awarded the tenant or occupant for any claim under this section, or to make a deposit with the clerk of such amount or such installments thereof from time to time as the court may direct, for the occupation of the premises. In determining said fair value, the court shall consider any evidence relative to the effect of any conditions claimed upon the use and occupation of residential premises. Such funds may be expended for the repair of the premises by such persons as the court after a hearing may direct, including if appropriate a receiver appointed as provided in section one hundred and twenty-seven H of chapter one hundred and eleven. When all of the conditions found by the court have been corrected, the court shall direct that the balance of funds, if any, remaining with the clerk be paid to the landlord. Any tenant or occupant intending to invoke the provisions of this section may, after commencement of an action under this chapter by the landlord, voluntarily deposit with the clerk any amount for rent or for use and occupation which may be in dispute, and such payments shall be held by the clerk subject to the provisions of this paragraph.

130. *See Massachusetts Department of Housing and Community Development, Governor's Housing Package at* http://www.state.ma.us/dhcd/Temp/GovPkg/Default.htm (last visited Jan. 25, 2002) [hereinafter *Governor's Housing Package*].

131. *See id.* At the time of publication, the Massachusetts House of Representatives had written a new bill on mandatory rent escrow, House Bill No. 4284, which was included in an act relating to the preservation of affordable housing in Massachusetts and was subsequently submitted to the Senate. The Senate then wrote its own version, Senate Bill No. 2119; both versions are currently in conference committee. *See The Commonwealth of Massachusetts Legislative Tracking*

A. Supporters of the Proposed Bill

The main advocates of this bill, and others like it, have been landlord groups around the Commonwealth led by SPOA.[132] SPOA's effectiveness in lobbying state legislators and the Governor is largely due to a paper it had published by the Pioneer Institute in 1999 entitled "The Road Home."[133] The Road Home exaggerates a small number of stories about landlords who have fallen prey to unscrupulous tenants by holding them out as regular occurrences instead of the anomalies that they actually are.[134]

The Road Home attacked Chapter 239, 8A by stating that the law currently "allows tenants to withhold the entire rent for even the smallest defects. It is an affirmative defense against eviction, blocking eviction until all code violations are repaired."[135] The paper does not explain, however, that any tenant who claims Code violations as a defense to summary process must prove these violations in court. Data shows that claimed Code violations rarely delay the eviction process, let alone result in repairs.[136]

System, *at* http://www.state.ma.us/legis/ltsform.htm (last visited Jan. 25, 2002).

132. *See The Road Home, at* www.pioneerinstitute.org/crg/bgc/invitechange.cfm#2000 (last visited Jan. 15, 2002) [hereinafter *The Road Home*]; *See also* Telephone Interview with Lenore Schloming, *supra* note 4.

133. *See id.* The Pioneer Institute for Public Policy Research defines itself as "a non-partisan, non-profit Boston-based think tank dedicated to challenging conventional wisdom by looking for fresh and innovative solutions to public policy problems affecting Massachusetts." www.pioneerinstitute.org/about (last visited Jan. 25, 2002). Each year the Institute runs "The Better Government Competition" in which it challenges citizens to come up with ideas of how to "improve the quality and cost-effectiveness of governmental services." *Id.* at www.pioneerinstitute.org/crg/competition/crg_competition.cfm (last visited Jan. 25, 2002).

134. *See Evictions, at* http://www.pioneerinstitute.org/crg/competition/crg_competition.cfm (last visited Jan. 25, 2002). In the fiscal year 1997, there were at least 50,000 nonpayment evictions in Massachusetts. *See also A Profile of Housing in Massachusetts: A 1998 Study by the University of Massachusetts,* University of Massachusetts Donahue Institute, at 19 [hereinafter Profile]. The number of eviction cases at the BHC sharply increased after 1994 mainly due to the end of rent control. *See id.* Statistics show that the percentage of landlords who are awarded possession in summary process cases is extremely high, and the amount of time in which the case is disposed of is extremely short. *See* Survey, *supra* note 116 and accompanying text. The fact that so many cases are disposed of in such a short amount of time in favor of the landlord necessarily implies that few tenants manipulate the system to the extent represented by SPOA.

135. *See The Road Home*, *supra* note 135, at 24.

136. *See* Survey, *supra* note 116 and accompanying text. A survey of sum-

The Road Home briefly mentions the Massachusetts consumer protection law (chapter 93A)[137] by stating that its ability to award treble damages and attorneys' fees "for any alleged unfair and deceptive practice, which includes knowingly renting an apartment with code violations," is asserted in "virtually every lawyer-assisted defense against eviction [which] alleges great malfeasance of the landlord regarding the condition of the apartment and makes chapter 93A claims for tens of thousands of dollars . . . ."[138] This statement as it stands is very misleading considering the fact that tenants are represented by counsel in only seven percent of summary process cases.[139]

The Road Home further states that many of the small property owners they refer to delay costly improvements until the last possible minute, "nurs[ing] along . . . leaky roof[s] and old plumbing just to squeeze out a few more years of life before doing big-buck capital replacements."[140] They argue that because many property owners operate "at a close margin,

---

mary process conducted by the Massachusetts Law Reform Institute reports that landlords are awarded possession in 88% of cases. The Survey also reveals that 73% of summary process cases reached a disposition in less than 16 days, 88% of cases reached a disposition in less than 30 days, and that "[a]greements account for the largest percentage of case outcomes, and, in [80% of these] cases, landlords are awarded possession." *See id.* The "agreements" that the Survey refers to are agreements made between the landlord and the tenant outside the courtroom. *See also* Russell Engler, *And Justice For All – Including The Unrepresented Poor: Revisiting the Roles of the Judges, Mediators, and Clerks*, 67 FORDHAM L. REV. 1987, 1989 (1999) (noting that significantly important rights were regularly waived "by unrepresented litigants. Unrepresented litigants appearing before mediators suffer a similar fate, after being deemed to have made an informed choice to participate in mediation in the first place. Where cases do not settle, unrepresented litigants also routinely forfeit important rights due to the absence of counsel").

137. *See* MASS. GEN. LAWS ch. 93A (2000).

138. *See The Road Home*, *supra* note 132, at 24.

139. *See* Kurtzberg, *supra* note 116 and accompanying text.

140. *The Road Home*, *supra* note 132, at 23. *The Road Home* goes on to applaud the practices of these small property owners by congratulating them on their style of management, which allows them to reduce monthly rental rates. *The Road Home* then warns of the dangers involved in this style of management due to the "difficult state laws: the state sanitary code, rent withholding, and the warranty of habitability." *Id.* at 23. These laws were all instituted in an attempt to empower the low-income tenant by forcing the landlord to keep the rental units in habitable condition. *See supra* note 21 and accompanying text. Landlord groups in favor of the proposed legislation view these laws as the obstacle which prevents them from managing their units in a way which would cut down on costs – if few or no repairs are made to the rental units, a large amount of the monthly rent is pure profit leaving less incentive to the owner to raise the rent. *See The Road Home*, *supra* note 132, at 23. This style of management is largely the reason the state enacted the sanitary code. *See supra* Part II, Section E and accompanying text.

[they] cannot possibly improve their property without a steady and reliable rent stream. Yet legally, owners are often helpless to evict non-paying, damaging or uncontrollable tenants."[141] A tenant, however, is not allowed to withhold rent for any reason other than those listed in Chapter 239, 8A, and a tenant who has committed waste or other damage to the apartment, or is being evicted for cause[142] cannot raise Chapter 239, 8A defenses and cannot win possession of the rental unit.[143] The Road Home considers current law to be in effect "housing theft,"[144] asserting that it allows tenants (consumers) to retain the defective product (the apartment with code violations) after receiving a refund (possession of the apartment) for the defect.[145] The Road Home asserts, contrary to black letter law,[146] that if a tenant contracts to rent an apartment with code violations, the agreed upon rent will be reduced because of the violations and, therefore, the tenant should have no legal basis for complaining about those code violations.[147] This type of behavior is exactly what the Supreme Judicial Court sought to abolish in *Hemingway*.[148] The purpose of Chapter 239, 8A and the sanitary

---

141. *The Road Home, supra* note 132, at 23.

142. *See* RESIDENTIAL LANDLORD/TENANT BENCHBOOK, 71, Flaschner Judicial Institute (1998). There are basically three types of evictions: (1) nonpayment of rent evictions; (2) fault (for cause) evictions that occur when the tenant does something wrong; and (3) no-fault evictions that occur when the landlord evicts the tenant for some other reason. An eviction for cause basically occurs when a tenant commits waste which is "defined as an unreasonable or improper use, abuse, mismanagement, or omission of duty touching real estate by one rightfully in possession which results in substantial injury to the premises." *Id. See also, e.g.*, Boston Housing Authority v. Guirola, 575 N.E.2d 1100 (Mass. 1991); Hodess v. Bonefont, 519 N.E.2d 258 (Mass. 1988); Spence v. Gormley, 439 N.E.2d 741 (Mass. 1982); Spence v. O'Brien, 446 N.E.2d 1070 (Mass. App. Ct. 1983); and Spence v. Reeder, 416 N.E.2d 914 (Mass. 1981).

143. A tenant can win possession only if the tenancy was terminated for nonpayment of rent or terminated without fault of the tenant (provided the tenant can prove his case). *See* MASS. GEN. LAWS ch. 239, § 8(A) (2000).

144. *The Road Home*, *supra* note 132, at 25.

145. *See id.*

146. *See* Boston Housing Authority v. Hemingway, 293 N.E.2d 831, 843 (Mass. 1973). The law clearly states that the warranty of habitability cannot be waived in exchange for lower rent. *Hemingway* also clearly states that the underlying purpose of 239, 8A is to create a remedy for the tenant to effectively enforce the state sanitary code, recognizing that this remedy is especially important for poor tenants who do not have easy access to legal representation. *See id.* at 839.

147. *See The Road Home*, *supra* note 132, at 25.

148. *See Hemingway*, 293 N.E.2d at 839. The SJC was fully aware of the inadequacies of tenant remedies prior to its decision in *Hemingway*: "[u]nlike the more cumbersome procedure established in c. 111, §§ 127A-127H, where the landlord can use delaying tactics by asking for repeated continuances, c. 239, § 8A's provisions force him to take the initiative by making repairs . . . ." *Id. See*

code is to provide every tenant with a habitable dwelling, and to force landlords to make repairs in the event that the dwelling does not comply with the code.[149]

The Road Home goes on to give its own "calculation" as to how much money is lost to the practice of rent withholding.[150] This number is not supported by any economic evidence, nor any evidence of what research was conducted to produce such a number.[151] In fact, most cases are settled outside of the courtroom by agreement[152] making it virtually impossible to track with any degree of certainty whether the tenant or the landlord comply with the agreement. Anyway, the loss is not to "rental housing" but to owners of rental housing. This simply means that the owners sustain a decrease in their profit margin, it is not indicative of what amount is put into the housing stock.[153] The argument implies that in order to maintain an adequate supply of low-income housing landlords should not be required to comply with the State Sanitary Code.[154] Tenants receiving Section 8 vouchers, however, live in low-income housing compliant with the Sanitary Code.[155]

---

also Haddad v. Gonzalez, 576 N.E.2d 658 (Mass. 1991). The SJC specifically held that it was not relevant whether or not the tenant paid less for the apartment because of the Code violations: "[t]he rule for computation of damages in this area gives a tenant no more than the benefit of the bargain because the implied warranty of habitability is part of the bargain. A landlord cannot nullify the implied warranty of habitability on a dwelling by giving his tenant a discount in rent." *Id.* at 668 (citations omitted).

149. *See Hemingway*, 293 N.E.2d at 839.

150. *See The Road Home*, *supra* note 132, at 26 stating "Massachusetts rental housing loses at least $78 million every year to the practice of rent withholding and its delay of eviction for non-paying tenants." *Id.* *The Road Home* goes on to maintain that half of that amount is unpaid rent, which is never recovered. *See id.*

151. *See The Road Home*, *supra* note 132, at 26 n.22.

152. *See* Fox, *supra* note 12, at 91. Most nonpayment of rent cases are settled outside of the courtroom, in fact, a survey conducted at the Boston Housing Court in the Spring of 1995 revealed that 80% of scheduled cases were handled "either through mediation with a housing specialist or through negotiation in the hallway." *Id.* *See also supra* note 134 and accompanying text.

153. *See* Edward H. Rabin, *The Revolution in Landlord-Tenant Law: Causes and Consequences*, 69 CORNELL L. REV. 517, 561 (1984) (discussing the economic ramifications of Sanitary Code enforcement). The predominant view among academic scholars seems to be that enforcement of the Sanitary Code will result in rent increases to cover costs and to maintain profit margins and the low-income housing stock will decrease hurting the low-income population. *See id.* at 558-63.

154. *See The Road Home*, *supra* note 132, at 23-25.

155. Section 8 Voucher program participants are required to submit to a yearly inspection by the administering agency. *See Section 8 Program Fact Sheet, at* http://www.hud.gov/section8.cfm (last visited Mar. 11, 2001). If the unit does

The Road Home further contends that tenants use the law to delay the eviction process by claiming sanitary code violations, which cost the landlord three months in lost rent.[156]

The law states that a landlord must terminate a tenancy before serving the tenant with a summons and complaint.[157] If the landlord does not want to give the tenant the right to "cure" the arrearage and restore the tenancy, he must send a rental period NTQ.[158] It is only after the expiration of the thirty days (or the fourteen days) that the landlord will be allowed to serve the tenant with a Summons and Complaint for summary process.[159] This process is called "summary process" because it is designed to give the landlord a quick and easy remedy for evicting a tenant.[160] The law is designed to protect both parties from abuse. The laws of summary process can be abused by both the tenant and the landlord. The fact that the potential for abuse exists, however, should not determine how any particular law is rewritten, rather, it should influence the protocol and performance of the agencies designed to prevent the abuses.

One of the solutions offered in *The Road Home* for the above-mentioned abuses (which they call "housing theft")[161] is mandatory rent escrowing.[162] This "solution" would prevent a tenant who is withholding rent due to Code violations from subsequently raising the violations as a defense to a nonpayment of rent eviction if the tenant does not deposit that withheld rent into an escrow account.[163] SPOA supports its position by maintaining that mandatory rent escrowing will separate those tenants legitimately withholding rent due to Code violations from those who are simply ma-

---

not pass the inspection, and the landlord does not remedy the Code violations, the administering agency may terminate the rental contract. *See id.*

156. *See The Road Home*, *supra* note 132, at 26-27.

157. *See* MASS. GEN. LAWS ch. 186, §§ 11, 12, 13, 15A, 17 (2000). A notice to quit (hereinafter "NTQ") for nonpayment of rent for a tenant at will must be at least fourteen days prior to eviction. *See id.* at § 11. This NTQ must contain language that informs the tenant of his right to "cure" the arrearage which, if done in a timely manner, will restore the tenancy. *See id.* at § 12. In order to terminate a tenancy-at-will for cause or for no cause, the landlord must send a rental period NTQ, which must terminate the tenancy on a date that rent is due. *See id.* at §§ 12, 17.

158. *See id.*

159. *See* UNIF. SUMM. PROCESS R. 2(b).

160. *See* Fox *supra* note 12, at 90-91 n.21. In her article, Fox points out the paradoxical nature of the summary proceeding – "[o]n the one hand, they grew out of a need to protect tenants from physical ejectment from their homes. . . . [o]n the other hand, they are by their mandate intended to serve landlords." *Id.*

161. *See The Road Home*, *supra* note 132, at 25.

162. *See id.* at 27.

163. *See id.*

nipulating the system.[164] The flaw in this argument is that SPOA presumes that all tenants who legitimately withhold rent, meaning those tenants who actually live in apartments with Code violations, save the money and have it available to them at any time. In reality, an overwhelming majority of tenants who face nonpayment evictions are low-income people who do not have easy access to legal counsel.[165] Statistics show that those tenants who are represented by legal counsel have a much higher success rate in terms of winning possession of the apartment.[166] The proposed amendment, as do its supporters, ignore the fact that most tenants are not represented by legal counsel and would have enormous difficulty complying with the proposed provisions without legal counsel.

Whether a tenant is legitimately withholding rent due to Code violations depends on the condition of the apartment, not the status of the tenant's bank account. It is illogical and improper to attempt to differentiate those legitimately withholding rent on the basis of whether they have the withheld money at the time of trial. A more appropriate basis would be, and is, the presentation of an inspection report from ISD, since the legal basis of rent withholding is the existence of Sanitary Code Violations in the subject premises.

The attempt to differentiate those legitimately withholding rent based on whether the money has been set aside is improper because it is not a rational contract-based theory. In other contract disputes, the party in breach is not made to post money with the court before receiving a trial on the merits. If, for example, a merchant accepted goods pursuant to a contract and failed to pay, the opposing party may not automatically presume that the reason is lack of funds. In fact, there could be any number of reasons for nonpayment such as the goods were defective, or they were not the goods ordered or expected. If the party who was not paid chose to sue, the breaching party (the party that did not pay) would not be expected or required to post the alleged amount due with the court before receiving a trial on the merits.

The Road Home further suggests that the amount deposited in escrow should be the contract rent of the apartment.[167] The flaw in this suggestion

---

164. *See* Schloming Telephone Interview, *supra* note 4.

165. *See* Kurtzberg, *supra* note 116 and accompanying text.

166. According to the Massachusetts Law Reform Survey, tenants statewide are represented by counsel in 7% of summary process cases; tenants statewide are awarded possession in 12% of summary process cases. *See* Survey, *supra* note 116, at 6-9. These statistics necessarily indicate that tenants who are represented by counsel have a much greater chance of winning possession.

167. *See The Road Home*, *supra* note 132, at 28.

is best articulated by Judge Daher in the *Slack* case.[168] The contract rent is an improper measure of what, if any, money should be deposited with the court because Chapter 239, 8A provides that the covenant to pay rent and the covenant to maintain the apartment are dependent on each other.[169] If the apartment does not meet the standards set forth in the Sanitary Code, then it is not worth the contract rent, rather it has a fair market value of a similar apartment with the same Code violations.[170] It would be preposterous to assert that an apartment with an infestation of rats and roaches has the same fair market value as one free of such an infestation. Therefore, a more accurate measure of what, again if any, money should be deposited with the court is the fair market value, which should be determined by a judge at an evidentiary hearing.[171]

The Road Home's position is that mandatory rent escrowing is a safeguard against abuse, and the "device of rent withholding would be turned into a bona fide mechanism to enforce owners to do demanded repairs."[172] Most tenants facing summary process proceedings are not represented by legal counsel.[173] Currently, over 80% of summary process cases end up in some type of non-judicial resolution.[174] Mandatory rent escrow will not better arm an unrepresented tenant who is poor and may not even speak English, nor will it alter the practice of housing courts.[175] Mandatory rent escrow will further disadvantage tenants, in what is already a severely lopsided dispersion of power. What will happen in most cases is that the tenant will have no defense, the case will be disposed of in favor of the landlord (as is already the case in over 80% of summary process cases),[176] and no repairs will be ordered or required before a new tenant moves in.

The Road Home further maintains that the practice of mandatory rent escrow will have "no untoward effects on tenants," by alleging that it would shift the tenants and their advocates (for those 7% who are fortunate enough to have legal counsel) away from litigation and toward mediation.[177] Currently, 95% of cases in Boston Housing Court go straight to mediation.[178] *The Road Home*'s position suggests that there is no need

---

168. *See* Slack, 74-SP-02335 at 9; *see also supra* note 68 and accompanying text.

169. *See Hemingway*, 293 N.E.2d at 840.

170. *See supra* note 53 and accompanying text.

171. *See supra* notes 52-58 and accompanying text.

172. *See The Road Home*, *supra* note 132, at 28.

173. *See supra* note 116 and accompanying text.

174. *See supra* note 116, 136 and accompanying text.

175. *See supra* notes 116, 152 and accompanying text.

176. *See supra* note 116 and accompanying text.

177. *See The Road Home*, *supra* note 132, at 28.

178. *See* Interview with Harvey J. Chopp, Esq., Court Administrator, The

for housing courts at all; by suggesting the favored "shift" from litigation to mediation as an "untoward effect" on tenants is simply absurd considering the fact that over 80% of cases do not result in any litigation.[179]

The truth is that SPOA, and other landlord groups, are seeking protection from tenants who do not pay their rent.[180] They argue that the practice of rent withholding can send a landlord into bankruptcy or can cause other economic hardship.[181] The answer to this problem however, is not rent escrowing. Rent escrowing will only deprive a tenant who has not complied with the mandatory provisions of a defense to the summary process action. Depriving a tenant of a defense is a last resort in the grand scheme of summary process.[182] The deprivation of a defense may ultimately result in a quicker disposition of the case at the trial level, but it will not solve the insufficiency of funds problem. In fact, a quicker disposition of the case at the trial level will be hard to accomplish as current research shows that over 70% of cases, from entry to disposition, last less than 16 days.[183] The arrearage, which in the end may result in foreclosure or the like, is not usually the result of a delay at the trial level. Rather, it is the landlord's delay in instituting the action that results in the greatest of arrearages.[184] In over 50% of nonpayment cases, the tenant is in arrears in excess of $1,000 before a notice to quit is even served by the landlord.[185] Rent escrowing is

---

Commonwealth of Massachusetts, Housing Court Department (Feb. 22, 2001).

179. *See* Survey, *supra* note 116. The Survey reveals that only 4% of summary process cases statewide result in judgments. *See id.* at 11.

180. Schloming Telephone Interview, *supra* note 4.

181. *See The Road Home*, *supra* note 132, at 23.

182. The road to Housing Court is a short one. In order to initiate a proceeding, a landlord must first terminate the tenancy if termination is not provided for otherwise in a written lease. *See* MASS. GEN. LAWS ch. 186, § 13 (2000). It is only after the expiration of this period (either fourteen days or thirty days) that a Summons and Complaint can be served on the tenant. *See generally* MASS. UNIF. SUMM. PROCESS R. The landlord decides what day the hearing will be held and writes it on the Summons and Complaint. *See id.* at Commentary to Rule 2. The hearing is held ten days from the entry of the Complaint, and the answer date is one week from the entry date. *See id.* at Rule 2(c). Monday is the entry day for summary process cases in Boston Housing Court. *See id.* In short, a landlord can be in and out of court (after winning possession of the unit) in less than two months from the time he decides to evict his tenant. *See id.*

183. *See* Survey, *supra* note 116, at 5.

184. I reviewed a small sample of cases at the Boston Housing Court [hereinafter Sample]. My research further showed that of all the nonpayment cases I reviewed, over 70% of the tenants had arrearages in excess of $1,000 by the time they were served with a Summons and Complaint. While this may not seem like a large amount, the tenants had rental subsidies and paid less than $400 per month in rent. *See id.*

185. *See id.*

not likely to impact a landlord's decision of when to initiate summary process proceedings; it is this delay that usually causes the economic burden on the property owners.[186]

B. "An Act Relative to Rent Withholding"

On May 23, 2001, the Massachusetts Joint Committee on Housing and Urban Development took testimony on a bill submitted by then Governor Celluci entitled "An Act Relative to Rent Withholding."[187] This bill sought to amend Massachusetts General Laws chapter 186 Section 14 by inserting a subsequent section, devoted to rent withholding and rent escrowing by a tenant.[188] Section 14 of Chapter 186 is entitled "Penalty on Lessor for Failure to Furnish Water, Heat, etc.; Waivers by Tenant Prohibited."[189] This statute addresses the covenant of "quiet enjoyment."[190] The proposed amendment required that, in the event of a material violation of the State Sanitary Code the tenant give written notice of his intention to withhold rent in escrow.[191] The amendment further required that this notice shall "identify each material violation or defect with specificity and shall be mailed or delivered to the landlord at least fourteen days prior to the due date for the first rental payment to be withheld . . . on account of such material violation or defect."[192]

Currently, the law does not require the tenant to give the landlord written notice of conditions of disrepair in the apartment; if the tenant has a problem a telephone call is sufficient to put the landlord on notice.[193] The issue of notice is of the utmost importance in a summary process case. One of the essential elements of a Chapter 239, 8A defense is that the landlord had notice of the conditions before the tenant began to withhold the rent.[194] There is currently no requirement that the tenant call ISD, or any other

---

186. *See id.*

187. *See* House Bill 3906 *at Governor's Housing Package*, *supra* note 130. The bill was submitted on February 6, 2001.

188. *See id.* at Section 1.

189. *See* MASS. GEN. LAWS ch. 186, § 14 (2000). Chapter 186 in general deals with "Estates for Years and At Will." *See id.*

190. *See supra* Section B and accompanying text.

191. *See Governor's Housing Package*, *supra* note 130.

192. *Id.*

193. *See* MASS. GEN. LAWS ch. 239, § 8A (2000). Section 8A does not require that the tenant give the landlord written notice of any conditions of disrepair. *See id.* The provision simply states that if the premises are in violation of the standard of fitness for human habitation set out in the State Sanitary Code, or any other building code or ordinance, proof that the violation existed at any one particular time can be presumed if a written notice is issued by the Board of Health. *See id.*

194. *See id.*

agency, in order to raise a Chapter 239, 8A defense; doing so simply eases the tenant's burden at trial.[195] Requiring the tenant to give the landlord written notice of any condition or disrepair, aside from an inspection report, would be unduly burdensome for the tenant.[196]

The amendment further requires the landlord to promptly remedy each material breach set out in the written notice.[197] If the landlord deems the breach immaterial, he may make arrangements for an inspection by the local Board of Health or other local code enforcement agency.[198] This is somewhat of a role reversal; usually the tenant is the party who contacts ISD to request an inspection. In the event that an inspection reveals no material breach of the State Sanitary Code the tenant will not be allowed to further withhold rent, and any rent withheld up until that point (in escrow) will thereafter be released.[199]

A most burdensome provision of this section of the bill is the requirement that the tenant deposit the withheld money with the Clerk Magistrate of the Division of the District Court Department, Municipal Court Department, or Housing Court Department within the jurisdiction of which the apartment is located.[200] This is burdensome because it requires the tenant to know which court is the appropriate court.[201] It also fails to specify what consequences will result if in fact the tenant chooses the wrong court, which is likely to happen considering that the landlord is the one who chooses the court in a summary process case.[202] The logical chronology of the process of rent withholding occurs, by definition, before the landlord has terminated the tenancy or served the tenant with a summary process Summons and Complaint. If, for example, the tenant deposits the withheld rent at the BHC, and the landlord subsequently serves the tenant with a summary process Summons and Complaint for West Roxbury District

---

195. *See* MASS. GEN. LAWS. ch. 239, § 8A (2000).

196. *See supra* note 116 and accompanying text.

197. *See Governor's Housing Package*, *supra* note 130.

198. *See id.*

199. *See id.*

200. *See id.*

201. A landlord may bring a summary process case in either the appropriate superior court, district court, Boston Municipal Court (hereinafter the "BMC") or an appropriate housing court, provided one exists, in the appropriate jurisdiction. *See* Rule 4 of the UNIFORM SUMMARY PROCESS RULES (stating that a party wishing to transfer a summary process action pending in another department other than a housing court department, shall file the appropriate form with both courts (the division in which it was filed, and the housing court department)).

202. A tenant can bring an affirmative case for damages in superior court, district court, BMC, or the housing court department, and is bound by the same procedural rules.

Court, what happens to the money sitting at the BHC? Will the transfer of funds delay the proceeding? These issues could possibly result in the opposite effect of the one desired, by causing more delays in the summary process proceeding.[203]

The tenant's right to withhold rent after the initial proceeding is conditioned upon the tenant providing reasonable access to the landlord in order to make repairs.[204] The right to withhold rent pursuant to the proposal is considered a right in addition to the tenant's rights pursuant to chapter 111, § 127L.[205]

The proposed amendment to chapter 186 did not specifically state, unlike the proposal in *The Road Home*, that the amount to be deposited with the court will be the contract rent.[206] However, what the proposed amendment did state was detrimental to Chapter 239, 8A. The amendment had the following provision: "[i]n the event that the landlord gives a notice of correction to the tenant as provided herein, the landlord may offer the tenant an amount suitable to compensate the tenant for diminution in the value of the premises while the violations and defects exist."[207] This provision completely changes the meaning of Chapter 239, 8A, the State Sanitary code, the consumer protection laws and the general evolution of landlord-tenant law in Massachusetts.[208] It further disregards the current judicial capacity to abate the arrearage in a summary process case for nonpayment of rent in which a tenant has a Chapter 239, 8A defense.[209]

First, the provision made no reference to what consequences the landlord would suffer if he failed to promptly remedy the violations, it simply stated that in such a case the tenant is allowed to further withhold rent.[210] In reality, a tenant can complain to the landlord about a rodent infestation, give the landlord two weeks to remedy the situation and upon his failure to do so, withhold rent. However, if the landlord refuses to exterminate for six months, the tenant may not by law, be entitled to any abatement of the rent for that period, absent any deduction in the rent the tenant may make

---

203. *See supra* note 160 and accompanying text.

204. *See Governor's Housing Package*, *supra* note 130, Section 1.

205. *See* MASS. GEN. LAWS ch. 111, §127L (2000) entitled "Right of Tenant to Reimbursement by Landlord for Certain Repairs to Cure Violations; Right of Tenant to Treat Lease as Abrogated, Pay Fair Rental Value and Vacate Premises; Waiver of Section Against Public Policy; Exception." Chapter 111 deals with Public Health Nuisances. *See id.*

206. *See Governor's Housing Package*, *supra* note 130, Section 1.

207. *Id.*

208. *See supra* Section II and accompanying text.

209. *See generally* Boston Housing Authority v. Hemingway, 293 N.E.2d 831 (Mass. 1973) and MASS. GEN. LAWS ch. 239, § 8A.

210. *See Governor's Housing Package*, *supra* note 130, Section 1.

as a result of exterminating on his own. This is contradictory to the principle that the covenant to pay rent is dependent on the landlord's covenant to maintain the property as required by the State Sanitary Code.[211] If the landlord is permitted, at his leisure, to make repairs with no penalty imposed for his failure to do so promptly, this amendment is nothing more than a step in the direction of eradicating Chapter 239, 8A completely.

Part 2 of Section 1 of the amendment set forth the manner in which the landlord is required to notify the tenant of the mandatory rent escrowing provisions.[212] This section gives the landlord the responsibility of informing the tenant which court has jurisdiction over the property.[213] The provision fails, however, to recognize that residents of the City of Boston, as well as many other areas of the state, have at least three different courts to which they could turn.[214] It also fails to address the issue of language. A recent study conducted by the Donahue Institute at the University of Massachusetts reveals that two-thirds of the Commonwealth of Massachusetts' racial/ethnic minority residents are "exclusively dependent on the residential rental market for housing."[215] If the landlord rents to non-English speaking tenants, is he required to provide the information on rent escrowing in a language the tenant will understand?[216] The only saving grace of this section of the amendment is that a tenant will be allowed to raise a Chapter 239, 8A defense to a nonpayment summary process action despite failure to comply with part 1 of Section 1 if the landlord fails to notify the tenant of the escrow requirement provided, however, that the court finds

---

211. *See* Berman & Sons, Inc. v. Jefferson, 396 N.E.2d 981, 984 (Mass. 1979).

212. *See Governor's Housing Package*, *supra* note 130, Section 1.

213. *See id.*

214. *See supra* note 201.

215. *See Profile*, *supra* note 134. The Study also reveals that Latino homeownership in Massachusetts is only 18.7%, while the percent of Latino homeownership nationally is 42.4%. *See id.* at 8. *See also* the Department of Housing and Community Development's HUD Community Planning and Development 2000-2004 Consolidated Plan and 2000 One-Year Action [hereinafter Consolidated Plan] (stating that the large immigration of foreign-born households into Massachusetts has resulted in a growing number of non-white immigrant households where English is the second language). Further, 41% of these immigrants do not have a high school diploma. *See id.* at 1-4. This immigrant population also has a significant portion of households of lower income with a significant number whose only household head is a single female. *See id.* The Study's conclusion is that the number of immigrants making up families in poverty is rising. *See id.*

216. The most recent bill drafted by the House committee on Bills in the Third Reading included a provision requiring that "the landlord . . . provide to the tenant a written statement . . . of tenant rights and responsibilities in English and in the tenant's primary language." *See* H.R. 4284 § 23, 3d Read. (Mass. 2001).

that the tenant did not act in bad faith in withholding the rent.[217]

Section II of the proposed amendment sought to strike the second and fifth paragraphs, and portions of the third and fourth paragraphs of Chapter 239, 8A.[218] The most startling change in the bill is that it literally sets the Commonwealth's progress in landlord-tenant law back over twenty-five years.[219] This legislation proposes that which the *Hemingway* court sought to avoid. This proposal seeks to take the Commonwealth back to times

---

217. *See Governor's Housing Package*, *supra* note 130, Section 1, part 2.

218. *See* MASS. GEN. LAWS ch. 239, § 8A. The proposed version of Chapter 239, 8A reads in pertinent part as follows (including only the first and second paragraphs):

> Except as hereinafter provided in any action under this chapter to recover possession of any premises rented or leased for dwelling purposes, brought pursuant to a notice to quit for nonpayment of rent, or where the tenancy has been terminated without fault of the tenant or occupant, the tenant or occupant shall be entitled to raise, by defense or counterclaim, any claim against the plaintiff relating to or arising out of such property, rental tenancy, or occupancy for breach of warranty, for a breach of any material provision of the rental agreement, or for a violation of any other law. The amounts which the tenant or occupant may claim hereunder shall include, but shall not be limited to, the difference between the agreed upon rent and the fair value of the use and occupation of the premises, and any amounts reasonably spent by the tenant or occupant pursuant to section one hundred and twenty-seven L of chapter one hundred and eleven and such other damages as may be authorized by law having as its objective the regulation of residential premises.
>
> In the event that the owner, the landlord or the landlord's authorized representative or agent initiates summary process proceedings for nonpayment of rent pursuant to this section, the tenant or occupant shall have a defense to these proceedings for nonpayment of rent with respect to any amounts of rent, which have been withheld and escrowed on account of material defects or violations affecting the habitability or safety of rental housing in substantial accordance with the provisions of section 14A of chapter 186 of the general laws, and with respect to any amounts properly paid for repairs pursuant to the provisions of section 127L of chapter 111 of the general laws, provided, however, that if the tenant has not acted in good faith in withholding and escrowing rent or if the landlord has been denied reasonable access by the tenant for the purpose of inspection of conditions and corrective measures, the court may find this defense not to be available to the tenant. In the event a tenant or occupant has not substantially complied with the provisions of section 14A of chapter 186, or with the provisions of section 127L of chapter 111 he shall have no defense, on account of material defects or violations affecting the habitability or safety of rental housing, to summary process proceedings for nonpayment of rent, but his lack of a defense to summary process for nonpayment of rent shall not affect the merits of any counterclaim or action for damages or other relief which the tenant or occupant may have regarding any such defects or violations affecting the habitability or safety of rental housing.

Governor's Housing Package, *supra* note 130.

219. *See supra* notes 37-44 and accompanying text.

when caveat lessee was the norm.[220]

The fourth paragraph of the newly proposed Chapter 239, 8A ambiguously mentions the amount of rent to be deposited with the court by stating that "[i]n determining fair value of use and occupation of the premises, the court shall consider any evidence relative to the effect of any conditions claimed upon the use and occupation of the residential premises."[221] This provision however, only comes into effect once the case gets to the trial court level -- it has no effect on what the tenant is required to deposit before, or if ever, a summary process proceeding is instituted. This provision is somewhat misleading in that it purports to offset the harshness set out in the amendment to chapter 186; however, as mentioned, it only gives the judge the authority to determine the fair value if a summary process proceeding is actually initiated.[222]

The proposed bill (An Act Relative to Rent Withholding) is wrong, unnecessary and inefficient. Chapter 239, 8A, as it is currently drafted, allows judges to order that money be deposited into an escrow account before making a final judgment.[223] The main goal of the proposed legislation is to protect landlords from tenants who "unreasonably withhold their rent."[224] Mandatory rent escrow is ideally supposed to prevent economic disincentives and disinvestments.[225] Mandatory rent escrow, however, will not solve the real problem which is the delay in terminating a tenancy.[226] The only other delay which practically poses problems is the two-week postponement of the trial which occurs automatically when a tenant requests discovery along with an answer.[227] This problem could be solved if Massachusetts were to adopt a rule similar to the one found in the Federal Rules of Civil Procedure which would allow for automatic discovery, thereby eliminating the need for the two-week postponement.[228]

---

220. *See id.*

221. *See Governor's Housing Package*, *supra* note 130, Section 4.

222. *See id.*

223. *See* Jackson v. Grant and Jackson, 98-SP-01144 (Mass. Dist. Ct. 1998); Gracisse v. Joseph and Sims, 97-SP-05717 (Mass. Dist. Ct. 1997); Parkview Properties v. Diesenhof, 95-SP-9500414 (Mass. Dist. Ct. 1995). *But see* Dove v. Doyle, 95-SP-01915 (Mass. Dist. Ct. 1995) (money required to be deposited only after hearing).

224. Governor Celluci's letter to Massachusetts Legislators (Feb. 6, 2001), *Governor's Housing Package*, *supra* note 130.

225. *See The Road Home*, *supra* note 132, at 22-24. SPOA's position is that without a steady stream of income from rent, owners are not able to make repairs and are in fact discouraged from investing in the property. *See id.*

226. *See* Sample, *supra* note 184 and accompanying text.

227. *See* UNIFORM SUMMARY PROCESS RULES, Rule 7(b).

228. *See* FED. R. CIV. P. 26.

C. "Pretrial Rent Deposit Program"

In September of 1994, the Governor of California signed a bill entitled "Pretrial Rent Deposit Pilot Project."[229] This bill was proposed and subsequently passed as a result of landlord interest groups' lobbying efforts for a faster eviction system.[230] The project was designed to be a "[t]est"[231] and was only implemented in certain courts throughout the state.[232] In the designated "test" areas, landlords were able to choose to prosecute an eviction case through either the Pilot Project or through the traditional unlawful detainer system (California's version of summary process).[233] The Pilot Project's supporters were concerned with the slow-down of the unlawful detainer system, and with what the L.A. Times described as "Tenants from Hell."[234] Like the summary process system in Massachusetts, the unlawful detainer system was designed to aid landlords in expediting the eviction process in order to gain possession in a matter of weeks.[235] A California eviction manual[236] articulates that landlords desire to evict tenants as soon as possible because most tenants are usually "indigent [and] judgment-proof . . . [rendering] an award of damages of rent to be meaningless."[237] The landlord groups in California were particularly troubled by what they described as "delay tactics" tenants were utilizing, mainly by frivolously

---

229. Cal. S.B. 690, 1995-1996 Leg., Reg. Sess. (Kopp 1994).

230. *See* Bruce Zucker, *California Pretrial Rent Deposit Pilot Project: A Legal and Empirical Analysis of the System in Action in the Los Angeles Municipal Court (Central Division)*, 35 CAL. W. L. REV. 159, 162 (1998).

231. *See id.* The legislation stated that between July 1, 1995 and July 1, 1999, certain designated southern California courts would participate in the experimental pretrial rent deposit program. *See* CAL. CIV. PROC. CODE § 1167.2 (West 1982 & Supp. 1998) *repealed by* Stats.1998, c. 931 (S.B.2139), § 119.

232. *See id.*

233. *See* Zucker, *supra* note 230, at 170. The main difference between the two systems is the requirement that a pre-trial hearing be held prior to the tenant receiving a trial on the merits. *See id.* "The primary function of the pretrial hearing is for the court to make a determination as to whether the landlord has pled a prima facie case evidencing a cause of action in unlawful detainer and whether the tenant has any basis, along with supporting evidence, for defending the action." *Id.* at 172. Basically, if the landlord proves his prima facie case, and the tenant is not likely to prevail, the court will order the tenant to deposit money with the court if the tenant wishes to proceed with a trial. *See id.*

234. Stephanie O'Neill, *Tenants from Hell: Professional Deadbeats, "Petition Mill" Scam Artists Imperil Small Rental Property Owners Unfortunate Enough to Select Them as Renters*, L.A. TIMES, Aug. 8, 1993, at K1.

235. *See* Zucker, *supra* note 230, at 160.

236. *See* MYRON MOSKOVITZ, CALIFORNIA EVICTION DEFENSE MANUAL (1998).

237. *Id.* at 1.2.

invoking the IWH defense.[238]

The Pretrial Rent Deposit Program was codified into law on July 1, 1996.[239] The Pilot Project's main goal was to reduce the time in which a residential landlord can regain possession of the rental unit by requiring that tenants comply with certain pre-trial procedures, including the "presentation of an offer of proof at a pre-trial hearing, as well as posting up to two-weeks' rent with the court, in order that the tenant may receive a trial on the merits."[240] The California legislation is similar to the Massachusetts bill relative to rent withholding, however, the California legislation covers many areas that the Massachusetts bill does not. For instance, the California legislation specifically details how the pretrial hearing will be conducted,[241] how and when attorneys may participate in the process, and how much rent will be deposited upon order of the court.[242]

The California legislation, although more detailed than its Massachusetts counterpart, is not flawless.[243] More importantly, however, an empirical study[244] of the two systems showed that the Pilot Project failed to meet its intended goals.[245] The empirical study also noted that the "most signifi-

---

238. *See* Zucker, *supra* note 230, at 162-63.

239. *See id.* at 162.

240. *Id.* The goal was to reduce the time from filing the complaint to regaining possession by 50%. *See* CAL. CIV. PROC. CODE § 1167.2 (West 1982 & Supp. 2001) *repealed by* Stats.1998, c. 931 (S.B.2139), § 119.

241. *See id.* The pretrial hearing must be held between eight and thirteen days after the landlord files proof of service of the summons and complaint. *See id.* The pretrial hearing is designed to be brief and informal – the court may only consider declarations, documentary evidence, the parties' oral testimony, and the defendant's answer or other response to the complaint (not the counterclaims). *See id.* The goal of the hearing is to "dispense justice properly, fairly, and inexpensively." CAL. CIV. PROC. CODE § 1167.2 (West 1982 & Supp. 2001).

242. *See id.* The legislation states that attorneys are generally excluded from the pretrial hearing in Los Angeles (one of the designated courts). *See id.* If the court does make an order requiring rent to be deposited, it must be done within two days after the pretrial hearing. *See id.* If the tenant fails to make the deposit, judgment for possession is entered for the landlord. *See id.*

243. *See* Zucker, *supra* note 230, at 169 (noting that the statute does not appear to be "thoroughly thought through by its authors . . . its wording has ambiguities and it contains provisions that appear internally inconsistent").

244. Zucker's study looked only at cases filed in the Los Angeles Municipal Court [hereinafter LAMC] in the calendar year 1997. *See id.* at 175.

245. *See id.* at 188. The reduction in time was only 14.8%, not 50% as proposed. *See id.* In addition to the goal of reducing time, other objectives included the following: no more than five percent of tenants who lose should file notices of appeal; no more than one percent of the rental units should have outstanding Code violations; the total cost of processing Pilot Project cases must be equal to or less than the cost to process the traditional unlawful detainer cases; and the new system

cant measure of success turn[ed] on whether 'due process' protections for tenants [were] maintained through the implementation of the Pilot Project eviction system."[246] These protections were already very limited in the traditional unlawful detainer system.[247] The results of the study show that a program such as the Pilot Project will only serve to further narrow a tenant's procedural rights, and therefore cannot be said to "maintain" due process protections.[248]

Only 347 landlords chose to participate in the Pilot Project in LAMC, whereas 30,850 were filed in the traditional unlawful detainer system.[249] It is not clear, other than the restrictions on attorney participation at the pretrial hearing, why so few landlords chose to participate in the Pilot Project. What is clear, however, is that landlords participating in the Pilot Project were represented by counsel in only 19.6% of the cases, whereas 72.5% of landlords proceeding in the unlawful detainer system were represented by counsel.[250] As for tenants, 1.2% were represented by counsel in the Pilot Project cases and 2.8% were represented in the unlawful detainer system.[251] The results of both systems are similar to Massachusetts results in terms of final judgments – 97.3% of judgments in the Pilot Project were entered in favor of the landlord, and 98.4% of judgments were entered in favor of the landlord in the unlawful detainer system.[252]

Zucker summed up his study by stating that a system such as the Pilot Project, which requires a pretrial hearing and/or a pretrial rent deposit as a condition to a tenant receiving a trial on the merits, caused significant procedural problems for the court and the parties involved – "[a]ll of the judicial energy and resources required to administrate this new system may not be worth the 14.8% reduction in time that appears to have been achieved."[253]

---

should not adversely effect the parties' abilities to present their cases (due process protections). *See id.* at 169, 175. The above objective regarding outstanding Code violations would not be feasible in Massachusetts, as the housing stock is much older than average housing and as a result many units are substandard and contain lead-based paint. *See* Consolidated Plan, *supra* note 215, at ii.

246. Zucker, *supra* note 230, at 188.

247. *See id.* at 188-89.

248. *See id.* at 189.

249. *See id.* at 175.

250. *See* Zucker, *supra* note 230, at 176-78.

251. *See id.*

252. *See id. See also* Survey, *supra* note 116.

253. Zucker, *supra* note 230, at 189.

IV. RECOMMENDATIONS

A. Impact of Proposed Bill

The Governor's bill, if passed, will have disastrous effects on the low-income rental population. In fact, it is the low-income population[254] who will be most affected by this law, as it is estimated that at least 90% of the cases brought in BHC on any given Thursday are nonpayment of rent cases.[255] Usually, the only tenants who raise defenses and/or counterclaims to the Summary Proceeding are those who are represented by counsel.[256] In fact, most tenants do not even answer the complaint they receive.[257]

Housing affordability in Massachusetts is a very serious problem.[258] In fact, nearly one-third of the entire population of Massachusetts is experiencing affordability problems.[259] "As a result [of this crisis], evictions for

---

254. Approximately 24% of residential tenancies in Massachusetts are subsidized by the government, either through project-based development and operating assistance, or through tenant-based rental assistance. *See* Consolidated Plan, *supra* note 215, at 1-8. This is a very significant portion of the rental population. It is also estimated that 41% of all rental households in Massachusetts have affordability problems. *See id.* at 1-9. Data from the 1990 census indicated that "there were 543,966 very low income households (less than 50% of median income), another 295,331 low income (51-80% of median) and another 185,692 moderate income households (81-95% of median)" in the state of Massachusetts. *Id.* at 1-8.

255. *See* Chopp Interview *supra* note 178.

256. *See id.*

257. *See id.* The law was actually amended in the early 1980s to address a serious problem: hundreds of tenants were being defaulted and denied their day in court for failing to serve an answer according to Rule 3 of the Massachusetts Uniform Summary Process Rules, which states that a defendant shall prepare a written answer which should deny every statement in the complaint which is in dispute, and shall put forth any defenses and/or counterclaims the defendant may have. *See* MASS. UNIF. SUMM. PROCESS R. 3. The answer is supposed to be filed with the court and served upon the plaintiff no later than the first Monday after the Monday entry day, which is also the Monday before the scheduled hearing. *See id.* Rule 10 of the Massachusetts Uniform Summary Process Rules was subsequently amended to read: "[i]f the defendant appears but has failed to file an answer, no default shall enter and the court shall postpone the trial date one week from the original trial date, unless the plaintiff consents in writing to an immediate trial." MASS. UNIF. SUMM. PROCESS R. 10.

258. *See* McCormack Institute, *Situation Critical: Meeting the Housing Needs of Lower-Income Massachusetts Residents, A Study by the Center for Social Policy*, University of Massachusetts Boston, at Introduction (2000) (hereinafter "McCormack Study").

259. *See id.* at 1. The McCormack Study also reveals that about 35,000 residents of the state are homeless and another 670,000 are shelter-poor – they have places to live but "face a squeeze between their housing costs and incomes that leaves them unable to meet their non-housing needs at even a low level of ade-

nonpayment of rent have risen across the state."[260] Part of the McCormack Study focuses on evictions and foreclosures.[261] The results of their study show that while not all summary process cases ultimately result in eviction, many tenants who are unable to pay their rent simply move out after receiving a notice to quit, even before any court action has been initiated.[262] This information, combined with statistics from the courts, implies that of the 950,000 renter households in Massachusetts, over 5% of all renters are evicted as a result of inability to pay the rent for their homes.[263] The repeal of rent control in 1994 has also resulted in a drastic increase in the number of evictions across the state.[264] The number of homeless people has also dramatically increased in the past decade.[265]

Mandatory rent escrow will ultimately result in more evictions. Tenants who do not have access to legal representation will find it virtually impossible to comply with the proposed amendments. The low-income population makes up an overwhelming majority of defendants in Housing Court every week;[266] denying these people the right to a defense will most certainly result in eviction. Small property owners are demanding that the legislature amend Chapter 239, 8A in order to protect a relatively small number of plaintiffs who bring summary process actions at the expense of a large number of defendants who cannot effectively defend themselves. If amended, the result would not be equitable – landlords are already in a superior economic bargaining position by the sheer fact that they own property. Further, many tenants, especially those receiving some kind of subsidy, are often at the mercy of their landlords due to their fear of losing

---

quacy. Most of the shelter-poor are either elderly, or working single-parent families, and/or people of color." *Id.*

260. *Id.*

261. *See id.* at 6.

262. *See id.*

263. *See id.*

264. *See Profile*, *supra* note 134, at 19. Evictions in Massachusetts, between 1993 and 1995, reportedly increased 53%. *See* McCormack Study, *supra* note 258, at 6.

265. *See* McCormack Study, *supra* note 258, at 8. In 1999, 10,500 Massachusetts families were homeless – an increase of 112% from 1990. *See id.*

266. *See* Sample, *supra* note 184. Examination of weekly trial calendars for the Boston Housing Court shows that the Boston Housing Authority alone averages 32 cases per week, and on average 67% of the cases in Boston Housing Court every week are brought by either the Boston Housing Authority or other big management companies. There is no way to effectively compile data on what percent of cases any given week are Section 8 Voucher recipients, but the sheer fact that only 7% of tenants statewide are represented by counsel shows that the majority of tenants do not have enough money to hire a private attorney. In Boston Housing Court, about 14% of tenants have legal representation. *See id.*

their rental subsidy.

B. Alternatives

Small property owners' main concern is money, especially owners who occupy units in their buildings and rely on the rent from the other units to finance the mortgage.[267] These property owners believe that rent escrow will solve this problem.[268] The most effective way to solve the problem, however, is not rent escrow, but the establishment of some type of emergency restitutionary fund for landlords who face foreclosure as a result of a tenant's failure to pay rent.[269] Zucker borrows the theory of the restitutionary fund from the criminal justice system.[270] Massachusetts could implement a program designed to help landlords in need by adding a $40 fee (still less than the filing fee for a new civil action in Massachusetts District Courts, which is $150) to the current sixty-dollar filing fee for a summary process action. Statewide, this program could raise almost $1,000,000.[271]

---

267. *See The Road Home*, *supra* note 132, at 19.

268. *See id.* at 27.

269. *See* Zucker, *supra* note 230, at 194.

270. *See id.* Zucker's other proposed remedies include: (1) alternative dispute resolution, (2) increased court sanctions, statutory damages, or criminal penalties for frivolously invoked defenses, (3) vexatious unlawful detainer litigant statute which would state that a tenant who enters an appearance in at least three unlawful detainer actions in a three-year period that result in a judgment of possession for the landlord is a "vexatious unlawful detainer litigant" and would therefore have to file a motion with the court in any subsequent action to prove that there is a reasonable probability that he will prevail in order to use the discovery process, issue subpoenas, or file motions, and (4) priority setting for landlord hardship – the "small landlord remedy." *See id.* at 190-96. Alternative dispute resolution, or mediation is not an appropriate proposal for BHC mainly because 95% of scheduled cases currently are sent straight to mediation. *See* Chopp Interview, *supra* note 178. The increased court sanctions and statutory damages proposal is also an inadequate remedy as it is unlikely to result in any financial gain to the landlord due to the fact that many tenants are judgment-proof. *See* Zucker, *supra* note 230, at 192. The vexatious unlawful detainer statute would also fail to aid the landlord, because as Zucker points out, very few tenants "truly employ delay tactics to abuse the system. Therefore, the use of [such a statute] should be the anomaly, rather than the rule." *Id.* at 194. Zucker's final proposal is complementary to the program I have suggested Massachusetts implement in that it seeks to serve small property owners, many of whom live in the buildings they own and particularly rely on the income from rent.

271. The Housing Court Department's General Statistics for the fiscal year 2000 reported that 23,484 new summary process cases were filed. The Statistics also show that statewide 8,544 criminal cases are filed – these cases are almost exclusively brought by the Inspectional Services Department, or other local Boards of Health against landlords who fail to repair Code violations. *Id.* If this

Landlords, upon written request, could either be awarded or loaned money from this fund depending on the circumstances.

The fund could even be accessible to landlords who are not at risk of foreclosure, but need the money to make repairs to their buildings, provided they sufficiently prove that the money will be used to make repairs, and that they cannot possibly make the repairs without financial help. In this case, the landlord would receive a loan from the fund and would repay the loan with interest, thereby creating another source of income for the fund. The state would have to administer the fund and supervise how the funds are being used in order to ensure that landlords who receive loans for property repairs are actually using the funds to make the repairs.

This extra forty-dollar increase would not create an undue burden on small property owners, considering the fact that in Boston Housing Court close to 70% of the scheduled trials are brought by big management companies.[272] Therefore, these management companies would be the main source of the funds. A restitutionary fund would not only serve landlords, but would help tenants by improving the existing housing stock. This sort of fund could also serve landlords by cutting down on Chapter 239, 8A defenses – if the landlord is able to maintain his property and repair conditions and Code violations, tenants will not be able to take advantage of Chapter 239, 8A to the same degree as they are currently able to do.

V. CONCLUSION

The evolution of the landlord-tenant relationship from caveat lessee to consumer protection fashioned a new angle to Contracts law.[273] This evolution was necessary for the protection of our cities and towns, not to mention our citizens.[274] Mandatory rent escrow is in effect retrograde-evolution, as it seeks to erode the social progress this Country has made over the last hundred years by creating a requirement which exists in no other realm of contracts law[275] – it de facto discriminates against the lower-income population by requiring them to pay in order to have their

---

number were included, an additional $42,720 could be raised. *Id.*

272. *See* Sample, *supra* note 184.

273. *See supra* note 17 and accompanying text.

274. *See supra* Section I, part E and accompanying text.

275. *The Road Home* claims that a similar mechanism exists in the form of a real estate lien. This is not, however, an appropriate analogy, as a lien acts merely as security and does not require an owner to liquidate any funds unless a judgment is rendered against him. Therefore, the party upon whom the lien is imposed is not required to actually pay anything before having his day in court. *See The Road Home*, *supra* note 132, at 27.

day in court.

Chapter 239, 8A currently allows Judges to make pre-trial orders.[276] The fact that judges do not often make them is indicative of their usefulness – or lack thereof. Bruce Zucker concluded his article about the California Pretrial Rent Deposit Program by stating that "it seems that the traditional system is perfectly capable of handling court evictions in California in a very timely fashion."[277] The program did not work in California and mandatory rent escrow will not work to serve small property owners in Massachusetts either.

Serena D. Madar

276. *See supra* notes 79, 128 and accompanying text.
277. Zucker, *supra* note 230, at 197.

Concurrent Protection of Products By Patent and Trade Dress: Use of the Functionality Doctrine in *Marketing Displays, Inc. v. TrafFix Devices, Inc.*

I. INTRODUCTION

Intellectual property is a significant aspect of most large and small corporations in today's high technology market.[1] It, in fact, seems to have a direct and positive impact on both personal economic growth and national economic competitiveness.[2] Indeed, it is often stated that one of the primary purposes of the patent laws is to encourage scientific advancement and innovation.[3] Additionally, patents allow the public to take advantage of significant technology by permitting copying of an inventor's work at the end of the twenty-year term of the patent[4] and competing with the said inventor in the market.

Certainly, however, the inventor possesses the right to protect her product after the expiration of her patent and to continue selling the product in a responding market without interference by an infringing user. Extending the monopoly created by a patent would greatly alter the original purpose of the patent laws.[5] However, in addition to patents, trademark interests in a patented product may have developed, and manufacturers are entitled to protect their trademark and trade dress rights so that consumers are not

---

1. *See* DONALD S. CHISUM ET AL., PRINCIPLES OF PATENT LAW 50-51 (1998).

2. *See id.*

3. *See, e.g.*, Andrea Falk, Comment, *Harmonization of the Patent Act and Federal Trade Dress Law: A Critique of* Vornado Air Circulation Systems v. Duracraft Corp., 21 IOWA J. CORP. L. 827, 830 (1996).

4. *See* 35 U.S.C. § 154 (1994) (stating length of patent protection is twenty years).

5. The U.S. Constitution provides that inventors "by securing for limited Times . . . [have] exclusive right to their . . . Discoveries." U.S. CONST. art. I, § 8, cl. 8. This has been interpreted as giving the inventor a monopoly over her works for a statutorily defined length of time. *But see* Thomas F. Cotter, *Is This Conflict Really Necessary?: Resolving an Ostensible Conflict Between Patent Law and Federal Trademark Law*, 3 MARQ. INTELL. PROP. L. REV. 25, 29 (1999) (arguing that it is erroneous to view a patent as a monopoly).

confused as to source.[6] What if the product configuration, protected as trade dress or trademark, is such that the public would be unduly confused by an infringing product that too closely resembles the original inventor's patented work? Is the public entitled to protection from such confusion? Or is it more important that the inventor be excluded from extending his exclusive rights? Would a trade dress, provided for an unencumbered number of years, extend a monopoly in that market? Is there *really* a conflict between these two ideas, arising under individual legal frameworks?[7]

On June 26, 2000, the United States Supreme Court granted certiorari to the Sixth Circuit Court of Appeals in order to resolve some of these pending questions.[8] The case was decided on March 20, 2001.[9] In particular, the Supreme Court granted certiorari to address whether federal trade dress protection extends to product configurations covered by expired utility patents.[10] The case, *Marketing Displays, Inc. v. TrafFix Devices, Inc.*,[11] originated in the Sixth Circuit and was one of many from various circuits that presented the issue of the extent of trade dress protection.[12] It was expected that the Supreme Court's decision would resolve the split that existed in the circuits on this issue.[13]

The importance of the issue as framed by the Court resides in the nature of patent protection, its virtual monopoly on a market for a term, and the effect that trade dress protection will essentially have on extending this monopoly indefinitely. It is the purpose of this Note to conduct an analysis of whether a true conflict exists between the patent laws and the Lanham Act, which is the source of trade dress protection.[14] In particular, the *Marketing Displays*[15] case will be the vehicle for the analysis. The briefs presented to the Supreme Court will be used to illustrate the inherent differences between trademark and patent protection that lead this author to con-

---

6. *See generally* 15 U.S.C. §§ 1051-1127 (1994).

7. The legal frameworks referred to are the Patent Act, 35 U.S.C. §§ 1-376, and the Lanham Act, 15 U.S.C. §§ 1051-1127.

8. *See* TrafFix Devices, Inc. v. Marketing Displays, Inc., 530 U.S. 1260 (2000).

9. *See* TrafFix Devices, Inc. v. Marketing Displays, Inc., 121 S. Ct. 1255 (2001).

10. *See id.*

11. 200 F.3d 929 (6th Cir. 1999) (hereinafter *Marketing Displays III*).

12. *See, e.g.*, Vornado Air Circulation Sys., Inc. v. Duracraft Corp., 58 F.3d 1498 (10th Cir. 1995); Sunbeam Prods., Inc. v. West Bend Co., 123 F.3d 246 (5th Cir. 1997); Thomas & Betts Corp. v. Panduit Corp., 138 F.3d 277 (7th Cir. 1998); Midwest Indus., Inc. v. Karavan Trailers, Inc., 175 F.3d 1356 (Fed. Cir. 1999).

13. *See infra* notes 89-134 and accompanying text.

14. *See* the Lanham Act, 15 U.S.C. §§ 1051-1127 (1994).

15. *See Marketing Displays III*, *supra* note 11, at 929.

clude that trade dress and patent protection are independently available, and both may be employed to protect any one invention. Further, a Supreme Court decision passed down prior to the *Marketing Displays* decision, *Wal-Mart Stores, Inc. v. Samara Bros., Inc.*,[16] will assist in illustrating the predictability of the direction in which the *Marketing Displays* case turned. It will be shown that the Supreme Court in fact decided the *Marketing Displays* case in accord with already existing legal analysis, despite the Court's apparent attempt to abbreviate the issue as originally presented.

This Note will be broken into the following sections: Part II will provide the background and applicable law by presenting the historical foundation of the *Marketing Displays* case, then a legal framework.[17] Part III will include a discussion of the circuit split via differing case law.[18] Part IV will introduce the arguments as presented to the Supreme Court by both Marketing Displays and TrafFix, the two parties to the Sixth Circuit case, and will include an analysis of the merits of these arguments.[19] Part V will be an analysis of the case in light of the current state of the law, including an explanation of the Supreme Court's decision.[20] And Part VI will be the author's conclusion, including a critique of the Supreme Court's decision with respect to the issues in the case.

II. BACKGROUND

A. *History of the* Marketing Displays *Case*

Marketing Displays, Incorporated (MDI) is the manufacturer of wind-resistant sign stands that are used as traffic signs. The signs are used especially in construction and sold under the trademark WINDMASTER, which was obtained in July 1977.[21] The inventive nature of the sign is that the use of spring mounts make the signs mobile enough to resist strong winds.[22] This particular design was patented in the mid-1960s, with the dual spring mounts claimed in the patent.[23] All applicable patents, however, expired on May 16, 1989.[24] MDI continued to use the product on the market, precisely as it had under the patents when they were enforceable.[25]

In 1994, TrafFix Devices, Incorporated (TrafFix) introduced a similar

16. 529 U.S. 205 (2000).
17. *See infra* notes 21-87 and accompanying text.
18. *See infra* notes 88-134 and accompanying text.
19. *See infra* notes 135-269 and accompanying text.
20. *See infra* notes 270-316 and accompanying text.
21. *See Marketing Displays III*, *supra* note 11, at 932.
22. *See id.*
23. *See id.*
24. *See id.*
25. *See id.*

product in the market and was issued a trademark for the name WINDBUSTER in the United States in 1996.[26] As with MDI's product, TrafFix sold its wind-resistant signs in the highway construction industry.[27] In fact, the TrafFix product was a reverse-engineered[28] version of the MDI product so that it could be sold in competition with the MDI sign stands at the end of the term of patent protection.[29] TrafFix proceeded to register its mark before using the WINDBUSTER name, and the United States Patent and Trademark Office (PTO), having found no similar registered mark, published the new mark in its Official Gazette on March 14, 1995.[30] No one opposed the mark during the prescribed time period, and the mark was allowed by the PTO and issued on January 9, 1996.[31]

It was after the prescribed period, in July 1995, that MDI reacted to the potentially infringing mark and filed its initial complaint.[32] The complaint alleged Lanham Act trademark infringement, violations of federal trade dress law, and unfair competition under Michigan common law.[33] This complaint was later amended to include a claim under Lanham Act unfair competition.[34] TrafFix counterclaimed, alleging Lanham Act unfair competition and federal antitrust violations.[35] In January 1997, the district court granted summary judgment to MDI on its trademark claims, granting a permanent injunction against TrafFix's use of the infringing WINDBUSTER mark, and TrafFix's antitrust counterclaim was dismissed.[36] Then in June 1997, the district court separately addressed the

---

26. *See* Marketing Displays, Inc. v. TrafFix Devices, Inc., 967 F. Supp. 953, 956 (E.D. Mich. 1997) (hereinafter *Marketing Displays I*).

27. *See Marketing Displays III*, *supra* note 11, at 932.

28. *See id.* Reverse-engineering is essentially re-engineering the same product with the same configuration, and, in this case, TrafFix actually used the MDI traffic stand to develop an improved product during the re-engineering process. *See* Petitioner's Brief at 6, TrafFix Devices, Inc. v. Marketing Displays, Inc., 530 U.S. 1260 (June 26, 2000) (No. 99-1571).

29. *See Marketing Displays I*, *supra* note 26, at 956.

30. *See id.* The procedure of publishing the mark in the Official Gazette is called an opposition, and it requires the publication of the mark by the Patent and Trademark Office before issuance of the mark so that persons may oppose applications for marks that they determine may be harmful to them. *See* 15 U.S.C. §§ 1062(a), 1063 (1994).

31. *See Marketing Displays I*, *supra* note 26, at 956.

32. *See id.*

33. *See id.* The court immediately dismissed MDI's claim of state unfair competition. *See id.*

34. *See id.*

35. *See id.*

36. *See Marketing Displays I*, *supra* note 26, at 965.

trade dress infringement claim brought by MDI.[37] Finding that MDI had failed to develop secondary meaning in the sign stand design and, determining that even if MDI had proved secondary meaning that the dual spring design was functional, the court granted summary judgment to TrafFix on the trade dress claim.[38]

The Sixth Circuit Court of Appeals decided the issues on appeal in December of 1999.[39] It upheld the district court's permanent injunction, but reversed as to MDI's trade dress and unfair competition claims, stating that MDI had presented genuine issues of material fact.[40] It was from this decision that TrafFix appealed to the Supreme Court of the United States, and the Court granted certiorari, wishing to address the issue of whether by attempting to protect its trade dress of the WINDMASTER mark, MDI was also lengthening the patent protection it once possessed before the expiration of its patents in 1989.[41]

Prior to the Supreme Court decision, the circuit courts were split in the determination of this issue.[42] The circuits that have opposed trade dress protection after a utility patent disclosure have done so primarily because of the public policy purposes behind the patent laws.[43] Contrarily, circuit courts that have supported the availability of trade dress protection reason that trade dress and patent protection are two separate intellectual property rights and can be had exclusive of each other.[44] These competing policies and the cases in which they developed cannot be fully comprehended without an analysis of the purposes of patent protection as opposed to trademark protection.

B. *Patent Law*

The ability of an inventor to obtain exclusive patent rights is rooted in the U.S. Constitution.[45] In light of the constitutional allowance, the Patent

---

37. *See* Marketing Displays, Inc. v. TrafFix Devices, Inc., 971 F. Supp. 262, 264 (E.D. Mich. 1997) (hereinafter *Marketing Displays II*).

38. *See id.* at 276.

39. *See Marketing Displays III*, *supra* note 11, at 929.

40. *See id.* at 942.

41. *See* TrafFix Devices, Inc. v. Marketing Displays, Inc., 530 U.S. 1260 (2000).

42. *See infra* notes 89-134 and accompanying text.

43. *See, e.g.*, Vornado Air Circulation Sys., Inc. v. Duracraft Corp., 58 F.3d 1498 (10th Cir. 1995); Zip Dee, Inc. v. Dometic Corp., 931 F. Supp. 602 (N.D. Ill. 1996).

44. *See, e.g.*, Sunbeam Prods., Inc. v. West Bend Co., 123 F.3d 246 (5th Cir. 1997); Thomas & Betts Corp. v. Panduit Corp., 138 F.3d 277 (7th Cir. 1998); Midwest Indus., Inc. v. Karavan Trailers, Inc., 175 F.3d 1356 (Fed. Cir. 1999).

45. The Intellectual Property Clause, U.S. Const. art. 1, § 8, cl. 8.

Act was passed by Congress in 1952.[46] Two types of patents are available: utility patents and design patents.[47] Design patents have a duration of fourteen years and are obtained to protect nonfunctional design features.[48] Utility patents, also known as functional patents, have a duration of protection of twenty years from the date of filing.[49] The two types of patents differ in that a design patent is not available where any aspect of the patented design is for a useful purpose. Both types of patents will not issue unless it is shown that the invention is novel, useful and non-obvious.[50]

It is the public policy purposes for issuing patents that are of particular interest in distinguishing patent rights from trademark rights. The primary purposes for issuing patents are to promote invention, reward invention and to permit ideas to become a part of the public domain for use after an inventor has secured an exclusive right for the life of the patent.[51] The current term for the life of a patent is twenty years.[52] Hence, an inventor enjoys an exclusive right to promote and profit from her patent for a fixed

---

46. *See* 35 U.S.C. §§ 1-376 (1994).

47. "Patents for processes, machines, manufactures and compositions of matter are generally called 'utility' patents." 1 DONALD S. CHISUM, CHISUM ON PATENTS § 1.01, at 1-9 (2001). This is to be distinguished from design patents, which protect "new, original and ornamental design for [articles] of manufacture." *Id.* Design patents are issued for 14-year terms, while utility patents are issued for 20-year terms. *See id.* There is an obvious overlap of patent law and trademark law with respect to design patents, since design patents are concerned with the way an invention looks. *See id.* § 1.04, at 1-297. Both cover the "nonfunctional features of useful objects." *Id.* This article does not address this particular issue, but focuses on the potential overlap of utility patent protection and trademark protection in the form of trade dress. Thus, reference made hereafter to a patent is reference to a utility patent, as defined in this footnote.

48. *See* 35 U.S.C. § 173 (1988).

49. *See* 35 U.S.C. § 154 (1994).

50. *See* 35 U.S.C. §§ 101-03 (1994). The statute reads, in part:

> Novelty requires the invention to be new. An invention is obvious if previously granted patents, published writings, or a combination of patents or publications would have shown a person, having ordinary skill in the field of the invention, how to make the invention, even if the prior patents or publications did not specifically cover the exact invention.

Gwendolyn Gill, Comment, *Through the Back Door: Attempts to Use Trade Dress to Protect Expired Patents*, 67 U. CIN. L. REV. 1269, 1273 (1999). The usefulness requirement will be discussed in more detail. *See infra* notes 295-98 and accompanying text.

51. *See* Theodore H. Davis, Jr., *Copying in the Shadow of the Constitution: The Rational Limits of Trade Dress Protection*, 80 MINN. L. REV. 595, 616 (1996) (quoting Quick Point Pencil Co. v. Aronson, 567 F.2d 757, 762 (8th Cir. 1977)).

52. *See* 35 U.S.C. § 154 (1994).

term of years, then the invention becomes available to the public for use in a freely competitive market.[53]

C. *Trademark Law*

The criteria for protection of a trademark are that it be used to distinguish goods or services such that a single source is identifiable.[54] A relatively significant difference between patent protection and trademark protection is that the latter is not for a fixed term, but is available on a renewable basis so long as the mark is being used in a regular manner.[55] An additional difference lies in the purpose of the relative rights: "[u]nlike patent law, the purpose of trademark and trade dress law is to prevent customer confusion and protect the value of identifying symbols, not to encourage invention by providing a period of exclusive rights."[56] Due to the purposes of trademark protection, it is unnecessary to limit the term for which it is protected.[57]

Trade dress protection differs from trademark protection, although there seems to be a significantly disappearing distinction between trade dress and trademarks.[58] Trade dress used to be considered unfair competition, but since trade dress requires the same factual conclusions as trademark protection, the two have been combined, so that protectable trade dress can

---

53. *See* Gill, *supra* note 50, at 1275-76.

54. *See* J. THOMAS MCCARTHY, 1 MCCARTHY ON TRADEMARKS AND UNFAIR COMPETITION § 3:1, at 3-2 (4th ed. 1996 & Supp. June 2001).

55. *See id.* § 6:8, at 6-11.

56. *Id.* § 6:3, at 6-6. *See also* Davis, *supra* note 51, at 603 ("the purpose of trademark law was *not* to reward creativity and ingenuity").

57. *See* Kevin E. Mohr, *At the Interface of Patent and Trademark Law: Should a Product Configuration Disclosed in a Utility Patent Ever Qualify for Trade Dress Protection?*, 19 HASTINGS COMM. & ENT. L.J. 339, 344-45 n.16 (1997):

> Since [trademarks] are intended primarily to advance consumer awareness rather than any body of knowledge, it would make little sense to limit the term of a trademark. The longer a mark is used in connection with a product, the stronger the consumer's association of that product with a particular producer, source or expected quality becomes. Although in certain instances a competitor will need a mark to compete effectively in the market in which the mark is used, there is usually an abundance of words or symbols to go around for everyone who wants to compete.

Id.

58. *See* MCCARTHY, *supra* note 54 § 8:1, at 8-4 (commenting that the difference between trade dress and trademarks is historical). "Today, many types of designations protectable as 'trade dress' are also registrable as 'trademarks.'" *Id.*

be a registrable trademark.[59]

Federal trade dress protection is asserted under the Lanham Act.[60] "[T]he protection of trademarks and trade dress under § 43(a) serves the same statutory purpose of preventing deception and unfair competition. There is no persuasive reason to apply different analysis to the two."[61] In 1999, this section was amended essentially to authorize its use to claim infringement of unregistered trade dress.[62] Trade dress has evolved over time to include three types: (1) "overall appearance of labels, wrappers, and containers used in packaging a product;"[63] (2) "totality of any elements in which a product or service is . . . presented;"[64] and, most recently, (3) the product's configuration, or its shape and design.[65] In each of these types, "the elements making up the alleged trade dress must have been *used* in such a manner as to denote product source."[66] There is no trade dress protection for a mere aesthetic feature.[67]

A determination that a trade dress has been infringed depends upon whether the alleged infringing trade dress is likely to cause confusion to a potential buyer, and therefore, that a finding of unfair competition is warranted.[68] The totality of the product is relevant, not simply the individual details in the product design.

Trade dress is protectable if three requirements are met:

> [1] it is either inherently distinctive or has acquired a secondary meaning, [2] it is nonfunctional, and [3] a likelihood of confusion can be established between the trade dress of the plaintiff as applied to his goods and/or services, when compared to the trade dress of the defendant, as applied to his goods and/or services.[69]

---

59. *See* Two Pesos, Inc. v. Taco Cabana, Inc., 505 U.S. 763, 775 (1992).

60. *See* 15 U.S.C. § 1125(a) (Supp. V 1994).

61. *Two Pesos, Inc.*, 505 U.S. at 773.

62. *See* Lanham Act § 43(a), 15 U.S.C. § 1125(a) (Supp. V 1994). The 1999 Amendment to the Lanham Act also requires that the party asserting trade dress rights for an unregistered trade dress has the burden of proving that the trade dress is not functional. *See* Martin P. Hoffman, *Trademarks, Copyrights, and Unfair Competition for the General Practitioner and the Corporate Counsel*, SE44 A.L.I.-A.B.A. 161, 174 (2000).

63. MCCARTHY, *supra* note 54, §8:1. Product packaging is considered the oldest form of recognized trade dress. *See id.*

64. *Id.*

65. *See id.* § 8:5.

66. *See id.* § 8:1 (emphasis in original).

67. *See id.* § 8:1. Aesthetic features are not protected because they do not cause a consumer to identify the product with the source of the product. *See id.*

68. *See* MCCARTHY, *supra* note 54, at § 8:1.

69. *See* Hoffman, *supra* note 62, at 178. *But see* Wal-Mart Stores, Inc. v.

Inherent distinctiveness is dependent upon the trade dress being arbitrary, fanciful or suggestive.[70] Even if the trade dress is inherently distinctive, secondary meaning is necessary such that consumers recognize the dress and affiliate it with a particular product and a particular manufacturer.[71]

Each of the above descriptions is relevant in an analysis of trade dress protection, but in comparing trade dress with patents, the requirement of functionality is particularly important. There is an assumption that a trade dress is functional if it embodies "a design feature which is superior to other available designs and thus provides a competitive advantage to the user."[72] Thus, for one to successfully assert trade dress rights, it must be shown that the trade dress or trademark features are nonfunctional.[73] Generally,

> a product feature is functional and cannot serve as a trademark, if it is essential to the use or purpose of the article or if it affects the cost or quality of the article, that is, if exclusive use of the feature would put competitors at a significant non-reputation-related disadvantage.[74]

A functional feature cannot be protected, even in light of the existence of consumer confusion, or secondary meaning, and others are free to copy the alleged trade dress.[75] As one scholar notes:

> [The] [r]ule against protection of functional features is intended to avoid creation of perpetual monopoly in trademark, or trade dress, that would circumvent and/or conflict with patent protection, which is of limited duration.

Samara Bros., Inc., 529 U.S. 205 (2000) (holding that trade dress protection now always requires proof of secondary meaning, regardless of its status as either inherently distinctive or descriptive).

70. *See* Hoffman, *supra* note 62, at 179. *See also* Abercrombie & Fitch Co. v. Hunting World, Inc., 537 F.2d 4, 9 (2d Cir. 1976) (stating that the taxonomy for determining the strength of a trademark, from weakest to strongest, include generic, descriptive, suggestive, arbitrary, or fanciful; and also that generic marks are not protectable).

71. *See* Hoffman, *supra* note 62, at 179. Secondary meaning can be shown by a number of factors, including consumer recognition, the ability of the consumer to identify the source and the advertising or sales techniques used by the trade dress owner that focus on the trade dress. *See id.*

72. Trademark Manual of Examining Procedure (TMEP) § 1202.03.

73. *See* MCCARTHY, *supra* note 54, § 7:63.

74. Qualitex Co. v. Jacobson Prods. Co., 514 U.S. 159, 165 (1995) (holding that the green/gold color of a dry-cleaning press pad was protectable as trade dress since the color was not functional, but served as a source-identifying feature).

75. *See* Hoffman, *supra* note 62, at 189.

> Also, patent protection requires [a] level of innovation that is new, useful, and unobvious. Similar standards do not apply to trade dress or trademark.[76]

Trade dress protection and trademark registration are restricted so that their extent does not reach that of a patent, where functional features are the purpose of the limited protection.[77]

This discussion illustrates how two kinds of intellectual property developed for differing public policy purposes. As mentioned, and importantly, while patents enjoy a limited term of protection, trademarks and trade dress have potentially unlimited durations of protection.[78] Because of this, issues of monopolization and unfair competition are clearly implied, and the circuit courts were left to decipher the correct application of trade dress protection.

D. *Brief Discussion of Supreme Court Cases that Precede the Circuit Split Cases*

It is of relevance that the Supreme Court has addressed the preliminary question of whether trademark protection can be extended to unpatented product designs with respect to state unfair competition laws.[79] The Sears-Compco doctrine arose out of two cases that held that federal patent laws preempt state unfair competition laws.[80] In *Sears*, the plaintiff had patents on a lamp, and, after achieving commercial success, Sears began to sell identical lamps to compete with the plaintiff.[81] Reversing the lower court's decision, the Supreme Court determined that state unfair competition laws cannot prevent copying because under federal patent law, an unpatentable article is in the public domain.[82] Likewise, the *Compco* case led to a similar conclusion that a lighting fixture could be copied since it was not enti-

---

76. *See id.*

77. *See* MCCARTHY, *supra* note 54, at §§ 7:63, 7:64.

78. *See supra* notes 53-5 and accompanying text.

79. *See* David W. Opderbeck, *An Economic Perspective on Product Configuration Trade Dress*, 24 SETON HALL LEGIS. J. 327, 336 (2000).

80. *See* Sears Roebuck & Co. v. Stiffel Co., 376 U.S. 225 (1964) (stating that federal patent laws prevail to the extent that state laws allow protection to articles that cannot be patented under federal law); Compco Corp. v. Day-Brite Lighting, Inc., 376 U.S. 234, 238 (1964) (holding "if the design is not entitled to a design patent or other federal statutory protection, then it can be copied at will").

81. *See* Sears Roebuck & Co. v. Stiffel Co., 376 U.S. 225 (1964). *See also* Opderbeck, *supra* note 79, at 336.

82. *See* Opderbeck, *supra* note 79, at 337. *See also Sears Roebuck & Co.*, 376 U.S. at 231 (stating "[a]n unpatentable article, like an article on which the patent has expired, is in the public domain and may be made and sold by whoever chooses to do so").

tled to any federal statutory protection.[83]

This seemingly restrictive case law was later modified and clarified to say that the *Sears-Compco* doctrine was not intended to preempt all trade dress protection via state law for all unpatented articles.[84] The *Bonito Boats* decision reinforced the *Sears-Compco* doctrine in that federal patent protection preempts state unfair competition laws, but allows state protection of unpatented designs in situations where "consumer confusion is likely to result; the design 'idea' itself may be freely exploited in all other contexts."[85] The Court pointed to the protections of the Lanham Act, § 43(a), as a reminder that a balance between patent law and unfair competition is codified.[86] The Court later held that trade dress protection is not inherently anti-competitive.[87] Thus, the Court firmly resolved against upholding state unfair competition laws if they conflicted with the federal patent laws. With this backdrop, the circuit courts were left to decipher exactly where the clash between patent law and *federal* trade dress protection resolved itself.

---

83. *See Compco Corp.*, 376 U.S. at 238. Interestingly, the *Sears* case and the *Compco* case were decided by the Supreme Court on the same day, but on different grounds. The *Sears* decision was grounded in the Supremacy Clause, U.S. Const. art. VI, cl. 2, while the *Compco* case found support in the Intellectual Property Clause, U.S. Const. art. I, § 8, cl. 8. Both cases decided that the state laws at issue were unconstitutional, but the *Compco* case clearly gives the Intellectual Property Clause substantially more weight, which would arguably make it a violation of the Constitution to copy functional features. *See* Davis, *supra* note 51, at 609-13.

84. *See* Bonito Boats, Inc. v. Thunder Craft Boats, Inc., 489 U.S. 141, 146 (1989) (holding that a Florida statute prohibiting the copying of unpatented boat hulls was invalid because the statute gave the original boat hull designer rights against all the world, similar to what would be granted in a federal patent, but not fulfilling the necessary requirements of a patent).

85. *Bonito Boats, Inc.*, 489 U.S. at 158. *See also* Davis, *supra* note 51, at 618 (stating that *Bonito Boats* really only stands for the proposition that the Intellectual Property Clause, Art. I, § 8, cl. 8 of the Constitution, does not "affirmatively restrict . . . state action" in protecting intellectual property rights). It was after the *Sears-Compco* and *Bonito Boats* decisions that courts in trade dress cases have generally required nonfunctionality. *See* Davis, *supra* note 51, at 619. This may be the most significant aspect of the aforementioned line of cases. *See, e.g.*, Qualitex Co. v. Jacobson Prods. Co., 514 U.S. 159 (1995) (commenting on the functionality doctrine's ability to prevent the inhibition of competition since producers will be unable to control a useful feature); Two Pesos v. Taco Cabana, Inc., 505 U.S. 763 (1992) (stating that Lanham Act protection depends upon functionality).

86. *See Bonito Boats, Inc.*, 489 U.S. at 146. *See also* Opderbeck, *supra* note 79, at 341.

87. *See Two Pesos, Inc.*, 505 U.S. at 763.

III. THE CIRCUIT SPLIT

With the aforementioned developments in the law, it is no wonder that the circuit courts have disagreed regarding the ability to assert trade dress rights despite the already existing utility patent protection. The Fifth, Second, Seventh, Tenth and Federal Circuit Courts of Appeals have contributed to the discussion involving the extent of trade dress protection and its probable ramifications.[88] Each of the contributing decisions will now be discussed in turn.

A. *Opposing Trade Dress Protection in the Wake of Expired Patents*

The Tenth Circuit's opinion was expressed in a 1995 decision involving the spiral design of a grill on a fan.[89] The court addressed whether product configuration could be protected as trade dress.[90] Vornado had obtained a utility patent on the fan, and the spiral design of the grill was claimed in the patent application.[91] The lower court determined that the fan grill was nonfunctional, thus leading to the conclusion that Vornado's grill design was protectable as trade dress.[92] The district court reasoned that Vornado's competitors did not need to use the disclosed design, and that consumers were likely to be confused by the use of a similar design by Duracraft.[93] The court enjoined Duracraft from further use of its fan.[94]

But the Appellate Court reversed, stating, ". . . this case clearly shows that trade dress protection can directly interfere with the public's ability to practice patented inventions after the patents have expired, and that it undermines the principle that ideas in the public domain should stay there."[95] The court felt that the purposes of the Lanham Act, section 43(a), did not include protecting against consumer confusion that arises from the copying of configurations that are "significant parts of patented inventions."[96] The court stated:

> [W]here a disputed product configuration is part of a claim in a utility patent,

---

88. *See infra* notes 89-134 and accompanying text.

89. *See* Vornado Air Sys., Inc. v. Duracraft Corp., 58 F.3d 1498 (10th Cir. 1995).

90. *See id.*

91. *See id.* at 1500.

92. *See id.* at 1501.

93. *See id.* at 1502. *See also* Mohr, *supra* note 57, at 394-95.

94. *See Vornado Air Sys., Inc.*, 58 F.3d at 1502.

95. *Id.* at 1508.

96. *Id.* at 1509.

> and the configuration is a described, significant inventive aspect of the invention, see 35 U.S.C. § 112, so that without it the invention could not fairly be said to be the same invention, patent law prevents its protection as trade dress, even if the configuration is nonfunctional.[97]

Thus, Vornado, the court concluded, could not receive trade dress protection for its fan grill because it was a significant inventive element of the patented product.[98] It is of particular relevance that the configuration at issue was claimed in the patent obtained.[99]

The court's reasoning rested on the Supreme Court's interpretation of the meaning of functionality.[100] Functionality depends on the needs of the competitor to make an equally competitive product.[101] If a trade dress is functional, it does not merit protection, even if it is inherently distinctive or has acquired secondary meaning.[102] Thus, a viable argument was created in the Tenth Circuit to illustrate that patent protection should not be unduly extended under trade dress law where the product configuration is a sig-

---

97. *Id.* at 1510. *But see Intellectual Property- Patent and Trade Dress Law- Tenth Circuit Applies "Significant Inventive Aspect" Test to Determine Whether Utility Patent Precludes Trade Dress Protection,* 109 HARV. L. REV. 1457, 1458 (1996) (commenting on the general test devised by the Tenth Circuit that allowed them to conclude that "patent concerns precluded trade dress protection for the spiral grill" and determining that the court erred by failing to define the test more specifically); Michael S. Perez, *Reconciling the Patent Act and the Lanham Act: Should Product Configurations Be Entitled to Trade Dress Protection After the Expiration of a Utility or Design Patent?,* 4 TEX. INTELL. PROP. L.J. 383, 411-12 (arguing that the test set forth by the Tenth Circuit is flawed because "it would preclude trade dress protection for significant patented features in situations where the public would not be denied the benefit of the inventions," and it is "too narrow because it would permit trade dress protection for design features that are inextricably tied to patented processes where the designs do not contribute to the overall 'inventiveness' of the process").

98. *See Vornado Air Sys., Inc.,* 58 F.3d at 1510.

99. This is particularly relevant because it allowed the Tenth Circuit to easily conclude that the spiral fan design was a significant inventive component of the patented invention without conducting a functionality test. *See Vornado Air Sys., Inc.,* 58 F.3d at 1509-10. *But see* Respondent's Brief at 31, TrafFix Devices, Inc. v. Marketing Displays, Inc., 530 U.S. 1260 (2000) (No. 99-1571) ("[I]t is submitted that the Vornado Court did not recognize that distinctive, nonfunctional product features can be separated from the utilitarian features which are the proper subject of a patent. Despite noting that 'availability of equally satisfactory alternatives for a particular feature' is 'often the fulcrum on which Lanham Act functionality analysis turns', the Vornado Court ignored any functionality analysis in reaching its result." *Id.*)

100. *See Vornado Air Sys., Inc.,* 58 F.3d at 1507.

101. *See* Qualitex Co. v. Jacobson Prod. Co., 514 U.S. 159, 164 (1995).

102. *See* Two Pesos, Inc. v. Taco Cabana, Inc., 505 U.S. 763, 775 (1992).

nificant component in the formerly patented article.[103]

B. *Allowing Trade Dress Protection in the Wake of Expired Patents*

The Fifth Circuit had a contrary interpretation of the conflict between patent and trade dress protection in *Sunbeam Products, Inc. v. West Bend Co.*[104] *Sunbeam* involved an alleged copy of the configuration of a stand mixer that incorporated certain functional features that were previously named in utility patents.[105] The court focused on the overall shape and configuration of the mixer, concluding that it was nonfunctional, and found without merit the argument by West Bend that certain features were functional, therefore making them not entitled to trade dress protection.[106] The test employed by the court, dubbed the "utilitarian standard of functionality,"[107] focused on whether the features of the configuration would impede the competing market.[108] Of greatest importance to the court was that there were other manufacturers who competed with the Sunbeam product.[109] The court's rebuttal to West Bend's argument that functional features cannot receive trade dress protection substantially said:

> [T]he trade dress claim in the instant case entails a combination of several features, with individual utility patents granted to certain features. In this regard, the principle that trade dress is the "total image" of a product, rather than its isolated components, is dispositive . . . . [G]ranting trade dress protection to the arbitrary product configuration of the American Classic Mixmaster® would not restrict any competitors from copying the individual functional features after the individual utility patents have expired, but would simply prohibit competitors from copying the arbitrary combination of features embraced in [the mixer]. Hence, there is no conflict between patent law and trademark law in the instant case.[110]

The court concluded that protecting the trade dress would not frustrate competition.[111]

The Fifth Circuit did not particularly take into consideration the pur-

---

103. *See supra* notes 89-102 and accompanying text.

104. Sunbeam Prods., Inc. v. West Bend Co., 123 F.3d 246 (5th Cir. 1997).

105. *See id.* at 256.

106. *See id.* at 256-57.

107. *Id.* at 255.

108. *See id.*

109. *See id.* at 257.

110. Sunbeam Prods., Inc. v. West Bend Co., 123 F.3d 246, 256 n.19 (5th Cir. 1997).

111. *See id.* at 257.

poses of the patent laws, as Sunbeam's product was nonfunctional in nature.[112] Consequently, the Fifth Circuit's interpretation that trade dress must be considered in totality allowed it to distinguish the utility patents from the trade dress, unlike the Tenth Circuit's decision that the utility patent precluded the ability of the manufacturer to obtain trade dress protection at all.

In 1998, the Seventh Circuit directly addressed the apparent conflict between trademark protection of product configurations and patent law, and, agreeing with the Fifth Circuit, found that trade dress was available. The conflict arose between manufacturers of cable ties – "small nylon belts used to bundle wires" comprised of a head on one end that has a slit and locking mechanism.[113] A patent was obtained on the cable tie in 1965, expiring in 1982.[114] The utility patent claimed the head of the tie as one of its elements, but it did not claim the shape of the head; that was merely disclosed in the patent specifications.[115] This fact allowed the court to distinguish *Vornado*, since in that case the fan grill was not part of the claims in the patent.[116] This distinction seems to be the sole purpose for the Seventh Circuit's lack of uniformity with the *Vornado* decision.

Based on the lack of claims as to the tie head, and in light of Supreme Court decisions that were cited as supporting the distinctions between patent law and trademarks,[117] the court overturned the district court's ruling,

---

112. *See id.* at 256.

113. Thomas & Betts Corp. v. Panduit Corp., 138 F.3d 277, 282 (7th Cir. 1998), *cert. denied*, 119 S. Ct. 336 (1998). Note that the parties to this Seventh Circuit case wrote amici briefs in the *TrafFix* case for their respective arguments. *See infra* note 154 and accompanying text.

114. The life of a patent used to be seventeen years, but was increased to twenty years from the time of filing beginning in 1995. *See* 35 U.S.C. § 154(a)(2) (1994). The term of a design patent remains fourteen years from the date of issue. *See* 35 U.S.C. § 173 (1994).

115. *See Thomas & Betts*, 138 F.3d at 282.

116. *See id. See also* Mohr, *supra* note 57, at 398. A patent application includes a claims section as part of the specification that "is particularly written to describe in exact detail the actual invention for which the inventor seeks patent protection. A patent is granted solely for the invention described in the claims, not for what may be described elsewhere in the specification." Gill, *supra* note 50, at 1274-75.

117. Each of the Supreme Court cases is distinguished as to the case at bar. *See, e.g.*, Bonito Boats, Inc. v. Thunder Craft Boats, Inc., 489 U.S. 141 (1989) (no trade dress protection granted under a *state* unfair competition law, since the federal patent laws would preempt state law); Sears, Roebuck & Co. v. Stiffel Co., 376 U.S. 225 (1964) (monopoly ends when patent protection ends, so unfair competition laws of the *state* cannot attempt to extend the rights); Singer Mfg. Co. v. June Mfg. Co., 163 U.S. 169, 179 (1896) (the machine's shape was dictated in the patent and covered all components, so trade dress was not an option); Scott Paper

finding that there "is no *per se* prohibition against features disclosed in a patent receiving trademark protection after the patent has expired."[118] The Court went further in stating "a configuration of an article cannot receive trademark registration if its purpose is to contribute functional advantages to the article or if the configuration results from functional considerations."[119] However, in light of this conclusive statement of law, and after distinguishing this case from the Tenth Circuit's *Vornado* decision,[120] the court found that there was no conflict with patent laws with respect to the Thomas & Betts trade dress because the shape of the invention was not claimed in the prior patent.[121] It seems that the Seventh Circuit may have agreed with the Tenth Circuit, if not for the fact that the trade dress in question (the shape of the head of a cable tie) was never claimed in the previous patent, and, as a result, even if nonfunctional, there existed no conflict between the trademark protection and the patent protection, even if the trade dress was nonfunctional.

The question of conflict presented itself yet again, this time in the Court of Appeals for the Federal Circuit.[122] The invention at issue was a curved winch post mounted to the front of trailers, necessary to draw watercraft onto the trailers for tow.[123] The district court granted an injunction to Karavan, the defendant, based upon the Tenth Circuit *Vornado* case.[124] The district court concluded that because the winch was discussed in a prior patent and claimed within a corresponding claim of that patent, the winch was a "significant inventive aspect"[125] of the patent, thereby rendering it unprotectable trade dress.[126] The appeals court found this conclusion to be erroneous because there was no determination as to whether the winch in

---

Co. v. Marcalus Mfg. Co., 326 U.S. 249 (1945) (dicta stated that patent protection ends a monopoly in favor of public use, but the case was decided on assignor estoppel, not trademark law, so was not dispositive). The *Thomas & Betts* court discussed each of the foregoing to show that this case could allow trade dress protection as exclusive from the former utility patents. *See Thomas & Betts*, 138 F.3d at 285-87.

118. *Thomas & Betts*, 138 F.3d at 288.

119. *Id.*

120. *See supra* notes 89-103 and accompanying text for a discussion of *Vornado*.

121. *See Thomas & Betts*, 138 F.3d at 289.

122. *See* Midwest Indus., Inc. v. Karavan Trailers, Inc., 175 F.3d 1356 (Fed. Cir. 1999).

123. *See id.* at 1357.

124. *See id.* at 1358.

125. Vornado Air Circulation Systems, Inc. v. Duracraft Corp., 58 F.3d 1498, 1510 (10th Cir. 1995) (court held that Lanham Act protection was not available when the component was significant in the utility patent).

126. *See Midwest Indus., Inc.*, 175 F.3d at 1358.

question was functional, which was necessary to decipher whether trademark protection was at all available.[127]

The court's test included a careful look into the functionality of the winch post, which led to a determination that although the winch post was a significant inventive aspect in the prior patent claims, that was *not* the same as determining it to be functional.[128] The functionality test employed by the Federal Circuit is that stated by the Supreme Court in *Qualitex*: "that Midwest's right to exclude others from using that feature 'would put competitors at a significant non-reputation-related disadvantage.'"[129] Functionality is stated as the sole test in an inquiry into the relationship between the Lanham Act and the patent laws.[130] Its purpose is summarily presented in the *Qualitex* case:

> The functionality doctrine prevents trademark law, which seeks to promote competition by protecting a firm's reputation, from instead inhibiting legitimate competition by allowing a producer to control a useful product feature . . . . If a product's functional features could be used as trademarks . . . a monopoly over such features could be obtained without regard to whether they qualify as patents and could be extended forever (because trademarks may be renewed in perpetuity).[131]

The Federal Circuit conscientiously focused not on whether a utility patent has been granted in the past, but rather on the functionality of the features questioned.[132] The significance of the prior patent is merely evidence of whether the asserted trade dress is functional.[133] The Federal Circuit used Supreme Court precedent to establish that there is not a conflict between the patent laws and the Lanham Act, although none of those cases mentioned address the issue on point.[134]

The *Marketing Displays* case arose in the midst of the circuit split and raised a similar question. The arguments presented to the Supreme Court by each side passionately argue, and add to, the logic that emerges from the Tenth Circuit *Vornado* decision and those decisions opposing it. It already seems clear, however, that the Supreme Court was in a position to either maintain the functionality doctrine or do away with it, as it is a

---

127. *See id.* at 1361, 1364.

128. *See id.* at 1364.

129. *See id. See also* Qualitex Co. v. Jacobson Prod. Co., 514 U.S. 159, 165 (1995) (setting forth the test for functionality).

130. *See Midwest Industries, Inc.*, 175 F.3d at 1364.

131. *See id.* at 1362, *quoting Qualitex*, 514 U.S. at 164-65.

132. *See id.*

133. *See id.*

134. *See id.* at 1363.

source of confusion for all of the circuit courts.

IV. ARGUMENTS BEFORE THE SUPREME COURT

A. *Petitioner TrafFix's Position*

The foregoing discussion of the circuit split[135] laid the foundation on which the parties of the Sixth Circuit case[136] built their arguments for the Supreme Court. The pre-existing issue for the Circuit Courts of Appeals in applying trademark law under the Lanham Act was magnified by Congress' 1999 amendment to the Lanham Act expressly allowing trade dress protection for unregistered trade dress.[137] Despite the long line of Supreme Court cases that directly assert the limited protection of the patent laws as a monopoly, a new question is at stake, one that arises from congressional intent to protect trade dress.[138]

The Petitioner, TrafFix Devices, posited that trade dress protection should not be allowed where the article seeking protection was the subject of an expired utility patent.[139] The Petitioner argued for a per se rule against trade dress protection for expired utility patents. Thus, the Petitioner attempted to establish that MDI cannot be entitled to any further protection since its patents for the traffic sign stands expired before TrafFix copied it, and any different interpretation would be allowing MDI to neglect its end of the patent bargain.[140] Specifically, TrafFix argued that

---

135. *See supra* Part III, notes 88-134 and accompanying text.

136. *See Marketing Displays III*, *supra* note 11, at 929.

137. *See* 15 U.S.C. § 1125(a)(3) (Supp. V 1994) (acknowledges "civil actions for trade dress infringement under this chapter for [unregistered] trade dress"). *See also* Petitioner's Brief, *supra* note 28, at 26.

138. *See* 15 U.S.C. § 1125(a)(3) (Supp. V 1994).

139. *See* Petitioner's Brief, *supra* note 28, at 9.

140. *See* Petitioner's Brief, *supra* note 28, at 9. The MDI patents leave no room for doubt that the sign stand was novel, useful, and nonobvious, since a patent will not be granted if these three requirements are not met. *See* 35 U.S.C. §§ 101, 102, 103 (1994). The Petitioner discusses prior litigation in which MDI was involved to provide further proof that the product configuration at issue is completely covered by the expired utility patents. *See* Petitioner's Brief, *supra* note 28, at 22. Specifically, *Sarkisian v. Winn-Proof Corp.*, 697 F.2d 1313 (9th Cir. 1983), proved very valuable to TrafFix even though a per se rule was not adopted. Since the Supreme Court conducted an inquiry into functionality, *Sarkisian* was used to show the parts of the sign stand that were considered functional by MDI.

"[t]he public's right to copy from expired patents should not be subordinated to a factbound inquiry - dressed in terms of functionality - into whether there is a 'competitive need' to copy a particular configuration, at a particular point in time, in a particular market."[141] This argument is in fact bolstered by a prior Supreme Court decision, which found support in the purposes for the granting of a patent itself.[142] Two issues of particular interest are raised: the right to copy as part of the patent bargain and the doctrine of functionality. Each of these issues will be discussed in turn.

i. *The Patent Bargain*

A recurring theme throughout the Petitioner's Brief involved a statement made in *Bonito Boats, Inc. v. Thunder Craft Boats, Inc.*[143] There the Court declared that the underlying patent system is a "carefully crafted bargain" that balances an individual's exclusive right to invent and patent, while maintaining that the ultimate right will pass to the public.[144] The bargain "encourag[es] the creation and disclosure of new, useful, and nonobvious advances in technology and design in return for the exclusive right to practice the invention for a period of years."[145] Further, it is in "consideration of its disclosure and the consequent benefit to the community" that a patent is granted.[146] In fact, the Court in *Bonito Boats* addressed the question of state unfair competition laws that may conflict with federal patent protection.[147] However, the Petitioner argued that there is no difference between that discussion and the one currently before the Court involving federal Lanham Act issues. Whether the conflict arises between state and federal laws or two federal statutes, upon expiration of the patent, it is to return to the public domain for the public's benefit.[148]

The Petitioner argued that the Supreme Court not only repeatedly set

---

141. Petitioner's Brief, *supra* note 28, at 11.

142. *See id.* A number of Supreme Court cases have stood for the proposition that patents are to be granted for a limited duration, as set forth by the patent laws originating in the Constitution. TrafFix used many of these cases in its brief to support the proposition. *See* Petitioner's Brief, *supra* note 28, at 14-18. *See also* United States v. Steffens, 100 U.S. 82 (1879); Graham v. John Deere Co., 383 U.S. 1 (1966); Kendall v. Winsor, 62 U.S. 322 (1858); Sears, Roebuck & Co. v. Stiffel Co., 376 U.S. 225 (1964). It is worth noting that the Court in these cases was commenting on the accepted purpose of a patent, not on the potential overlap of protection from trademark law.

143. *See* 489 U.S. 141 (1989).

144. *See id.* at 150.

145. *Id.* at 150-51.

146. *Id.* at 151.

147. *See id.*

148. *See* Petitioner's Brief, *supra* note 28, at 16.

forth the purposes and limitations of the patent bargain, but it has also upheld the public's right to copy and use what was once protected in expired patents, since the right to copy is the benefit the public received for making the patent bargain.[149] "[T]he Court has repeatedly emphasized that unfair competition or trademark law may not be applied in a manner that would defeat the public's right to copy and use the subject of expired patents, or create what are in effect patent monopolies of unlimited duration."[150] The Petitioner used a number of cases to develop this history of the Court's decisions.[151] The Petitioner first likened its case to that of *Singer*, decided by the Court in 1896.[152] In *Singer*, the Court would not enjoin a competitor from producing a sewing machine that had the same overall appearance as Singer's, since Singer's utility patent had expired, and the monopoly protection on the distinctive character of its sewing machine expired with it.[153] The Court stated that "on the termination of the patent there passes to the public the right to make the machine in the form in which it was constructed during the patent."[154]

The Petitioner found further support for its proposition in a 1938 Supreme Court case, *Kellogg Co. v. National Biscuit Co.*[155] The *Kellogg* case was a dispute over Kellogg's right to make shredded wheat in a pillow-shaped form where National Biscuit had previously held product and process patents on the same shaped shredded wheat.[156] The Court essentially followed the *Singer* precedent and affirmed Kellogg's right to manufacture pillow-shaped shredded wheat. The only condition placed on Kellogg was that it somehow distinguish its product so that the consumer would in no way be confused by the similarity of the products.[157] Based on *Singer* and

---

149. *See id.* at 17-18.

150. *Id.* at 17.

151. In addition to those that will be discussed with particular detail, *see* Scott Paper v. Marcalus Mfg. Co., 326 U.S. 249, 256 (1945) and *Bonito Boats, Inc.*, 489 U.S. at 152.

152. Singer Mfg. Co. v. June Mfg. Co., 163 U.S. 169 (1896).

153. *See* Petitioner's Brief, *supra* note 28, at 18. *See also Singer*, 163 U.S. at 179.

154. *Singer*, 163 U.S. at 185. One amicus argued that *Singer* provides a bright-line rule that items formerly covered by patents necessarily enter the public domain, and therefore no part of the item could be the subject of Lanham Act protection. *See* Amicus Brief of Panduit Corp. in support of Petitioner at 4-7; Marketing Displays, Inc. v. TrafFix Devices, Inc., 530 U.S. 1260 (2000) (No. 99-1571). *But see* Amicus Brief of Thomas & Betts, at 12; Marketing Displays, Inc. v. TrafFix Devices, Inc., 530 U.S. 1260 (2000) (No. 99-1571) (stating *Singer* is not dispositive because it predates the Lanham Act).

155. *See* 305 U.S. 111 (1938).

156. *See id.* at 113, 117.

157. *See id.* at 120. Kellogg distinguished its product by using cartons that

Kellogg, the Petitioner concluded that "[i]f anything, the public's right to copy an article is only stronger where . . . it has enjoyed patent protection and, therefore, the public has conferred a patent monopoly upon the inventor *in exchange* for the right to copy the product upon expiration of the patent. That is the essence of the patent bargain."[158]

The amici in support of the Petitioner also had concerns over extending patent protection in light of the purposes set forth for issuing patents. The thrust of the arguments again resided in the bargain theory of patent: the weight of the constitutional right to grant patents found in the Intellectual Property Clause is heavier and cannot be overcome by the Lanham Act, enacted by Congress pursuant to the Commerce Clause.[159] A particularly potent point was that the patent bargain includes the good will associated with the patented product that, along with the patented invention, passes into the public domain at the end of the patent term.[160] If the good will of the product is part of the patent, and good will is essentially secondary meaning as necessary for trade dress protection, it follows that when the patent expires, there can be no trade dress protection.[161] Not only is the good will part of the patent, but it would be inequitable for the original patent holder to break her contract, so to speak, by then getting trade dress protection.[162] Further, the patent holder is not completely disadvantaged at the expiration of her patent, since the patent holders are the original owners and can use that fact in advertising to the public.[163] The original patent owner can assert the superiority of her product, having been the first to manufacture it. Since the patent owner has been given this advantage, if used to the fullest, lack of trade dress protection will have little adverse economic effect.[164]

The Petitioner's argument on the patent bargain was dependent upon the Court equating state unfair competition laws with the federal Lanham

---

displayed the Kellogg name. *See id.* at 120-21.

158. Petitioner's Brief, *supra* note 28, at 20.

159. *See* Amicus Brief of Malla Pollack Supporting Petitioner at 7, Marketing Displays, Inc. v. TrafFix Devices, Inc., 530 U.S. 1260 (2000) (No. 99-1571) [hereinafter Amicus Brief of Pollack].

160. *See id.* at 6-7 (arguing that the Court in *Kellogg* held that good will was included in the plaintiff's investment in a patent, and when that expired, all of the rights of protection did as well).

161. *See id.*

162. *See id.* at 9.

163. *See id.* at 10.

164. *See id.* at 11-12 (arguing that economics should not even be part of the discussion in considering trade dress protection after expired patents, since the right to copy stems from the Constitution).

Act,[165] as each of the cases presented by the Petitioner in support of its contentions about the patent bargain involved a state unfair competition claim. It seems fair to assume, however, that there is a difference between *state* unfair competition laws and *federal* recognition of trade dress protection. Certainly there is an identical purpose between the two, as the Lanham Act provides "national protection of trademarks in order to secure to the owner of the mark the goodwill of his business and to protect the ability of consumers to distinguish among competing producers,"[166] and state unfair competition laws attempt to prevent the same on a local level. But surely there is some weight in the fact that the Lanham Act is federally mandated protection of trademarks and trade dress. In other words, although state law can, and often is preempted by federal statutes, federal statutes are deemed "capable of coexistence, [and] it is the duty of the courts . . . to regard each as effective."[167]

Perhaps one of the most significant arguments that was set forth in the Thomas & Betts Amicus Brief supporting the Respondent is that the Petitioner's case analyses used to support this proposition were not dispositive in determining that there should in fact be no trade dress protection for expired patents.[168] Many of the cases used by the Petitioner predate the Lanham Act, so they are not precisely on point.[169] Those pre-Lanham Act cases used by Petitioner to fashion its per se rule are also unpersuasive since they "were all attempts to restrict the sale of utilitarian products through rights arising under state law. And in each case, [the Supreme] Court invalidated such claims based upon federal preemption under the Supremacy clause."[170] The MDI case differs in that there are two compet-

---

165. *See* Petitioner's Brief, *supra* note 28, at 25. The line of Supreme Court cases discussed earlier in this section deal with federal patent laws preempting state unfair competition laws, while the Lanham Act is another federal law enabling the protection of trade dress. *See supra* notes 143-58 and accompanying text.

166. Petitioner's Brief, *supra* note 28, at 25 (quoting Park 'N Fly, Inc. v. Dollar Park & Fly, Inc., 469 U.S. 189, 198 (1985)).

167. Respondent's Brief, *supra* note 99, at 13 (quoting Morton v. Mancari, 417 U.S. 535, 551 (1974)). It has been argued that the balancing of the federal statutes in this case is precisely what has forced the courts to create new tests to determine whether trade dress protection should be granted. *See* Gill, *supra* note 50, at 1290. However, as will be addressed, if the courts use doctrines already in place, such as the functionality doctrine, there is no need for them to create their own tests. *See infra* notes 297-314 and accompanying text.

168. *See* Thomas & Betts Amicus Brief, *supra* note 154, at 12-15.

169. *See, e.g.*, Singer Mfg. Co. v. June Mfg. Co., 163 U.S. 169 (1896); Kellogg Co. v. National Biscuit Co., 305 U.S. 111 (1938); Scott Paper Co. v. Marcalus Mfg. Co., 326 U.S. 249 (1945). The Lanham Act was passed in 1946. *See* 15 U.S.C. § 1051-1127 (1994).

170. Thomas & Betts Amicus Brief, *supra* note 154, at 12-13. The cases to

ing federal rights involved. In fact, the Lanham Act was not interpreted to protect the design of a product until after its enactment.[171] Thus, Thomas & Betts argued the Petitioner's primary support in its argument against trade dress protection is at best conjecture as to what the Supreme Court would have decided had there been a Lanham Act with which to contend. More importantly, the conjecture is futile because the Court is bound to decipher the conflict before it, so there is nothing to suggest which way the former cases would have come out since there was no conflict of federal laws.[172]

By recognizing trade dress protection under section 43(a) of the Lanham Act, it is true that Congress probably did not intend to extend the life of a patent monopoly, but it is far-reaching to conclude that Congress intended there should be no protection of product configuration at all, since the Lanham Act says just the opposite. It is presumed that Congress was aware of the extent of patent protection when they willingly extended the Lanham Act to protect trade dress in 1999.[173] Further, unfair competition laws and patent protection have almost always co-existed.[174] Thus, it is questionable whether the Petitioner is accurate in the presumption that prior Supreme Court cases involving state unfair competition laws are dispositive on the same issue if raised under the Lanham Act.

The Supreme Court cases used by the Petitioner raise questions in addition to the issue of state unfair competition versus federal law. Substantively, *Singer*[175] and *Kellogg*[176] seem to present some questionable precedent.[177] *Singer* proposes that the whole of a patented article passes into the public domain upon expiration of the patent, including the goodwill of the maker, design of the product and configuration of the product, despite clear palming-off by the competitor.[178] Even with clear examples of palm-

---

which Thomas & Betts are referring are *Sears*, *Compco*, and *Bonito Boats*, discussed *supra* notes 80-87 and accompanying text.

171. *See* Thomas & Betts Amicus Brief, *supra* note 154, at 13. It was in the 1970s and 1980s that the appellate courts began to recognize product configuration protection. *See Trade Dress Developments*, 624 PLI/Pat 599, 606 (2000).

172. *See* Thomas & Betts Amicus Brief, *supra* note 154, at 13-14.

173. *See* Petitioner's Brief, *supra* note 28, at 26.

174. *See* Bonito Boats, Inc. v. Thunder Craft Boats, Inc., 489 U.S. 141, 166 (1989) (patent protection and unfair competition laws have existed for almost 200 years).

175. 163 U.S. 169 (1896).

176. 305 U.S. 111 (1938).

177. See *supra* notes 149-158 for Petitioner's interpretation of why these two cases preclude trade dress protection for an expired utility patent.

178. *See Singer* 163 U.S. at 185. The defendant in the *Singer* case not only manufactured machines identical to Singer's, but also "numbered its machines with serial numbers in the millions in the same manner as Singer, even though the defendant had not manufactured millions of machines, and added a 'dummy' tension

ing-off by the defendant, the court would not allow trade dress protection to Singer because of expired utility patents.[179] Perhaps the Court's interpretation presupposes that no more than one form of intellectual property protection can be afforded to a single product. After the amendment of the Lanham Act, it seems presumptuous to conclude that any element of an invention that has been revealed in a patent necessarily becomes part of the public domain, even if it would unjustifiably confuse consumers.[180] As consumer markets expand and products become more readily available for purchase, such as on the Internet, the public's right not to be confused grows in importance, since consumers should be entitled to trust the quality of a product associated with a particular name. Hence, a right to copy a patented product exactly as it exists, configuration included, seems increasingly irrational.

Because more than one federal law may apply and confer protection to an article, the inquiry should be into more than merely the balance struck by the patent bargain. The inquiry should include whether there is a balance between the differing intellectual property rights bestowed. A patented article, upon expiration of the patent, may have developed a particular identity protectable as trade dress, and neglecting to recognize this protection may lead to a greater infringement of rights, namely stripping the patentee of trade dress rights to which they are entitled.

Even more, the Court itself in the *Singer* and *Kellogg* cases seems to recognize this possibility. In *Singer*, while the Court held that the sewing machine and the name "Singer" both became part of the public domain at the expiration of the patents, the Court also required that the prevailing defendant clearly indicate the manufacturing source on its own machines so that consumers would not be confused.[181] If the Court wanted the de-

---

screw in the same location as a working tension screw on the genuine Singer machines." Opderbeck, *supra* note 79, at 332. It is difficult to imagine a case where consumers would be more confused as to source and quality than in a case such as this, where the defendant intentionally "borrows" the goodwill of the former patent holder.

179. *See Singer*, 163 U.S. at 185.

180. *See* Opderbeck, *supra* note 79, at 369. Further, if there was a per se prohibition of trade dress rights for features of expired patents, "a patentee with . . . narrow claims could disclose a variety of alternative designs which, though not patentable, would then preclude competitors from using those designs to distinguish their products. This would be equivalent to obtaining trademark rights 'in gross,' without accompanying goodwill, something trademark law prohibits." *Id.* at 369-70.

181. *See* Thomas F. Cotter, *Is This Conflict Really Necessary?: Resolving an Ostensible Conflict Between Patent Law and Federal Trademark Law*, 3 MARQ. INTELL. PROP. L. REV. 25, 42-43 (1999).

fendant to indicate the source, this probably meant that the name "Singer" had not become generic, but did have a secondary meaning to the interested public, so that trade dress protection would have been appropriate.[182] The Court weakened its decision in *Kellogg* in the same respect by requesting that Kellogg clearly indicate that it was the source of its version of shredded wheat, despite the fact that the "name and shape [of the shredded wheat] had entirely entered the public domain."[183] Thus, it is questionable whether these two Supreme Court cases can justify a per se rule that trade dress protection is unavailable, since they seem to lend some credence to the opposing perspective: that there is merit in preventing consumer confusion via trademark protection of the patented configuration.

ii. *The Functionality Doctrine*

There is yet another pressing argument that trade dress protection should not be available in the wake of patent protection. It revolves around the long-standing functionality doctrine.[184] In *TrafFix Devices*, the Petitioner questioned the meaning and application of "functionality" in the context of trade dress protection for product configuration. The Petitioner argued quite persuasively that:

> [T]his Court itself has admonished that functional - or "useful" - product features are not entitled to trademark protection under Section 43(a). Moreover, Congress recently amended Section 43(a) to provide that in an action for trade dress infringement, the plaintiff "has the burden of proving that the matter sought to be protected is not functional." As the Committee Report accompanying this amendment explains, "functional marks should be dealt with under the patent law." Indeed, the Report goes on to observe that the amendment seeks to discourage owners of patents that are "about to expire" from attempting to register their invention as "a trademark, which receives protection as long as the owner uses it in commerce." Cloaking the subject of an expired utility patent – such as MDI's dual-spring sign stand – with trade dress protection would frustrate Congress' express intent to reserve such protection for marks that are "not functional."[185]

---

182. *See id.*

183. *Id.* at 44. Thus, on one hand the Court is suggesting that the name "Shredded Wheat" became generic, and was therefore unable to receive trademark protection, while on the other hand, it is imposing upon the party seeking to benefit from the expired patent the responsibility of differentiating their product so consumers do not become confused as to the source. *See id.* at 43-44.

184. *See supra* notes 130-34 and accompanying text (discussing the functionality doctrine).

185. Petitioner's Brief, *supra* note 28, at 30 (citations omitted).

The functionality principle is precisely that which the Petitioner sought to abolish.[186] The alleged failure of this doctrine is shown in part by the contradictory interpretations of the meaning of functionality by the circuit courts:

> Some lower courts (including the Sixth Circuit below) have construed the concept of functionality so that it is possible for a product or its features to be functional in a literal sense – and, thus, be useful and meet the requirements of utility patent protection – but not in a legal sense due to factors extraneous to the product itself, such as the competitive need to imitate it . . . . The public's right to copy from expired utility patents should not be subordinated to this judge-made and counter-intuitive concept of 'functionality.'[187]

In other words, a defense of functionality in an infringement suit would not work because the patent only protects useful features, but these useful features are not considered functional under trademark law, so functionality does not prevent an overlap of the patent laws and the trademark laws.[188] This argument is consistent with the decision in *Vornado Air Circulation Systems*.[189] A finding by a court "that there is no 'competitive need' for the public to enjoy the fruits of a particular discovery, at a particular point in time, in a particular market" would unduly extend the monopoly granted by a patent for an indeterminate amount of time, and is contrary to the very purposes of the patent laws.[190] Since a product feature will not qualify for utility patent protection unless it is useful, and a product feature is functional if it is essential to the use of the article, it debatably follows that a product feature covered by a utility patent is functional and cannot be covered by trade dress protection.

Indeed, for some interested amici, the functionality doctrine was the de-

---

186. *See id.* at 30-33.

187. *Id.* at 32.

188. *See* Mohr, *supra* note 57, at 389. It is noted that:

> There is a potential gap . . . between what is functional and solely the province of patent law, and what is merely "useful" and thus protectable under trademark law. Unlike the functionality doctrine, patent law requires only that an invention have a utilitarian function . . . not that it be the best or even one of the best . . . . The utilitarian nature of a feature in an invention alone thus will not preclude trade dress protection.

Id. at 391.

189. Vornado Air Circulation Sys., Inc. v. Duracraft Corp., 58 F.3d 1498 (10th Cir. 1995) (holding that a component is functional if a competitor needs that component to make an equivalent and competitive product).

190. Petitioner's Brief, *supra* note 28, at 38.

termining factor in resolving that trade dress protection should be unavailable for expired utility patents.[191] The United States argued that "where an expired utility patent shows that the features alleged as trade dress contribute to the operation of the claimed invention, those features must be considered 'functional' and thus ineligible for Lanham Act protection."[192] This assertion is founded in the purposes of patent rights and trademark protection, the latter of which "must be something other than, and separate from, the merchandise to which it is applied."[193] Thus, functionality is the doctrine that differentiates trademark and patent law, since any other interpretation would allow monopoly rights to useful trade dress features for much longer than the Constitutional grant intended to permit.[194]

Part of the fear surrounding the functionality doctrine is probably due to the fact-specific inquiries necessary to determine if a product is functional.[195] This leaves an implication that the doctrine is not clearly enough defined so as to be effectively applied by all of the circuits in the same way. The doctrine of functionality is in place so that a feature that is utilitarian cannot be protected to the detriment of those who wish to compete with the product, but cannot, due to the limited number of configurations available.[196] The following four possible factors in determining functionality have been identified:

> (1) the existence of a utility patent disclosing the utilitarian advantages of the design; (2) the existence of advertisements and promotions which tout the functional and utilitarian advantages of the product; (3) the existence of other alternative designs; and (4) whether the design results from a . . . superior method of manufacturing the article.[197]

It seems most logical that courts apply this doctrine, using the existence of a prior utility patent as one factor, rather than making it a determinative inquiry that trade dress protection should not be allowed. The fact that part

---

191. *See* Amicus Brief of the United States Supporting Petitioner at 7, Traf-Fix Devices, Inc. (No. 99-1571) [hereinafter Amicus Brief of the U.S.]. *But see* Amicus Brief of Pollack, *supra* note 159, at 8-9 ("[T]he trademark doctrine of 'functionality' is irrelevant to material which was protected by a patent; the patent bargain is an independent, constitutionally mandated limit on former patent holders' rights to exclude.").

192. Amicus Brief of the U.S., *supra* note 191, at 7.

193. *Id.* at 9 (*quoting* In re Morton-Norwich Prods., 671 F.2d 1332, 1336 (C.C.P.A. 1982)).

194. *See id.* at 10-13.

195. *See* Falk, *supra* note 3, at 834 n.67.

196. *See, e.g.*, Qualitex Co. v. Jacobson Prods. Co., 514 U.S. 159 (1995).

197. Falk, *supra* note 3, at 834-35.

of an invention is revealed in a patent and is therefore useful does not mean that this aspect has not come to be identified with a particular source. This seems acutely apparent where trade dress protection is sought for part of an invention, not the whole patented item.[198] In fact, the functionality doctrine may even be utilized to minimize the conflict between the Patent Act and Lanham Act.[199] An argument to do away with the functionality doctrine and replace it with a per se prohibition neglects to consider that it would be just as difficult for the court to determine whether a product configuration is disclosed in a utility patent under a per se prohibition as it is to determine if that configuration is functional.[200] These considerations, opposed by the Petitioner, were embraced by the Respondent, Marketing Displays, in its brief advocating for trade dress protection.[201]

B. *Respondent Marketing Display's Position*

Despite Petitioner's seemingly convincing brief opposing the recognition of trade dress protection in light of expired utility patents, the Respondent was left with a strong argument in favor of such protection, and this argument was supported by multiple amici briefs presented to the Court.[202] In essence, the Respondent saw no conflict between unfair competition laws and the federal patent laws.[203] Each of these intellectual property protections stems from different policy purposes and protects different features of a product, and, therefore, the two are not mutually exclusive.[204] But how can this be, given TrafFix's argument identifying the ostensible overlap between patent protection and trade dress protection?

Each of the arguments set forth by the Petitioner was, predictably, op-

---

198. *See* Marketing Displays, Inc. v. TrafFix Devices, Inc., 532 U.S. 23 (2001).

199. *See* Falk, *supra* note 3, at 838-39 (arguing that the Tenth Circuit in *Vornado* erred in creating the "significant inventive component" test since they could have used the same evidence that made the fan grill a significant component to find it functional instead, thus vitiating the need to address the conflict between patent law and the Lanham Act).

200. *See, e.g.*, Todd R. Geremia, Comment, *Protecting the Right to Copy: Trade Dress Claims for Configurations in Expired Utility Patents*, 92 NW. U. L. REV. 779, 815 (1998) (arguing that a per se rule is feasible because courts can compare the specifications of the utility patents to determine whether configurations are covered by the patents).

201. *See infra* notes 202-269 and accompanying text.

202. *See, e.g.*, Amicus Brief of Thomas & Betts, *supra* note 154; Brief of Amicus Curiae The American Intellectual Property Law Association in Support of Respondent, *TrafFix Devices, Inc.*, (No. 99-1571) [hereinafter Amicus Brief of AIPLA].

203. *See* Respondent's Brief, *supra* note 99, at 1.

204. *See id.* at 10.

posed by the Respondent in its brief. Thus, what follows will be an analysis of the counter-arguments to the patent bargain theory and the doctrine of functionality.[205] Additionally, however, the Respondent and its amici asserted that there is no true right to copy, that the remedies differ immensely for patent and trademark infringement, and that there are practical reasons why a per se prohibition should not be in place. Each of these ideas will be discussed forthwith.

i. *Balancing Two Federal Rights: The Patent Act and the Lanham Act*

1. *The Patent Bargain Revisited*

The Respondent argued that "all of the various intellectual property rights, including patents, copyrights, trademarks, trade dress and trade secrets, can exist concurrently in the same product without trespassing on each other."[206] Both trademark and patent protection are supplied by federal law, authorized by the Constitution.[207] Because trademark and patent laws are enacted and authorized constitutionally, absent Congressional intent to the contrary, both can co-exist and one cannot preempt the other.[208]

The Respondent and its amici concluded that the Lanham Act and the Patent Act are complementary, and to find otherwise would undermine the purposes of the acts themselves.[209] Not only can trademark and patent rights exist simultaneously by way of federal law, but they each embody such differing purposes that there seems to be no conflict in upholding them separately.[210] The policies underlying the Lanham Act "[help] to

---

205. *See supra* notes 143-201 and accompanying text (delineating Petitioner's discussion of the patent bargain and the functionality doctrine).

206. Respondent's Brief, *supra* note 99, at 11.

207. *See id.* at 13. *See also* 15 U.S.C. §§ 1051-58 (authorizing unfair competition and trademark protection); U.S. CONST., art. I, § 8, cl. 3 (Commerce Clause granting authority to Congress to pass federal legislation codifying trademark law); and 35 U.S.C. §§ 1-376 (the patent laws).

208. *See* Respondent's Brief, *supra* note 99, at 13. *See also* Morton v. Mancari, 417 U.S. 535, 551 (1974) (courts must find both trademark and patent laws effective unless there is a differing congressional intent); Thomas & Betts Corp. v. Panduit Corp., 138 F.3d 277, 285 (7th Cir. 1998) ("federal trademark protection cannot be preempted by patent law."). *See also* Davis, *supra* note 51, at 641 (stating that Congress is free to create protections outside of the Intellectual Property Clause, provided "efforts to evade the Clause's limits . . . create qualitatively different protection.")

209. *See* Respondent's Brief, *supra* note 99, at 16-28.

210. *See supra* notes 45-78 and accompanying text (discussing the public

stimulate competition by providing consumers with a means to readily identify the sources of goods and services in the marketplace."[211] Trademark and trade dress protect the owner while encouraging the public to make improvements on the product protected by the mark.[212] On the other hand, patent protection provides incentive to the public to invent by allowing inventors a limited monopoly in return for which the public will receive full disclosure of the new invention.[213]

This discussion leads to an important distinction made by the Respondent with respect to the truly unique rights provided by trademark and patent protection. A patent grants an exclusionary right; it does not "grant its owner the right to make or market a product."[214] What is allowed is the right for the patentee, for a limited time, to make, use or sell the invention himself, excluding others from copying his invention. Contrarily, trade dress protection is only permitted if the product has acquired such a distinctive identity in the marketplace that its source should be protected.[215] Thus, trade dress does not protect the invention itself, only the invention insofar as it has come to be recognized.[216] The Respondent did not deny that there exists a patent bargain that should be respected. Rather, the argument was that the rights differ so drastically as between patents and trade dress that the bargain struck between the inventor and the government when a patent is granted is a non-issue in a trade dress case.

The Respondent did not stand alone in its belief that the patent bargain will not be disrupted by trade dress protection. The amici were duly concerned with the way that the Petitioner interpreted patent and trademark laws. The Thomas & Betts Corporation stressed that, "a patentee does not receive through the patent any right to practice an invention, let alone an exclusive right . . . . [T]he only right received through a patent is the negative right to exclude others from making, using, and selling the invention

---

policy purposes behind trademark protection as opposed to patent protection).

211. Respondent's Brief, *supra* note 99, at 14.

212. *See id.* at 13-14.

213. *See id.* at 15.

214. *Id.* at 16. Further,

> [a] patent does not confer on its owner any positive property rights such as the right to make, to use, or to sell the patented invention. Instead, the patent owner is granted the right to exclude others from making, using, or selling the patented invention without the patent owner's permission for the duration of the patent.

Gill, *supra* note 50, at 1275.

215. *See* Respondent's Brief, *supra* note 99, at 17-18.

216. *See* Lanham Act § 43(a), 15 U.S.C. § 1125(a) (Supp. V 1994).

defined by the patent's claims."[217] Further, Thomas & Betts argued:

> Since a patent never gives its owner the right to practice the invention during the patent term, this nonexistent right cannot 'enure' to the public upon expiration. All that occurs when a patent expires is that the patent itself ceases to be an impediment to the public's practice of the invention – that is, to the extent the public would have the right to practice the invention in the absence of the patent, the expiration of the patent removes that obstacle.[218]

In support of this argument, Thomas & Betts offered the Patent Act of 1952, which clarified that a patent excludes others from using an invention; it does not grant others the right to use the invention.[219] The patentee herself may be prevented from use of the patent if there are later patents that have not expired and would prevent use, or she may be prevented from using the patent due to other laws that will not allow use of the invention.[220] Thus, the public does not actually have a "right" to practice a patented invention upon expiration of the patent, and it would be an erroneous interpretation of the patent laws to say otherwise.

2. *Is There a Right to Copy?*

The former argument is inextricably entwined with the argument that the right to copy a patented product on the expiration of the patent is not an absolute right, but a right implied by the time limitation placed on patent protection.[221] "The right to copy the subject matter of an expired utility patent has always yielded to other equally compelling policies, such as the unfair competition laws or state police power interests."[222] The Respondent was critical of TrafFix's attempt to interpret some Supreme Court cases so narrowly that they seem to support preempting trade dress protection, since

---

217. Amicus Brief of Thomas & Betts, *supra* note 154, at 8. Note that Thomas & Betts was involved in similar litigation involving a patent for cable ties, and upon expiration of the patent, Thomas & Betts successfully sought trade dress protection in 1994 before the Seventh Circuit Court of Appeals. *See* Thomas & Betts Corp. v. Panduit Corp., 138 F.3d 277 (7th Cir. 1998).

218. Amicus Brief of Thomas & Betts, *supra* note 154, at 8.

219. *See* 35 U.S.C. § 154 (1994).

220. *See* Amicus Brief of Thomas & Betts, *supra* note 154, at 10. (using, as an example, Eli Lily & Co. v. Medtronic Inc., 496 U.S. 661 (1990), where the owner of a patent on a drug still could not promote the use of that drug because there was no FDA approval).

221. *See* Respondent's Brief, *supra* note 99, at 26-27.

222. *Id.* at 27.

none of them explicitly recognizes a right to copy.[223] This weighed heavily in MDI's favor: federal trademark protection would supersede any right to copy if there is no explicit right to copy in the first place.

But there is clear Supreme Court precedent, pre-dating the Lanham Act, that recognizes a right to copy the material in expired patents.[224] So, is there really a right to copy?[225] It would be inaccurate to deny a right to copy at all; perhaps the issue should be the extent of the recognized right. It seems particularly contrary to the Intellectual Property Clause and the purpose of granting patents to conclude that an item is able to be copied as-is at the close of the patent period.[226] Perhaps a more plausible interpretation of the "right to copy" is that the idea passes to the public so that others may improve upon the invention, create an individual version of the invention, even develop a more identifiable version for consumers. However, allowing the copying of an invention exactly as it exists simply because that invention has passed into the public domain does not seem to be what was envisioned by the Intellectual Property Clause.[227] A right to copy is really a limited right for the public to copy what genuinely passes to the public domain, not any feature or item that is protected by other laws.[228]

3. *Secondary Meaning*

---

223. *See id.* at 28-29.

224. *See supra* notes 149-83 and accompanying text discussing the *Singer* and *Kellogg* cases thoroughly. Another Supreme Court case is often cited as recognizing the right to copy. *See* Scott Paper Co. v. Marcalus Mfg. Co., 326 U.S. 249, 250 (1945) (holding that a patent assignor has a complete defense to an action for infringement where the alleged infringing device is that of an expired patent). Significantly, the Court discusses copying in the *TrafFix* decision. *See* TrafFix Devices, Inc. v. Marketing Displays, Inc., 121 S.Ct. 1255, 1260 (2001) (stating ". . . in many instances there is no prohibition against copying goods and products").

225. *See* In re Mogen David Wine Corp., 328 F.2d 925 (C.C.P.A. 1964) (holding there is no absolute right to copy subject matter covered in an expired patent); Leatherman Tool Group, Inc. v. Cooper Indust., 131 F.3d 1011 (Fed. Cir. 1997) (holding there is no affirmative right to copy). *See also* Ethan Andelman, *Thomas & Betts Corp. v. Panduit Corp.*, 14 BERKELEY TECH. L.J. 229, 233 (1999) (discussing that there is no right to copy even though it may appear that the Supreme Court pronounced otherwise).

226. *See Singer*, 163 U.S. at 169 (1896).

227. *See* Andelman, *supra* note 225, at 238-41 (arguing that "an absolute right to copy exists only for items *fully* in the public domain, i.e., not protected by *any* body of intellectual property law"). Thus, a patented article may pass into the public domain as far as patent law is concerned, but there are other areas of law that may still prevent outright copying. *See id.* at 240.

228. *See id.*

The American Intellectual Property Law Association (AIPLA) filed as amicus and took another approach in its belief that the patent laws and the Lanham Act are complementary.[229] AIPLA asserted that a safeguard already in place is the necessity of secondary meaning in acquiring trade dress protection.[230] In other words, a trade dress can only be protectable if it is source identifying, not simply a useful feature.[231] The argument is that a per se rule should not apply because some manufacturers go to great lengths to develop secondary meaning, even during the term of their patents, and if they are automatically precluded from trade dress protection at the end of this term, they may be losing much of the goodwill they have made concerted efforts to develop.[232] With this additional precaution in practice, there is arguably little need for concern about an overlap of the Lanham Act and the patent laws. AIPLA removes the question of trade dress protection from conflict with the patent bargain by magnifying the differences between these two rights.

AIPLA grasps the heart of the policy for allowing both trade dress and patent protection for a single product.[233] With secondary meaning now essential to any trade dress protection recognition, it is the party seeking this protection that will have to show that it is the source recognized by consumers of the product.[234] The conflict between patent rights and trademark rights is thereby reduced. The right that will extend beyond the expiration of any applicable patent will not be the exclusionary right that initially belonged to the patentee, but the producer's right to continue to profit from her invention by reducing consumer confusion in the applicable market. In other words, the product configuration that produces evidence of secondary meaning is worthy of trade dress protection, just as any other product - even one without a prior patent - has had to produce secondary meaning.[235] Thus, there is no bias against the patentee simply for having

---

229. *See* Amicus Brief of AIPLA, *supra* note 202.

230. *See* Amicus Brief of AIPLA, *supra* note 202, at 17. Secondary meaning was not always necessary for every assertion of trade dress protection. But the Supreme Court changed this recently in *Wal-Mart*. *See infra* notes 273-309 and accompanying text.

231. *See* Amicus Brief of AIPLA, *supra* note 202, at 17.

232. *See id.* at 18-20.

233. *See supra* notes 229-32 and accompanying text.

234. *See* Wal-Mart Stores, Inc. v. Samara Bros., Inc., 529 U.S. 205 (2000).

235. Some examples of registered trade dress prove illustrative. LIFESAVER candies with a center hole, the LEVI tag on jeans, and the three stripes on ADIDAS products are all registered trade dress. *See* MCCARTHY, *supra* note 54, § 8:7. If each of these configurations were to be registered today, registration would require that secondary meaning has developed. But it is not difficult to see that these configurations clearly indicate source, and therefore, secondary

received a patent, and there is no deprivation to the public, who are still free to develop a competing product so long as they distinguish it sufficiently.[236]

The Respondent and its amici set forth the inherent differences between patent and trade dress protection succinctly. They argued that the patent bargain was not the correct inquiry in these circumstances. The focus should be on the presence of secondary meaning and the balancing of the differing policies underlying patent and trademark law. The Respondent then turned to the functionality doctrine to continue its argument that trade dress protection should not be prohibited for patent holders.

ii. *The Functionality Doctrine*

Of course, the Respondent addressed the conflict that exists when trade dress appears to protect a functional feature. Respondent argued for a functionality test that has been supported by a number of the Circuit Courts of Appeals.[237] Essentially, a feature would be functional if "constraints limit the available alternatives to the extent that others need the same trade dress in order to fairly compete."[238] Thus, a trade dress is nonfunctional if "the number of alternatives is sufficient that appropriation of one poses no great threat to competition."[239] In other words, functionality of a trade dress is defined according to the alternative configurations that are available to competitors.[240] With such an approach, a functionality inquiry may consider the existence of a patent, but its existence would only be one factor in determining whether the feature or product configuration is in fact functional, or whether it has many alternatives so that competi-

---

meaning would be fairly easily developed. If there were former patents on these objects, secondary meaning could still have developed during the course of the patent protection, and the configuration should thus receive the benefit of that protection.

236. *But see* Amicus Brief of Pollack, *supra* note 159, at 10 (arguing that the patentee is not disadvantaged for having received a patent even if she is denied trade dress protection).

237. *See, e.g.*, Fun-Damental Too Ltd. v. Gemmy Indus. Corp., 111 F.3d 993, 1002 (2d Cir. 1997); Intsy*Bit, Inc. v. Poly-Tech Industries, 95 F.3d 663, 673 (8th Cir. 1996); Tools USA and Equip. Co. v. Champ Frame Straightening Equip., Inc., 87 F.3d 654, 657 (4th Cir. 1996). It seems that this test originated from the Supreme Court itself, according to the Respondent's interpretation. *See* Qualitex Co. v. Jacobson Products Co., 514 U.S. 159, 165-66 (1995).

238. Respondent's Brief, *supra* note 99, at 20.

239. *Id.* Evidence of the availability of other alternatives is generally only one factor used in determining functionality, but it seems to be the one factor upon which the courts place most emphasis. *See* Mohr, *supra* note 57, at 388.

240. *See* Respondent's Brief, *supra* note 99, at 21.

tion is not stifled.[241] If this is the case, MDI should not be precluded from trade dress protection unless their dual-spring design is functional in that competitors would be forced to use alternatives that lack the functional and economic advantages of the design.[242]

It is noted that the functionality discussion is buttressed by the Lanham Act itself, as amended in 1999.[243] Congress has acknowledged that trade dress protection is acceptable for nonfunctional product features even if a patent did exist on that product.[244] MDI, therefore, wanted to stress that they have trade dress protection for the overall nonfunctional aspects of the traffic sign stands, not the functional features that were previously covered by patent protection.[245]

The AIPLA adamantly opposed the adoption of a per se rule as pursued by the Petitioner, TrafFix, because of the existence of the functionality doctrine.[246] At the forefront, AIPLA emphasized that the rights protected by a patent were not the same as those protected by trademark law under the Lanham Act. The Petitioner, therefore, was incorrect in its desire to impose a per se rule, especially because the Tenth Circuit case upon which Petitioner relied was incorrectly decided.[247] The complicated test set forth by the Tenth Circuit in its interpretation of earlier cases would lead to a lack of uniformity where there is already a safeguard in place: the Lanham

---

241. *See id.* at 21-22.

242. *See* Amicus Brief of AIPLA, *supra* note 202, at 17:

> [I]f courts properly apply the functionality doctrine, a manufacturer cannot use trade dress litigation to preclude competitors from copying a configuration that is necessary in order to practice a useful invention no longer covered by a patent. If a product's configuration meets the rigorous criteria for trade dress protection . . . it should not be unenforceable merely because the manufacturer previously obtained a patent on an invention its product embodies.

243. *See* 15 U.S.C. § 1125(a)(3) (Supp. V 1994) (trade dress rights available for nonfunctional features of a product configuration).

244. *See* Respondent's Brief, *supra* note 99, at 27. The House Report on the 1999 amendment states, "the owner of a patent that is about to expire may try to register it as a trademark, which receives protection as long as the owner uses it in commerce." H.R. Rep. No. 250 (106th Cong., 1st Session 1999). It further states, however, that "functional marks should be dealt with under the patent law." *Id.* Thus, the distinction still comes down to functionality.

245. *See* Respondent's Brief, *supra* note 99, at 27.

246. *See* Amicus Brief of AIPLA, *supra* note 202, at 10.

247. *See id.* at 10-12 (stating that "the Tenth Circuit created a new test to be applied when a defendant raises an expired patent as a defense to trade dress infringement. In contrast to the well-developed functionality test . . . the test applied in [the Tenth Circuit case] requires courts to construe patent claims to decide whether an element is a 'significant inventive aspect.'").

Act requires nonfunctionality for trade dress protection.[248] Further, "if courts properly apply the functionality doctrine, a manufacturer cannot use trade dress litigation to preclude competitors from copying a configuration that is necessary in order to practice a useful invention no longer covered by a patent."[249]

It is questionable whether application of the functionality doctrine in a proper way will be enough of a safeguard to prevent extending a patent monopoly. There is no reason why the functionality doctrine cannot be a useful reality. First and foremost, specifications that are necessary for patent applications often include many features that are not functional with respect to the patent, but reveal one mode of practicing the patent.[250] This makes it particularly burdensome to the patentee to have a per se rule prohibiting later trade dress protection, since the patentee will want to reveal as little as possible in her application so that she may later assert trade dress interests. If the patentee reveals as little of her invention as possible, there is a possible adverse effect on the public, since the public will not benefit from a more complete disclosure.[251] Moreover, even if a feature is disclosed that later comes to be a source-identifier, the mere fact that it has been disclosed in a patent is not a good indicator that it is essential, or utilitarian, to the whole invention.[252] Essentially, because the intellectual property rights sought are so fundamentally different, whether trade dress rights are bestowed should not depend on a patent, but should depend on a fresh inspection of the requirements for acquiring trade dress.[253] Nonfunctionality is precisely one of these requirements. It should be a primary inquiry before granting trade dress rights, but whether there is a patent should be nothing more than evidence in the functionality inquiry.[254]

iii. *Different Remedies, Different Rights*

Among all of the other distinct differences between patent protection

---

248. *See* Amicus Brief of AIPLA, *supra* note 202, at 12-13.

249. *See id.* at 17.

250. *See* Andelman, *supra* note 225, at 242-43 (illustrating that the cable tie at issue in the *Thomas & Betts* litigation happened to disclose an oval head in the patent, but the head could have assumed a different shape, since there is nothing to indicate that this shape was chosen for utilitarian reasons).

251. This is not necessarily true, since the patent laws require a complete disclosure of the best mode of practicing the invention. *See* 35 U.S.C. § 112 (1988).

252. *See* Andelman, *supra* note 225, at 243.

253. *See id.* at 241-43.

254. This, in fact, seems to be the direction the Supreme Court goes in its analysis of this case. *See* TrafFix Devices, Inc. v. Marketing Displays, Inc., 121 S. Ct. 1255 (2001).

and trademark protection, the Respondent included a brief analysis of the different remedies that are available in the case of infringement of one or the other.[255] It is argued that trade dress protection cannot be the equivalent of a perpetual patent in light of the remedies available. In the infringement of a patent, complete injunctive relief is the normal remedy, while the infringement of a trade dress may lead to a number of remedies that would still allow the infringer to use the trade dress to some extent.[256] In fashioning a remedy, "courts have required the use of disclaimers, house-marks or trade names, changes in product appearance, or corrective advertising to curtail the likelihood of confusion."[257] Therefore, a court may simply require an infringer of a trade dress to alter the trade dress somewhat, but allow continued use in that altered state. The public benefits because there will be no product confusion, and, in this case, TrafFix would be able to compete in the relevant market by making some changes to its product so that it did not infringe upon MDI's trade dress rights.[258] Since remedies are the thrust of any infringement case, the fact that complete injunctive relief is standard in a patent case and not in trade dress cases is another indication that the trade dress protection is not simply an extension of patent protection as argued by the Petitioner.

iv. *Practical Effects of a Per Se Prohibition*

Finally, Respondent attempted to undercut the Petitioner's argument by addressing a practical effect of precluding trade dress protection for all expired utility patents.[259] Essentially, the concern is that a per se rule mandating that all expired utility patents be deemed functional would "upset the harmonious balance struck between the patent laws and unfair compe-

---

255. *See* Respondent's Brief, *supra* note 99, at 24.

256. *See id.*

257. *Id.* In order for infringement of a trade dress to be found, it must be shown that the infringing product creates a likelihood of confusion by the purchasing public. This is codified in 15 U.S.C. § 1125(a)(1)(A) (1994). In other words, if two products simply appear to be very similar, that is not enough to find infringement. The Respondent, in its brief on page 22, uses as an example of the test for likelihood of confusion a Sixth Circuit case that lists the following as factors for likelihood of confusion: the strength of the plaintiff's mark; the relatedness of the goods; the similarity of the marks; evidence of actual confusion; marketing channels used; likely degree of purchaser care; defendant's intent in selecting the mark; and the likelihood of expansion of the product line. *See* Frisch's Restaurants, Inc. v. Elby's Big Boy, 670 F.2d 642, 648 (6th Cir. 1982). Other circuits use similar tests in determining the likelihood of confusion, and it is because of this essential test that the Respondent argues the marked differences in remedies available.

258. *See* Respondent's Brief, *supra* note 99, at 25.

259. *See* Respondent's Brief, *supra* note 99, at 37-38.

tition laws."[260] As examples, Respondent refers to the Coca-Cola® bottle, the Volkswagen Beetle, and the Porsche 911 as famous shapes that were protected by expired utility patents, and now have trade dress protection that protects the public from confusion.[261] After all, "such a rule would create the nonsensical position that trade dress protection could be available for a product that was never disclosed in a utility patent -- although it met the statutory requirements for patentability -- but not for a product that was actually disclosed in a utility patent."[262]

It seems that this would in turn affect the patent system since "[t]he completeness of patent disclosures would be limited out of concern that the disclosure may later prevent enforcement of any trade dress rights in the same type of product which is the subject of the patent."[263] Amici to the Respondent addressed this issue as well.

Thomas & Betts asserted that there will be impractical effects if the position of the Petitioner is adopted. Specifically, there is strong opposition to a per se rule whereby no trade dress rights would be recognized if there was an expired utility patent that disclosed a product feature.[264] "In full effect, to prevail in a trade dress claim under the Lanham Act, a claimant would not only have to prove all existing elements of a claim, but also that the trade dress feature is novel," which is a patent requirement, not a trade dress requirement.[265] Because this per se rule would only apply to patent holders, there would also be an adverse effect on the patent system, since those seeking patent protection may be reluctant to do so and disclose in rich detail what could be used against them in a later dispute over trade dress.[266]

Extending this argument further leads to an interesting conclusion. Thomas & Betts suggests that the Petitioner desires all materials in the public domain be ineligible for protection under the Lanham Act. If this is so, not only products entering the public domain via expired patents, but, for example, those that became public through textbooks or foreign patents would also be ineligible for trade dress protection.[267] Otherwise, there

---

260. *Id.* at 38.

261. *See id.*

262. *Id.*

263. Respondent's Brief, *supra* note 99, at 40.

264. *See* Amicus Brief of Thomas & Betts, *supra* note 154, at 16-18.

265. *Id.* at 16.

266. *See id.* at 17.

267. *See id.* at 18. "Suppose a product feature appeared in *Popular Mechanics*, but the person who created the feature never obtained a patent. If one accepts the arguments of Petitioner . . . why should there be a distinction between product features which entered the public domain upon expiration of a patent and [product features] which entered the public domain through other channels . . . ?" *Id.*

would be discriminating effects against patent owners.

It is not clear that a per se rule prohibiting trade dress protection for expired utility patents would be any easier to apply,[268] although, it is argued, an advantage is that it is preferable because there will be no fear of litigation for those who wish to copy a product at the expiration of a patent.[269] However, the ease of application is only one factor to consider in formulating such a profound rule. While there might be a reduction in litigation, there would most certainly be reluctance in applying for patents at all in certain circumstances where the inventor suspects that the product would prove to be a success on the market. Thus, there are practical reasons for avoiding a per se law prohibiting trade dress protection for expired patents.

V. RECENT SUPREME COURT ACTIVITY

Although the Supreme Court has not directly addressed the issue of the competing policies of trade dress under the Lanham Act and the Patent Act before the *TrafFix* case, *Bonito Boats*[270] represents the first case in which the Court recognized that patent rights "represent . . . a 'competing' federal policy that must be considered when addressing a claim under Section 43(a) of the Lanham Act."[271] In order to better understand where the progression of the Court's investigation of trade dress protection has ended up, an analysis of the recent *Wal-Mart* case is helpful.[272]

A. *The Supreme Court and Trade Dress in* Wal-Mart v. Samara Brothers

The Supreme Court granted certiorari to the Second Circuit to hear the *Wal-Mart* case on the limited question of what must be shown to establish that a product's design is inherently distinctive for purposes of trade dress protection.[273] The case involved the design of children's clothing into

---

268. *See* Todd R. Geremia, *Protecting the Right to Copy: Trade Dress Claims for Configurations in Expired Utility Patents*, 92 NW. U. L. REV. 779, 815 (1998).

269. *See, e.g.*, Ronald J. Horta, *Without Secondary Meaning, Do Product Design Trade Dress Protections Function as Infinite Patents?*, 27 SUFFOLK U. L. REV. 113 (1993).

270. 489 U.S. 141 (1989).

271. William D. Raman, *Litigating Copyright, Trademark & Unfair Competition Cases for the Experienced Practitioner*, 624 PLI/Pat 599, 607 (2000).

272. *See* Wal-Mart Stores, Inc. v. Samara Brothers, Inc., 529 U.S. 205 (2000).

273. *Id.* at 1341.

which Samara, the clothing designer, claimed to have integrated particular designs so that the product would be identified with Samara.[274] As a test for determining if the trade dress was inherently distinctive, the Second Circuit asked "'whether the design was likely to be understood as an indicator of the product's source.'"[275] Based on this test, the Second Circuit affirmed the lower court's decision that "Wal-Mart's marketing of the knockoffs was willful piracy with an intent to deceive consumers as to the source of the product."[276] But the Supreme Court reversed, ultimately deciding that product design was not inherently distinctive.[277] The Court was explicit in stating that protection of product design could be had only where there is proof of secondary meaning.[278] Essentially, the Court was interested in whether the trade dress to be protected was so distinct that any consumer would identify the maker with the design.[279] With this decision, the Court slowed the growth of trade dress protection that has occurred in the recent past.[280] No longer is any trade dress product protectable without secondary meaning, even if the trade dress is arbitrary, fanciful or suggestive, all of which are protectable without secondary meaning under trademark law. It suffices to say that the Court considers all trade dress equivalent to descriptive trademarks.[281]

The Court was, therefore, forced to reconcile the *Wal-Mart* decision with its earlier decision in *Taco Cabana v. Two Pesos*, where it was held that inherently distinctive trade dress was protectable without secondary meaning.[282] Without particular difficulty, it was deciphered that the trade

---

274. *See* Samara Brothers, Inc. v. Wal-Mart Stores, Inc., 165 F.3d 120, 126 (2d Cir. 1998).

275. *See id.* at 125 (quoting Landscape Forms, Inc. v. Columbia Cascade Co., 113 F.3d 373, 378 (2d Cir. 1997).

276. *Supreme Court Refines Trade Dress Protection Guidelines:* Wal-Mart Stores v. Samara Bros., 6 No. 17 ANDREWS INTELL. PROP. LITIG. REP. 3, 3 (Apr. 5, 2000).

277. *See* Raman, *supra* note 271, at 610.

278. *Wal-Mart Stores, Inc.*, 529 U.S. at 211. The requirement of secondary meaning for all trade dress changes the prior rule that inherently distinctive trade dress was protectable upon its use. *See* Siegrun D. Kane, *Trade Dress After* Wal-Mart: *"Rolling Back" Product Design Protection*, 616 PLI/Pat 653, 657, 663 (2000). After *Wal-Mart*, "to achieve protection for trade dress consisting of an unregistered product design, plaintiff must prove that consumers perceive the design as a source indicator." *Id.* at 657.

279. *See* Amicus Brief of AIPLA, *supra* note 202 and accompanying text.

280. *See* Kane, *supra* note 278, at 663.

281. This is true at least in the case of unregistered trade dress. *See Wal-Mart Stores, Inc.*, 529 U.S. at 211.

282. *See* Taco Cabana Int'l., Inc. v. Two Pesos, Inc., 932 F.2d 1113, 1120 n.8 (5th Cir. 1991), *aff'd*, 505 U.S. 763 (1992), *reh'g denied*, 120 L. Ed. 2d, 947, 113

dress in *Taco Cabana* was product packaging, which does not need secondary meaning to be protected.[283] This has significance for future cases, including *TrafFix*, since it will have to be determined whether the trade dress sought to be protected is product packaging or product design, the latter of which will not be protected without secondary meaning.[284]

It is difficult to determine just how much the *Wal-Mart* decision influenced the Court's decision in *TrafFix*, although it is clear that this recent decision had some impact, and, as the following will indicate, the *Wal-Mart* decision aided as a predictor of where the Court was willing to go with trade dress. It appears that the Court was unwilling to extend protection of trade dress any further than its recent expansion, this in light of greater concern for securing free competition.[285] Perhaps the Court's focus upon secondary meaning can shed some light on an alternative solution to resolving conflict between patent rights and trade dress rights. An exploration of secondary meaning leads to a conclusion that it can -- and now must -- be a phase in an inquiry of whether trade dress protection is appropriate. If secondary meaning can be sufficiently established, the functionality of the product configuration should be explored, according to a specific test. If it is determined that the configuration is non-functional, trade dress protection should be available, despite the possibility of the existence of a prior patent. Each of the phases in this suggested analysis will now be set forth.[286]

i. *Secondary Meaning*

There are two possible ways to interpret the Court's decision in *Wal-Mart* in an application to cases of overlap between trade dress and patent law. One interpretation of the true suggestion in the *Wal-Mart* decision in relation to the *TrafFix* case is that "a manufacturer or producer would not be placed at a disadvantage by not having protection upon adoption under

---

S. Ct. 20 (1992).

283. *See Wal-Mart Stores Inc.*, 529 U.S. at 210.

284. It may seem fairly easy to make a distinction between product packaging and product design, but usually this is not the case. An example would be the purchase of a Coca-Cola® bottle. Some, and probably most consumers buy the bottle for the cola inside. But there are collectors that are interested only in the configuration of the bottle itself. Would a Coca-Cola® bottle be protectable as product design or product packaging? Would the consumer's intent suddenly become an issue? *See* Respondent's Brief, *supra* note 99, at 38.

285. *See Wal-Mart Stores, Inc.*, 529 U.S. at 214.

286. Note that perhaps the better order of inquiry is first, whether functionality is established, and next, whether there is a secondary meaning. This is suggested by the Court in the *TrafFix* decision. *See infra* notes 310-16 and accompanying text.

trade dress theories because it could initially secure protection by obtaining a design patent or copyright for the design."[287] Such logic, carried to an extreme suggests that "one may initially obtain a design patent or copyright on a product design and hold the world at bay under patent or copyright laws until secondary meaning is acquired, at which time it will hold the world at bay under trade dress theories."[288]

According to this interpretation, the establishment of secondary meaning sufficiently allows the patentee to protect her trade dress, just as any other party would be free to do, despite the prior existence of the patent. Additionally, fear of a monopoly for an indefinite period of time is unfounded, since under the patent laws, the patented invention will pass into the public domain to the extent that no secondary meaning has been established. Those configurations that have adequately developed as source-identifiers are then tested for functionality,[289] as defined under *trademark* laws, and if they are non-functional, the product is entitled to trade dress protection.

Contrarily, the reasoning in *Wal-Mart* can be used to illustrate that trade dress protection should *not* be available for the expired utility patent.[290] It is argued that

> [t]rade dress in unregistered product configurations requires secondary meaning. "Secondary meaning" in the appearance of the article is the "good will of the article." This good will of the article belongs to the public at the end of the patent term. The holder of an expired patent may not, therefore, claim an exclusionary right to "secondary meaning" in any aspect of the article formerly marketed under the patent's umbrella. Without secondary meaning, he may not claim trade dress.[291]

It follows, therefore, that a per se rule should be in place excluding trade dress protection for any product that has enjoyed patent protection, since secondary meaning is nothing more than good will.[292]

However, this interpretation is dependent upon an absolute right to copy expired patents. The judicially-created right to copy is not absolute, but is merely a right to copy that, in fact, does pass to the public under patent laws, not something that remains protected under other laws.[293] With only a limited right to copy, it follows that only a limited part of the invention

---

287. *See* Raman, *supra* note 271, at 611.
288. *Id.*
289. *See infra* notes 297-309 and accompanying text.
290. *See* Amicus Brief of Malla Pollack, *supra* note 159, at 6-7.
291. *Id.* (citations omitted).
292. *See id.*
293. *See supra* notes 221-28 and accompanying text.

passes into the public domain. Those elements of the invention that have secondary meaning are protected under the Lanham Act, so the goodwill associated with those configurations is still entitled to protection. The Lanham Act is precisely one of the "other" laws that is triggered at the close of a patent term.[294] It would be violated if the entire invention protected by patent automatically passed into the public domain.

Both of the aforementioned interpretations of the meaning of the *Wal-Mart* decision revolve around the product obtaining secondary meaning. The author proposes that if the inquiry remains with the acquisition of secondary meaning – or lack thereof – much of the conflict between trade dress protection and patent protection will be resolved.

Secondary meaning has been criticized as difficult to determine because its acquisition is generally a factual inquiry.[295] However, courts are adept in determining secondary meaning, as they frequently employ a test of secondary meaning in cases involving trademark.[296] Thus, it should remain the first step in a determination of trade dress, followed by whether the product is functional.

ii. *The Functionality Doctrine as a Solution, Not a Problem*

The Court in *Wal-Mart* imposed the requirement that all trade dress develop sufficient secondary meaning before protection is provided.[297] However, the *Wal-Mart* decision clearly did not in any way prohibit the conclusion that trade dress is available for expired patents.[298] Prior to the *TrafFix* decision, it was suggested that however the Court chose to resolve the *TrafFix* issue, the functionality doctrine would not be the test that the Court would adopt in its resolution, since the functionality doctrine lacks precision.[299] Just the opposite is true.[300] The Court employed the function-

---

294. *See* 15 U.S.C. § 1125 (Supp. V 1994).

295. *See, e.g.*, Todd R. Geremia, *Protecting the Right to Copy: Trade Dress Claims for Configurations in Expired Utility Patents*, 92 NW. U. L. REV. 779, 786 n.44 (1998) (stating secondary meaning often requires considering "whether a company encourages consumers to identify its product by its trade dress in its advertising, the length and manner of a company's use of the trade dress, and whether the public actually does identify the company's product through its trade dress").

296. *See generally* J. THOMAS MCCARTHY, MCCARTHY ON TRADEMARKS AND UNFAIR COMPETITION § 14.01 (4th ed. 1996).

297. *See Wal-Mart Stores Inc.*, 529 U.S. at 211.

298. *See supra* notes 273-96 and accompanying text.

299. *See* Perry J. Saidman, *Kan TrafFix Kops Katch the Karavan Kopy Kats? Or Beyond Functionality: Design Patents Are the Key to Unlocking the Trade Dress/Patent Conundrum*, 82 J. PAT. & TRADEMARK OFF. SOC'Y 839, 856-57 (2000).

ality doctrine to hold that trade dress protection was not available in this case.[301]

Indeed, it would have been troubling if the Court decided to do away with the functionality doctrine. First, the functionality doctrine is a way to even the playing field between those seeking trade dress rights who have expired patents and those who do not, since both will have to prove non-functionality before they secure trade dress protection.[302] This does not undercut the bargain theory of patents, since those features covered in the patent that are functional still pass to the public domain.[303] Further, the party seeking the trade dress protection is left with the burden of proving non-functionality,[304] a task that may be made more difficult if they once owned a patent on the invention. A functionality test also allows the differences between patents and trademarks to continue to exist. In a practical sense, the usefulness of an invention necessary to secure a patent[305] really does differ from the functional aspects precluded from trade dress protection.[306] Utility as a prerequisite to issuance of a patent is to insure that the invention is operable, or "functions for its intended purpose."[307] This is to say that the patented invention as a whole must be operable or it is unable

---

300. *See* TrafFix Devices, Inc. v. Marketing Displays, Inc., 121 S. Ct. 1255, 1259-60 (2001).

301. *Id.* at 1261.

302. *See* Hoffman, *supra* note 63, at 174.

303. *See* Malla Pollack, *The Owned Public Domain: The Constitutional Right Not to be Excluded – Or the Supreme Court Chose the Right Breakfast Cereal in* Kellogg v. National Biscuit Co., 22 HASTINGS COMM. & ENT. L.J. 265, 293-94 (2000) (arguing that holders of expired patents give no consideration in the patent bargain "[i]f the consideration question is reduced to functionality").

304. *See* Lanham Act § 43(a), 15 U.S.C. § 1125(a) (Supp. V 1994).

305. *See* 35 U.S.C. § 101 (1994) (stating "whoever invents or discovers any new and useful . . . manufacture . . . or any new and useful improvement thereof, may obtain a patent thereof, [subject to other requirements]").

306. *See* Geremia, *supra* note 295, at 801-02 (stating "it is the nonreputation-related, market-determined necessity of incorporating a product configuation, not its actual utility [as necessary in a patent context], that determines [a product's] functionality").

307. CHISUM ET AL., PRINCIPLES OF PATENT LAW 729 (1998). Not only is the usefulness requirement for a patent dependent upon it being operable, but obtaining a patent does not require the inventor to prove the superiority of the form of her invention over other products. *See* Geremia, *supra* note 295, at 787-88 n.53. The inventor need only show that hers is the best mode of the invention as she *subjectively* contemplates it to be. *See* 35 U.S.C. § 112 (1994). This is relevant since there may be, and often are, better ways to embody an invention than the way the patentee chose to do so. Since this is so, there is even more positive proof that protecting the product configuration will not impede competition by monopolizing, as others are free to use the invention and develop better methods or embodiments.

to be protected by patent. But it is possible for part of a patent to have developed meaning to the relative community that can now operate as trade dress. So long as this element of the patent is not one that is so substantially a part of the invention that it is functional, it is feasible that there can be later protection of the trade dress.[308]

Functionality in application may create substantial issues since it is primarily a factual question. However, courts are equipped to determine a question of fact if they focus on "whether the feature's availability to other manufacturers gives consumers the benefit of a competitive market, or if the item is incidental, merely distinguishes the brand, and other competitors can do without it."[309] Regardless of the fear over application of the functionality doctrine, it endures, as the Supreme Court readily illustrated.

B. *The Supreme Court Decision in* Marketing Displays v. TrafFix

As discussed, the arguments in the *TrafFix* case occurred in November 2000, and the Supreme Court decision was passed down March 20, 2001.[310] Its conclusion was that a former patent is strong evidence of functionality, and this being so, protection could not be provided to Marketing Displays for its dual-spring sign design.[311] Functionality is the premise on which this decision turns.[312] In fact, the Court states that "[f]unctionality having been established, whether MDI's dual-spring design has acquired secondary meaning need not be considered."[313] This is more than ample evidence that the Court is not only not going to do away with the functionality inquiry, but the functionality inquiry will become primary in determining whether trade dress is protectable in the wake of a patent, and sec-

---

308. There is some case law that supports this proposition. *See, e.g.*, Zip Dee, Inc. v. Dometic Corp., 931 F. Supp. 602 (N.D. Ill. 1996) (holding that trade dress protection was available for a slatted awning cover that was disclosed in a utility patent, but was not functional since it was a dependent element of another claim, not an independent claim itself).

309. Robyn L. Phillips, *Determining if a Trade Dress is Valid*, 29 IDAHO L. REV. 457, 477 (1992/1993).

310. *See TrafFix Devices, Inc.*, 121 S. Ct. at 1255.

311. *See id.* at 1260. The Court focused particularly on the fact that in 1983, MDI was the plaintiff in litigation related to the dual-spring design. In that case, MDI argued that springs spaced apart in a competing design infringed the patent despite the fact that their own patent had springs placed close together. MDI won this argument. This basically precludes MDI from now arguing that the dual-spring design was not covered as a functional feature in the patent. *See id.* at 1260-61. *See also* Sarkisian v. Winn-Proof Corp., 697 F.2d 1313 (9th Cir. 1983).

312. *See TrafFix Devices*, 121 S. Ct. at 1260-62.

313. *Id.* at 1262.

ondary meaning will only be explored once it has been determined that the product to be protected is non-functional.[314]

Despite the strength with which the Court employs the functionality doctrine, it seems that it was particularly pertinent in this case, so that Marketing Displays had little or no chance to meet its burden of proving non-functionality.[315] The Court does not extend its decision to hold that the Intellectual Property Clause of the Constitution precludes trade dress protection for a product covered by an expired utility patent.[316] In application, however, there seems to be little chance that the holder of an expired patent will successfully meet her burden of proving that features are non-functional. Notwithstanding this steep burden of proof, it appears that trade dress protection is indeed available for expired utility patents.

CONCLUSION

The Supreme Court's decision in *Marketing Displays, Inc. v. TrafFix Devices, Inc.*, provides guidance in addressing the question of whether trade dress protection should be available for a product configuration formerly the subject of a utility patent. This case was a proper vehicle to illustrate that trade dress protection should be a viable option for configurations, despite prior patents, since trade dress protection and patent protection are two separate and unique forms of intellectual property protection. Existence of a patent should only be one element in a functionality test to decipher whether the particular configuration in question is of a functional nature. If the configuration is functional, trade dress protection should not be afforded. Likewise, secondary meaning must also be established to show that the configuration is properly identified with a source. If both of these phases are satisfied, the owner of the product has proved that she has a trade dress worthy of protection.

Alison Marcotte

314. *See id.*
315. *See id.*
316. *See id.* at 1263.

Literal Interpretation and "Absurd" Results: *Commonwealth v. Wallace* and "Trial on the Merits"

Francis D. Doucette*

I.

Familiar to lawyers and laypersons alike, the phrase "trial on the merits"[1] has a rock-solid, reassuring ring to it. Regardless of the nature of the grievance or controversy, those four words hold out hope and the promise of legal redress, suggesting fairness, finality, a meaningful opportunity to be heard. Moreover, when the Legislature sees fit to insert those powerful and evocative words in a statute, it is reasonable to assume that the valuable legal rights embodied are inviolable, beyond tampering or attack. But in the recent case of *Commonwealth v. Wallace*[2] the Supreme Judicial Court of Massachusetts not only dashed those assumptions but abandoned a great many of its own hallowed principles along the way. Sadly, it is a decision which illustrates the lengths to which a court will go—including jettisoning the English language—in order to arrive at a socially acceptable result.

The focus of the case is the interpretation of statutory language, and "[s]tatutory interpretation is a pure question of law."[3] The underlying facts, however, are uncomplicated and easy to recite. In 1998 the defendant was charged in Quincy District Court[4] with violating a "permanent" restraining order sought and obtained by the Quincy Housing Authority in Norfolk Superior Court pursuant to Massachusetts General Laws Chapter 121B, section 32F. The order[5] prohibited the defendant (and five other

---

* B.A., University of Massachusetts, 1970; M.A., Harvard University, 1972; Ph.D., Harvard University, 1974; J.D., Northeastern University School of Law, 1981. Francis D. Doucette is a member of the Massachusetts and Virginia bars and has been in private practice since 1982. He has been a part-time clinical supervisor at Suffolk University Law School since 1990.

1. Defined in BLACK'S LAW DICTIONARY (7th ed. 1999) as "[a] trial on the substantive issues of a case, as opposed to a motion hearing or interlocutory matter."

2. 730 N.E.2d 275 (Mass. 2000).

3. Commonwealth v. Cintolo, 613 N.E.2d 509, 510 (Mass. 1993).

4. Commonwealth v. Wallace, No. 9856-CR-006424.

5. Norfolk Superior Court Civil Action No. 96-01692, Judgment Dated April 17, 1997, Judith A. Cowin, J.

named individuals) "from entering or remaining within one thousand (1000) feet"[6] of specified property of the Quincy Housing Authority.

The last sentence of the second paragraph of Chapter 121B, section 32F, states in forceful, seemingly unambiguous terms that "[n]o permanent [restraining] order shall be granted except as a final judgment after a trial on the merits."[7] And any order so issued, whether temporary or permanent, is required by the first paragraph of the same statute to "contain the following statement: VIOLATION OF THIS ORDER IS A CRIMINAL OFFENSE" (capitals in original).[8] Thus there is no question that Chapter 121B, section 32F, is a penal statute; the maximum penalties are a fine of up to $3,500 and imprisonment for up to two years in the house of correction, or both.

It was uncontroverted that the defendant never had a trial on the merits because he defaulted in Superior Court; even the final judgment prominently[9] acknowledged that fact. Indeed, at the time of the civil trial the defendant was incarcerated. Relying on this background and the plain wording of the statute, the defendant's attorney filed a motion and a memorandum[10] seeking dismissal of the complaint. The primary contention was that the "permanent" restraining order issued was a legal nullity since no trial on the merits had ever taken place, in contradiction of the statutory mandate. Accordingly, the Commonwealth could not prevail in a criminal trial because it could not establish a knowing[11] violation of a valid order. Secondarily, the defendant exhorted the court to respect the legislative intent in insisting on a trial on the merits and to abide by the maxim that criminal statutes must be strictly construed.

On these (or some other) grounds the district court judge, of course, could have dismissed the complaint sua sponte.[12] But the Commonwealth, assisted by counsel for the Quincy Housing Authority, opposed the motion to dismiss, and, after a hearing, the judge denied it. Conceding, however, that "these are questions of law which are important and where the law is in doubt,"[13] the judge, pursuant to Rule 34, reported two questions of law[14]

---

6. *Id.*

7. MASS. GEN. LAWS ch. 121B § 32F (2001).

8. *Id.*

9. *See* Norfolk Superior Court Civil Action No. 96-01692, Judgment Dated April 17, 1997, Judith A. Cowin, J.

10. *See* Commonwealth v. Wallace, No. 9856-CR-006424.

11. The defendant, however, was served in hand with a copy of the restraining order. *See* Commonwealth v. Wallace, 730 N.E.2d 275, 277 (Mass. 2000).

12. "We take this occasion to note that a judge has inherent authority to dismiss an indictment sua sponte. This is a necessary corollary of the judge's authority to process criminal cases." Commonwealth v. Jenkins, 727 N.E.2d 1172, 1174-75 (Mass. 2000).

13. Report pursuant to M.R.C.P. 34 of Mark S. Coven, J., Docket No. 9856-CR-

to the appeals court.

The Supreme Judicial Court then transferred the case on its own initiative, albeit skewering the action as procedurally defective on several fronts. The defendant, the court first maintained, should not be allowed to challenge the validity of the permanent restraining order in the motion to dismiss the criminal case. Such a challenge, in the court's view, amounted to an impermissible "collateral attack,"[15] an approach disfavored and rejected by the court in an appeal of a criminal contempt case just the year before.[16] The court went on to explain that other options and remedies[17] were available to the defendant, had he wished to contest the issuance of the restraining order; none of which the defendant pursued.

Additionally, the court even questioned the propriety[18] of the district court judge reporting the questions of law to the appeals court, intimating that the matter was not appropriate for interlocutory appeal. Indeed, ultimately the court ruled that "[t]he report is discharged as not being properly before us."[19] All of this caviling fairly begs the question why the court ever bothered to accept such a supposedly error-riddled case for review in the first place. Nonetheless, once having disinfected the air of these procedural irregularities, the court proceeded swiftly to an analysis of the defendant's main contention, agreeing with the district court judge that "the issues raised by the reported questions are important."[20]

II.

The analysis is abbreviated, occupying barely more than a page, and is more remarkable for what it leaves out than for what it includes. Throughout, the court seems uneasy or reluctant even to acknowledge that it is dealing with a criminal case. For example, the concept of liberty and the potential loss of that liberty never figure in the discussion. The court quotes or cites a total of six cases in the analysis, five of which are civil and all from Massachusetts, and never hints that special concerns are sup-

---

006424.

14. The two questions were: "(1) Can the Superior Court, pursuant to M.G.L.c.121B, sec.32A et seq., grant a permanent injuction on behalf of a public housing authority only after a trial on the merits? (2) Must the District Court dismiss a criminal complaint for a violation of a permanent injunction entered pursuant to M.G.L.c.121B, sec.32F where that injunction entered after the default of the defendant and not after a trial on the merits?" *Wallace*, 730 N.E.2d at 276-77.

15. *Id.* at 277.

16. *See* Commonwealth v. Dodge, 705 N.E.2d 612 (Mass. 1999).

17. *See Wallace*, 705 N.E.2d at 277.

18. *See id.* at 276 n.1.

19. *Id.* at 278.

20. *Id.* at 276-77.

posed to apply in the interpretation of a criminal statute. Obviously, the court knows better, as it has admitted elsewhere in analogous circumstances: "[t]he considerations underlying the proper application of a criminal statute are different."[21]

There is no mention in the analysis whatever of the celebrated rule that "[c]riminal statutes are to be construed strictly against the Commonwealth and in favor of the defendant, under the premise that '[n]o one may be required at peril of life, liberty or property to speculate as to the meaning of penal statutes."[22] No idle academic exercise, this rule—sometimes referred to as the rule of lenity—has its origins in the common law[23] and can be traced to the beginnings of the Republic.[24] The historical basis of the rule is generally assumed to be the cruel and barbarous punishments of earlier ages, which enlightened judges and courts sought to circumvent by fashioning and applying the practice of strict construction.[25]

To be sure, there is authority to the effect that the rule of strict construction, as an aid in interpretation, is limited to statutory language which admits of ambiguity. "There is no need of construction to ascertain the meaning of a statute where the language is clear and unambiguous and the intention of the law-making power is manifestly apparent therefrom."[26] A 1984 decision[27] of the United States Supreme Court reflects this sentiment: "The rule of lenity is of course a well-recognized principle of statutory construction . . . but the critical language of [18 U.S.C.S. § 1001] is not sufficiently ambiguous, in our view, to permit the rule to be controlling here."[28]

But that narrow position, limiting the rule of strict construction to cases clouded by ambiguous statutory language, is far from universally held. Almost two centuries ago, Chief Justice John Marshall observed that "[t]he rule that penal laws are to be construed strictly, is perhaps not much

---

21. Commonwealth v. Chapman, 744 N.E.2d 14, 17 (Mass. 2001).

22. Commonwealth v. McLaughlin, 726 N.E.2d 959, 966 (Mass. 2000) (citations omitted).

23. "It is required at common law that a criminal statute be strictly construed, and that any doubt be resolved in favor of the defendant." WHARTON'S CRIMINAL LAW (15th ed. 1993), § 12 and cases cited. *See also* PAUL H. ROBINSON, CRIMINAL LAW DEFENSES (1984), § 35(a)(2) n.6 (citing Livingston Hall, *Strict or Liberal Construction of Penal Statutes*, 48 HARV. L. REV. 748, 750-51 (1935): "The rule of strict construction was developed in England before the 19th Century, when courts sought to ameliorate the harsh penalties imposed for many offenses by narrowing the scope of a statute through strict construction.").

24. *See* United States v. Wiltberger, 18 U.S. 76 (1820).

25. *See supra* note 23 and accompanying text.

26. People v. Lund, 46 N.E.2d 929 (Ill. 1943).

27. *See* United States v. Rogers, 466 U.S. 475 (1984).

28. *Id.* at 484.

less old than construction itself."[29] A leading authority on Massachusetts criminal law asserts succinctly and without qualification that "a criminal statute must be construed strictly. The court may not enlarge a criminal statute by implication."[30] To illustrate, relying on that general principle, and making no finding or argument regarding ambiguity, the court in *Commonwealth v. Cintolo*[31] affirmed the allowance of a motion to dismiss by ruling, as a matter of law, that the defendant corporate president "was not the 'person having employees,' within the language of Massachusetts General Laws Chapter 149, section 150C."[32] In *Commonwealth v. Boone*,[33] which dealt with the definition of "escape" in Massachusetts General Laws Chapter 268, section 16, the court engaged in no discussion of ambiguity but firmly declined to enlarge upon the statute under review: "To create such a fiction on the language of s[ection] 16 alone runs contrary to the usual tenet of strict construction of a criminal statute."[34]

Of course, regardless of whether ambiguity is a factor, there are limits to the rule of strict construction. For example, in *Commonwealth v. Roucoulet*,[35] the defendant was arrested in a motor vehicle within 1000 feet of a school zone in possession of a "large quantity of cocaine, cutting agents,"[36] baggies, and the like. He elected to plead guilty to the indictment which charged him with trafficking in cocaine, but disputed, both at a jury-waived trial and on appeal, the separate indictment which charged him with violation of the school zone statute. Proof was lacking, the defendant contended, that he had any intention to distribute these drugs inside the school zone; on the contrary, even the Commonwealth's evidence established convincingly that the drugs were destined to be delivered to an undercover police officer then waiting at a co-defendant's home outside the prohibited 1000 foot area. The court, however, refused to read into the statute any requirement that there must be proof of an intent to distribute within the school zone. "The maxim that criminal statutes are to be strictly construed," Justice Greaney wrote, "does not mean that an available and sensible interpretation is to be rejected in favor of a fanciful or perverse one."[37] The more general rule, as stated in a United States Supreme Court decision[38] from 1976, is that "[a] criminal statute . . . is to be strictly con-

29. *Wiltberger*, 18 U.S. at 95.

30. JOSEPH P. NOLAN & LAURIE J. SARTORIO, 32 MASSACHUSETTS PRACTICE § 8 (3d ed. 2001).

31. 613 N.E.2d 509 (Mass. 1993).

32. *Id.* at 359.

33. 477 N.E.2d 1026 (Mass. 1985).

34. *Id.* at 1029.

35. 601 N.E.2d 470 (Mass. 1992).

36. *Id.* at 471.

37. *Id.* at 473.

38. Barrett v. United States, 423 U.S. 212 (1976).

strued, but it is 'not to be construed so strictly as to defeat the obvious intention of the legislature.'"[39]

III.

In *Wallace*, however, the court refrains from ever declaring, at least explicitly, whether it finds ambiguous the phrase "trial on the merits" in Chapter 121B, section 32F. The accepted wisdom is that "[a] term is ambiguous only if it is susceptible of more than one meaning and reasonably intelligent persons would differ as to which meaning is the proper one."[40] Certainly, when inclined, the court knows how to say that some word or combination of words is "not ambiguous"[41] or that the words of a statute "are plain."[42]

Whether this omission in *Wallace* is inadvertent or calculated, it obviously allows the court the luxury to bypass the rule of strict construction and deny the defendant the deferential treatment to which he would otherwise be entitled. "If the statutory language 'can plausibly be found to be ambiguous,' the rule of lenity requires the defendant to be given 'the benefit of the ambiguity.'"[43] Similarly, in the absence of an explicit finding of ambiguity, "there is no occasion to examine legislative history,"[44] and no justification to "look to extrinsic circumstances to determine the intent of the Legislature as to [the statute's] meaning."[45]

The polar opposite of ambiguous language is clear or plain language. "It is a well-established canon of construction that, where the statutory language is clear, the courts must impart to the language its plain and ordinary meaning."[46] As it should be, "the statutory language itself is the principal source of insight into the legislative purpose."[47] It would be ludicrous and futile to suggest that the Legislature was not familiar with the

---

39. *Id.* at 218 (quoting American Fur Co. v. United States, 2 Pet. 358, 367, 7 L. Ed. 450 (1829) and citing Huddleston v. U.S., 415 U.S. 814, 831 (1974)).

40. County of Barnstable v. American Fin. Corp., 744 N.E.2d 1107, 1109 (Mass. App. Ct. 2001).

41. *Roucoulet*, 601 N.E.2d at 473.

42. Daluz v. Dept. of Corr., 746 N.E.2d 501, 508 (Mass. 2001) (stating "[t]he words of G. L. c.30, s.58, are plain.").

43. Commonwealth v. Carrion, 725 N.E.2d 196, 198 (Mass. 2000) (quoting *Roucoulet*, 601 N.E.2d at 473).

44. *Roucoulet*, 601 N.E.2d at 472 n.7.

45. EMC Corp. v. Comm'r of Revenue, 744 N.E.2d 55, 57 (Mass. 2001).

46. Commonwealth v. One 1987 Mercury Cougar Auto., 600 N.E.2d 571, 573 (Mass. 1992). *See also* Caminetti v. United States, 242 U.S. 470, 485 (1917) (stating "[w]here the language is plain and admits of no more than one meaning, the duty of interpretation does not arise, and the rules which are to aid doubtful meanings need no discussion.").

47. Hoffman v. Howmedica, Inc., 364 N.E.2d 1215, 1218 (Mass. 1977).

legal meaning and significance of "trial on the merits" when it drafted section 32F of Chapter 121B. "It is to be presumed . . . that the legislature is acquainted with the law, that it has knowledge of the state of the law on subjects on which it legislates . . . ."[48] Beyond doubt, the phrase "trial on the merits" in section 32F of Chapter 121B is a laudable example of the deliberate use of clear, plain language. Only a sophist or contrarian would contend otherwise. But the court in *Wallace* declines to give those four words either praise or meaning.

Instead, the court, in its abbreviated analysis, defeats the defendant's claim by selecting and championing one dubious, obscure canon of statutory construction, at the expense of all other canons. It borrows and quotes this canon from a 1979 civil case:[49] "[a] literal construction of statutory language will not be adopted when such a construction will lead to an absurd and unreasonable conclusion . . . ."[50] The subject matter of the 1979 civil case—a wrangling over the interpretation of the phrase "any town" in a dispute between two towns concerning the reimbursement of veterans' benefits—obviously has nothing to do with criminal jurisprudence and the potential loss of liberty at stake in the instant case.

Nonetheless, with that doctrine as its sole guide, the court acknowledges but disagrees with the defendant's argument that the phrase "trial on the merits" in Massachusetts General Laws chapter 121B, section 32F, should be given a literal interpretation. "[S]uch an interpretation," the court declares flatly, "would lead to an absurd result."[51] What would that absurd result be? "It would mean," the court endeavors to explain, "that no permanent injunction could ever issue against an individual who defaulted in a civil action where a restraining order was sought . . . ."[52]

Even if that consequence is conceded, is that not a matter for the Legislature to address, by an appropriate amendment to Chapter 121B, section 32F? As the Massachusetts Supreme Judicial Court has stated in various forms when confronted with similar problems before, "[t]he solution to the dilemma . . . is in the hands of the Legislature."[53] But ignoring for the moment the avenue of legislative correction of a possibly flawed statute, is the court being honest when it depicts the present state of affairs in

---

48. Commonwealth v. Boarman, 610 S.W.2d 922, 924 (Ky. App. 1981). *See also* Commonwealth v. Smith, 710 N.E.2d 1003, 1005 (Mass. App. Ct. 1999) *aff'd*, 728 N.E.2d 272 (Mass. 2000) (stating "[f]urther, '[t]he Legislature must be assumed to know the preexisting law and the decisions of [the Supreme Judicial C]ourt.'").

49. *See* Town of Lexington v. Town of Bedford, 393 N.E.2d 321 (Mass. 1979). A more recent expression of this rule may be found in Johnson Lumber v. Woodscape Homes, Inc., 746 N.E.2d 533, 536 (Mass. App. Ct. 2001) (citing various authorities).

50. *Town of Lexington*, 393 N.E.2d at 327.

51. *Wallace*, 730 N.E.2d at 278.

52. *Id.*

53. *Dodge*, 705 N.E.2d at 616.

Wallace as an absurd result? Or, rather, is it a result the court simply does not like and refuses to let stand? Is the decision based on accepted, well-settled principles of constitutional criminal law? Or, rather, is it the product of political correctness and social engineering? No fewer than four public housing authorities filed amicus briefs in the case.[54] Of course, "[t]he longstanding rule is that an amicus curiae is not a party to the case . . . ."[55]

The court adopts an alarmist tone and hastens to point out that one of the principal purposes of the statute is "to allow a housing authority to create a safe public housing environment, by protecting tenants from the unlawful actions of nontenants."[56] Such a goal is clearly commendable, but has it been thwarted or made impossible by the statutory requirement of a trial on the merits prior to the issuance of a permanent restraining order? The court implies as much but provides no data whatsoever in support. Indeed, the opinion is utterly devoid of statistics, crime reports, or accounts of frustrated efforts by public housing authorities to obtain permanent restraining orders to curb incidents of violence or disorders. Rather, without any such data and without apology, the court elevates public policy concerns over the rights of the defendant, paying scant attention to any other concerns. In the process, the court in essence treats "trial on the merits" as surplusage and all but deletes the phrase from Chapter 121B, section 32F.

Oddly, the court never defines the operative label "absurd," apparently viewing the word as self-defining.[57] Nor does the court ever define "trial on the merits" or discuss the legal history or etymology of the phrase. While "not dispositive," "[d]ictionary definitions are helpful . . . in statutory interpretation cases."[58] Equally disturbing, the court fails to cite a single authority, from any jurisdiction, where this particular canon of statutory construction was resorted to, dispositively or otherwise, in a criminal case.[59]

There are, moreover, no treatises on statutory construction mentioned or cited in *Wallace*. The pre-eminent work on the subject informs us that

---

54. *Wallace*, 730 N.E.2d at 276 n.2.

55. *In re* Harvard Pilgrim Health Care, Inc., 746 N.E.2d 513, 518 (Mass. 2001).

56. *Wallace*, 730 N.E.2d at 278.

57. "Absurd" is defined as "[o]ut of harmony with reason or propriety; incongruous, unreasonable, illogical." OXFORD ENGLISH DICTIONARY 43 (2d 1989).

58. Commonwealth v. Smith, 710 N.E.2d 1003, 1005 n.4 (Mass. App. Ct. 1999), *further appellate review granted*, 714 N.E.2d 825 (Mass. 1999), *aff'd* 728 N.E.2d 272 (Mass. 2000).

59. The single criminal case cited in the analysis, *Commonwealth v. Williams*, 691 N.E.2d 553 (Mass. 1998), is inapposite, having nothing to do with claims of absurd results but dealing rather with the proper interpretation of the juvenile transfer statute, the former Massachusetts General Laws Chapter 119, section 61.

"there is authority for applying the plain meaning rule even though it produces a harsh or unjust result or a mistaken policy as long as the result is not absurd."[60] Little further elucidation is provided, however, and the overwhelming majority of cases cited in support of that general proposition are civil.

But even if it is agreed that the principle applies equally to criminal cases, where does a court draw the line between "mistaken policy" and "absurd results?" What is the difference between the two? Does a result become absurd simply because a court says so? Why is the situation in *Wallace* an absurd result rather than an accepted statutory "anomaly?"[61] What, finally, is absurd about a defendant in a criminal case, facing a potential loss of liberty of two years if convicted, asking that the clear, plain language of a statute be adhered to? Isn't this a far more reasonable request than the case of a defendant urging, to quote Justice Greaney, a "fanciful or perverse"[62] interpretation of a statute?

Admittedly, adherence to the statutory requirement of a trial on the merits would likely impose a hardship on officials of public housing authorities and others, including the courts. This would be particularly true if it could be demonstrated persuasively that individuals like the defendant in *Wallace* default in large numbers when these civil trials are scheduled. But, no such figures documenting court defaults appear in the text or the footnotes of the court's opinion. By the court's own standards, hardship, in these circumstances, is not a factor that should be taken into account. "We cannot interpret a statute so as to avoid injustice or hardship," the same court has written, "if its language is clear and unambiguous and requires a different construction."[63]

It is not uncommon for appellate courts, when scrutinizing statutory language under challenge in criminal cases, to arrive at results which do not appear to serve society's interests or carry out the apparent legislative intent. Yet, there is not, and should not, be a rush to label all these results as absurd. "We have no doubt," Justice Wilkins wrote in *Commonwealth v. Boone*, about the escape of the defendant (determined to be a sexually dangerous person and serving a term of from one day to life), "that the law

---

60. NORMAN J. SINGER, STATUTES AND STATUTORY CONSTRUCTION, 2A § 46.01 at 83 (5th ed. 1992).

61. *See* DaLuz v. Dept. of Correction, 746 N.E.2d 501, 508-09 (Mass. 2001) (stating "[w]e recognize that the statutory language [MASS. GEN. LAWS ch. 30, §58] appears to creatye an anomaly: a partially disabled employee injured by violence of a prisoner or a patient may recover more assault pay benefits than a totally disabled employee.").

62. *Roucoulet*, 601 N.E.2d at 473. For further discussion of *Roucoulet*, see *supra* notes 35-37 and accompanying text.

63. Rosenbloom v. Kokofsky, 369 N.E.2d 1142, 1144 (Mass. 1977) (citing Milton v. Metro. Dist. Comm'n, 172 N.E.2d 696, 700 (Mass. 1961)).

should provide a penalty for what the defendant did. No statute, or combination of statutes, however, does so."[64] More recently, the Massachusetts Appeals Court has ruled that handwritten notes containing sexually graphic language and propositions allegedly left (along with money) by an adult in the desk of a fifth-grade student do not constitute "printed material" for purposes of the obscenity statutes.[65] Similarly, the Supreme Judicial Court has concluded that the statute[66] which criminalizes the act of making annoying telephone calls is not violated when an individual instead sends a series of allegedly annoying facsimile transmissions to a television station.[67]

Further examples and variations of such results abound in the criminal law, all of which could easily be viewed as absurd since the net effect is that the offensive conduct goes unpunished. But it is not the function of the courts to be writing or re-writing laws or labeling as absurd results which do not conform to the court's way of thinking. "Courts are not at liberty to impose their views of the way things ought to be simply because that's what must have been intended, otherwise no statute, contract or recorded word, no matter how explicit, could be saved from judicial tinkering."[68]

IV.

It is in the very last sentence, however, of the analysis in *Wallace* that the court stands its own reasoning upside down and exposes how shallow and disingenuous its position really is. "A sensible construction of the provision in question," the court writes, "is that no permanent [restraining] order shall be granted until there has been an *opportunity* for a trial on the merits."[69] Having waited until the end to volunteer that interpretation, the court adds, in a footnote, that "[w]e need not speculate as to the reason for the language in the statute."[70] But the glaring reality is that the court is doing precisely that, speculating, qualifying, "tinkering" with the legislative language. Putting it another way, the inescapable conclusion is that, despite its silence on the subject, the court indeed *does* find ambiguous the phrase "trial on the merits" in Chapter 121B, section 32F. To achieve its desired objective, the court cannot bring itself to admit this ambiguity.

---

64. *Boone*, 477 N.E.2d at 1029.

65. *See* Commonwealth v. O'Keefe, 723 N.E.2d 1000, 1005 (Mass. App. Ct. 2000).

66. MASS. GEN. LAWS ch. 269, § 14A (1998).

67. *See* Commonwealth v. Richards, 690 N.E.2d 419 (Mass. 1998).

68. Kilpatrick v. Superior Court of Arizona, 466 P.2d 18, 27 (1970).

69. *Wallace*, 730 N.E.2d at 278 (emphasis in original).

70. *Id.* at 278 n.3.

When there is ambiguity in the language of a penal statute, the defendant by a long, courageous tradition is entitled to the benefit of that ambiguity. This court, however, lacks the courage and intellectual honesty to confer on him that benefit.

Even more egregious, the court in this shattering conclusion to its analysis violates bedrock principles of judicial and legislative separation of powers. Only four years earlier, the same court practically boasted of its ability to recognize the limits of its authority by quoting strong language from its earlier cases: "We have traditionally and consistently declined to trespass on legislative territory in deference to the time tested wisdom of the separation of powers as expressed in art. XXX of the Declaration of Rights of the Constitution of Massachusetts even when it appeared that a highly desirable and just result might thus be achieved."[71]

Suggesting or implying that the legislature really meant an "opportunity" for a trial on the merits is more than an innocent "trespass on legislative territory." "We are not permitted to add words to a statute that 'the Legislature did not see fit to put there,'"[72] the same court has spoken. "We must construe the statutes as they are written."[73] The real, practical difference between trial on the merits and an "opportunity" for trial on the merits is not a subtle nuance. It is a difference in kind, not in degree, a wholesale revamping of the statute, a diminution from statutory insistence on the real thing to a mere "opportunity" for the real thing. There is not a whisper, a hint in Chapter 121B, section 32F, that that is what the legislature intended.

In *Commonwealth v. Smith*,[74] a very different but instructive case, the Massachusetts Appeals Court agreed with the prosecution that "the defendant's conduct [digital penetration and oral intercourse with his eleven-year-old daughter], if true, is shocking and abhorrent."[75] Yet, the court affirmed the defendant's motion to dismiss the two indictments charging incest (additional indictments charged indecent assault and battery and assault and battery) on the grounds that the defendant's alleged acts did not constitute "sexual intercourse" within the meaning and purview of Chapter 272, section 17. "We have no right," Justice Smith wrote, "to read into the incest statute 'a provision which the legislature did not see fit to put there. . . .' To do so would amount to judicial legislation, which is

71. Pielech v. Massasoit Greyhound, Inc., 668 N.E.2d 1298, 1302 (Mass. 1996) (citations omitted).

72. Commonwealth v. Russ R., 744 N.E.2d 39, 43 (Mass. 2001) (citations omitted).

73. Brennan v. Election Comm'rs of Boston, 39 N.E.2d 636, 638 (Mass. 1942).

74. 710 N.E.2d 1003 (Mass. App. Ct. 1999).

75. *Id.* at 1005.

forbidden by Article 30 of the Massachusetts Declaration of Rights."[76]

By urging that the legislature really meant an "opportunity" for trial on the merits and not an actual trial on the merits in Chapter 121B, section 32F, the court in *Wallace* for all practical purposes added to the statute "a provision which the Legislature did not see fit to put there."[77] That interpretation may be a sensible one, in the eyes of many, but the crucial language needed to support that interpretation is altogether missing. Moreover, there are other, equally sensible interpretations of the statute, as written. Statutory insistence on a trial on the merits may convey, for example, a legislative disdain for permanent restraining orders. It is just as plausible to assume that the legislature was expressing a preference that permanent restraining orders be issued only after all parties have had their day in court.

Far from being a careful, thoughtful analysis, the decision in *Wallace* is the judicial equivalent of a "quick fix." It is a triumph of political correctness and expediency over the rights of a criminal defendant to the orderly process of law. It represents not only judicial activism running amuck and trampling over the explicit wishes of the legislature but a cavalier disregard for the sanctity of the English language. It is an erosion of individual rights that society in the long run can ill afford.

76. *Id.* at 1006 (citations omitted).

77. *Id.* (citing King v. Viscoloid Co., 106 N.E. 988, 989 (Mass. 1914)).

Is There Anything "Legal" About Extralegal Action? The Debate Over Dorr's Rebellion

Paul M. Thompson*

Most histories are about change. Those that have been written about Dorr's Rebellion, an uprising in Rhode Island in 1842, are in line with this general rule. Nearly all the books and articles that have been written about Dorr's Rebellion claim that it marked a turning point in the constitutional, legal, and political history of one of America's smallest states. The story goes like this: before Dorr's Rebellion, Rhode Islanders lived under an English Charter. After the Rebellion, they abandoned that outdated document and adopted their own American constitution. Before the Rebellion, Rhode Island was the only state in the union that still limited the vote to property holders. After the Rebellion, these property qualifications were generally swept away. Before the Rebellion, a majority of Rhode Islanders were shut out of the state's legal and political institutions. After the Rebellion, however, Rhode Island's formal institutions became accessible to those who were previously excluded. One historian, Patrick Conley, neatly summed up the changes that took place in Rhode Island as a result of Dorr's Rebellion. Prior to 1842, Conley noted, Rhode Island was a "democracy in decline,"[1] while after the Rebellion, it had joined the modern democratic world.

This article takes a different approach to Dorr's Rebellion. It focuses not on how things changed, but on how they stayed the same. Change was inevitable in Rhode Island in 1842. Even those who opposed the Rebellion conceded that reforms were needed to make the state's legal institutions more open and more responsive to the needs of its citizens.[2] By

---

* Law Clerk, United States District Court, Western District of Pennsylvania; J.D. 1998, Harvard Law School.

1. *See generally* PATRICK T. CONLEY, DEMOCRACY IN DECLINE: RHODE ISLAND'S CONSTITUTIONAL DEVELOPMENT, 1776-1841 (1977).

2. *See* WILLIAM M. WIECEK, THE GUARANTEE CLAUSE OF THE U.S. CONSTITUTION 96 (1972). Wiecek notes that "[t]he *sine non qua* of the [opponents of Dorr's Rebellion] . . . was legality: reforms were conceded, however reluctantly, but they could be brought about only in a legitimate way." *Id. See also* Letter from John

highlighting the changes that occurred during, and as a result of, the crisis in Rhode Island, most authors have overlooked the real significance of Dorr's Rebellion. It is not the ends achieved that makes Dorr's Rebellion a memorable event. Instead, it is the means used by reformers to achieve their ends that makes Dorr's a notable snapshot in American legal history -- and these means were not new.

Those who supported the Rebellion in Rhode Island were defenders of extralegal action.[3] They believed that Americans had a legal right to make and enforce law without proceeding through the formal legal institutions set up by the state. Accordingly, in 1842, Thomas Wilson Dorr and his supporters [hereinafter the "Dorrites"] met in their own extralegal convention, drafted and adopted their own extralegal constitution, and established their own extralegal government. The Dorrites did all of this without the approval of the existing government of Rhode Island.[4] They found support for their actions in the *Declaration of Independence*, the adoption of the *United States Constitution*, and the writings of the founding fathers.[5] To the Dorrites, the American Revolution had legitimized a certain kind of legal activity – namely, extralegal action. To them, extralegal action was, and had always been, "legal."

Those who opposed Dorr's Rebellion had a different view. Looking at the same history and interpreting the same documents, they drew a very different conclusion about the legality of extralegal action. The American Revolution was not about means -- resistance and revolution -- to Dorr's opponents. It was about ends: establishing free, republican governments in America. Following independence from Britain, and the establishment of constitutional governments in America, riots, mobs, and other forms of extralegal action were no longer acceptable. In the middle of the nineteenth century, the only legitimate role for the people was to obey, and not resist, their government.[6] To the opponents of Dorr's, any other conduct was strictly illegal. It was this debate that raged in Rhode Island in 1842, and this debate was not new.[7]

Extralegal action was the foundation of the colonial resistance during the 1760's and 1770's. Decades before the actual outbreak of war with

---

Whipple to John Tyler (Apr. 9, 1842), *in* INTERFERENCE OF THE EXECUTIVE IN THE AFFAIRS OF RHODE ISLAND, 28TH CONG., H.R. REP. NO. 546, at 669 (1844) [hereinafter BURKE'S REPORT].

3. *See infra* Part I, for a detailed discussion of the meaning of extralegal action.

4. *See infra* Part II.

5. *See infra* Part III, for a detailed discussion of the legal theory espoused by the supporters of Dorr's Rebellion.

6. *See infra* Part IV.

7. *See infra* Part I.

Britain, Americans in all the colonies took to the streets to protest unlawful acts of the British government. Out of these extralegal mobs emerged extralegal governments. In the 1770's and 1780's, few patriots questioned the legality of extralegal action. After the Revolution, however, serious questions about the practice began to emerge. Even Samuel Adams, perhaps the most vocal proponent of extralegal action before and during the Revolution, began to question its continuing vitality.[8] Although Americans debated the legality of extralegal action throughout the 1780's and 1790's, they never resolved the conflict. By the end of the eighteenth-century, old and new ideas about extralegal action sat side-by-side and in constant tension.[9] This tension resurfaced in 1842 during Dorr's Rebellion.

The story of Dorr's Rebellion cannot be told on a blank canvas -- it must be placed in the proper historical and political context. Parts I and II of this article add background and bits of color to our portrait of Dorr's Rebellion. Part I discusses the tradition of extralegal action in America, while Part II presents the facts and circumstances that led many Rhode Islanders to turn to extralegal action in 1842. Parts III and IV, then, fill in the details of our portrait. Part III shows how the Dorrites embraced the legacy of extralegal action and Part IV presents the argument of the opponents of Dorr's Rebellion. By the end of the article, our portrait of Dorr's Rebellion is complete. It reveals a nation in conflict about the legitimacy of extralegal action even in the middle of the nineteenth century. It is a portrait that is remarkable for its similarity to the past.

With few notable exceptions, most historians have viewed the legacy of extralegal action as an eighteenth-century anomaly that somehow evaporated or changed beyond recognition after the end of the American Revolution. Yet, as this article explains, more than fifty years after the close of the war with Britain, this eighteenth-century anomaly still survived and many Americans still debated its significance. During Dorr's Rebellion, supporters and opponents of the uprising wrote hundreds of pamphlets, drafted thousands of letters, and delivered numerous speeches about the legality of extralegal action. Despite this broad and lengthy debate, the Americans of 1842, like their eighteenth-century counterparts, did not resolve their dispute. No doubt, many things did change during Dorr's Rebellion -- but one thing did not. Well into the nineteenth-century, Americans were still struggling to resolve an age-old conflict -- the legality of extralegal action -- that was deeply embedded in their founding

---

8. *See* Letter from Samuel Adams to Noah Webster (Apr. 30, 1784), *in* 4 THE WRITINGS OF SAMUEL ADAMS 305 (Harry A. Cushing ed., 1968).

9. *See* Paul M. Thompson, *The Reaction to Shays' Rebellion*, 4 MASS. LEGAL HIST. 37, 39, 56 (1998).

moment.

I. RESISTANCE, REVOLUTION AND THE LEGACY OF EXTRALEGAL ACTION

The moment the American colonies ended their war with Great Britain, the new American nation faced a less tangible, but equally formidable, threat at home. If the colonists could resist British efforts to enforce the Stamp Act in 1765, could not American citizens do the same to laws passed by the Continental Congress in 1785? And if the people had the right "to alter or to abolish" their existing government, and "institute new government" in its place,[10] could not Americans overthrow their new government and construct a better one at any time?[11] For the Americans of the revolutionary generation, these questions were not trivial. The radical Bostonian Samuel Adams agonized over these very issues and feared that the young republic would not survive.[12] Virginian Henry Lee also worried that America would fall prey to the twin "horrors of anarchy and licentiousness,"[13] unless it tamed the threat posed by its revolutionary founding. Adams and Lee, however, were not alone. By the middle of the 1780's, many American leaders feared that their new nation -- a nation fused in revolution and steeped in a tradition of popular resistance to authority -- would itself succumb to revolution at home.

These fears were rooted in tradition and history. Throughout the eighteenth-century, Americans shared an enduring belief in the rights of revolution and resistance to authority.[14] Prior to the Revolution, popular up-

---

10. THE DECLARATION OF INDEPENDENCE para. 1 (U.S. 1776).

11. *See* Pauline Maier, *Freedom, Revolution, and Resistance to Authority, 1776-1976, in* FREEDOM IN AMERICA: A 200 YEAR PERSPECTIVE 25, 26 (Norman A. Graebner ed., 1977) [hereinafter Maier, *Freedom*].

12. *See* William Pencak, *Samuel Adams and Shays' Rebellion*, 62 NEW ENG. Q. 63, 65 (1989). *See also* PAULINE MAIER, THE OLD REVOLUTIONARIES: POLITICAL LIVES IN THE AGE OF SAMUEL ADAMS 30, 31 (1980) [hereinafter MAIER, THE OLD REVOLUTIONARIES].

13. Letter from Henry Lee to George Washington (Sept. 8, 1786), *in* 8 LETTERS OF MEMBERS OF THE CONTINENTAL CONGRESS, 1785-89, at 463 (Edmund C. Burnett ed., 1936).

14. *See* PAULINE MAIER, FROM RESISTANCE TO REVOLUTION: COLONIAL RADICALS AND THE DEVELOPMENT OF AMERICAN OPPOSITION TO BRITAIN, 1775-1776, at 3, 27 (1973) [hereinafter MAIER, RESISTANCE]; *see also* Pauline Maier, *Popular Uprisings and Civil Authority in Eighteenth-Century America*, 27 WM. & MARY Q. 138 (3rd Ser. 1993) [hereinafter Maier, *Popular Uprisings*]. Professor Maier's work on the rights of resistance and revolution is the most comprehensive, and it is relied on heavily throughout this article.

risings and riots were common in America.[15] For instance, in 1765, a crowd of Bostonians took to the streets in opposition to the Stamp Act. In 1773, the Tea Act fueled a similar reaction. When these limited acts of resistance failed to change the tyrannical ways of the British King, Americans exercised what they believed to be their ultimate right: the right of revolution. From resistance to revolution: such was the logic of rebellion.[16] Americans of the revolutionary generation understood this logic and were prepared to exercise their rights of resistance and revolution to retain their freedom. "[W]hat country can preserve its liberties," Thomas Jefferson asked rhetorically, "if their rulers are not warned from time to time that their people preserve the spirit of resistance?"[17]

This "spirit of resistance" was comprised of two distinct, but related rights: a right of revolution and a right of resistance. The right of revolution was grounded in the contractual nature of government.[18] Eighteenth-century Americans believed that they had a contract with their King in which they agreed to obey him as long as he acted for the common good. When the King abused his power, he breached this contract and was stripped of all lawful authority. Power then reverted back to the people who were free to overthrow their government and put a new one in its place.[19] This justification for revolution emerged in England during the seventeenth century,[20] but blossomed within the American colonies. In

---

15. *See* Gordon S. Wood, *A Note on Mobs in the American Revolution*, 23 WM. & MARY Q. (3rd Ser.) 635, 635 (1966). *See also* PAUL A. GILJE, RIOTING IN AMERICA 51-53 (1996); 1 PAMPHLETS OF THE AMERICAN REVOLUTION, 1750-1765, at 581-82 (Bernard Bailyn ed., 1965); GORDON S. WOOD, THE CREATION OF THE AMERICAN REPUBLIC 319-28 (1969) [hereinafter WOOD, CREATION]; John P. Reid, *In a Defensive Rage: The Uses of the Mob, the Justification in Law, and the Coming of the American Revolution*, 49 N.Y.U. L. REV. 1043, 1053 (1974); Arthur Meier Schlesinger, *Political Mobs and the American Revolution, 1765-1776*, 99 PROC. AM. PHIL. SOC'Y 244, 244 (1955).

16. *See* BERNARD BAILYN, THE IDEOLOGICAL ORIGINS OF THE AMERICAN REVOLUTION 94-143 (1967).

17. Letter from Thomas Jefferson to William Smith (Nov. 13, 1787), *in* 12 THE PAPERS OF THOMAS JEFFERSON 356 (Julian P. Boyd ed., 1955).

18. *See* MAIER, RESISTANCE, *supra* note 14, at 27-28. The colonists justified revolution only in extreme circumstances. *See* JOHN LOCKE, SECOND TREATISE ON GOVERNMENT (1690), *reprinted in* TWO TREATISES OF GOVERNMENT 414-15 (Peter Laslett ed., 2d 1970).

19. *See* JULIE MOSTOV, POWER, PROCESS, AND POPULAR SOVEREIGNTY 55-60 (1992). Mostov writes: "When the King [ ]puts himself in a state of war with the community, acting 'beyond right,' he loses the immunity and authority given to him by law. The government dissolves, authority reverts back to the people, and they then are free to establish a new government." *Id.* at 59.

20. *See* MAIER, RESISTANCE, *supra* note 14, at ch. 2; *see also* Pauline Maier, *The Road Not Taken: Nullification, John C. Calhoun, and the Revolutionary Tradition*

1776, Americans described their Revolution as a break in "the original contract between king and people."[21] Therefore, when the colonists asserted their right "to alter or to abolish" their existing government, "and to institute new government,"[22] they were merely putting into practice what they had long before accepted in theory.[23]

The right of resistance was also grounded in contract, but it acknowledged that every breach of the contract need not result in full-scale revolution.[24] Often, the best response to a single unlawful act by the King was some kind of limited opposition to that act alone. For instance, Jefferson noted that resistance to unlawful authority prevented the "degeneracy of government, [ ] nourishe[d] a general attention to the public affairs," and kept rulers attentive to the rights of the people.[25] By exercising their right of resistance, the colonists hoped that they could stomp out tyranny before it took root.[26] The right of resistance, therefore, was a necessary middle ground between complete obedience to oppressive rulers and full-scale revolution.

Although both revolution and resistance seemed to permit direct opposition to the King and his authority, eighteenth-century Americans did not consider either act to be illegal. Rather, Americans considered acts of resistance and even revolution as "extralegal" in both substance and form.[27] First, revolution and resistance were acts against a King who himself was acting without authority.[28] Eighteenth-century Americans knew that when

---

in South Carolina, 82 S.C. HIST. MAG. 1, 2 (1981).

21. WOOD, CREATION, *supra* note 15, at 270.

22. THE DECLARATION OF INDEPENDENCE para. 1 (U.S. 1776).

23. *See* PAULINE MAIER, AMERICAN SCRIPTURE: MAKING THE DECLARATION OF INDEPENDENCE 135-37 (1997).

24. *See* Maier, *Freedom*, *supra* note 11, at 28; MAIER, RESISTANCE, *supra* note 14, at ch. 2.

25. Letter from Thomas Jefferson to James Madison (Jan. 30, 1787), *in* 11 THE PAPERS OF THOMAS JEFFERSON, *supra* note 17, at 93. *See also* Harris G. Mirkin, *Rebellion, Revolution, and the Constitution: Thomas Jefferson's Theory of Civil Disobedience*, 13 AM. STUD. 61 (1972). Mirkin argues that Jefferson saw resistance as beneficial to society because it prevented the need for full-scale revolution. Acts of resistance "forced the society to deal with problems before they assumed proportions that would justify real revolution." *Id.* at 64.

26. Without the right of resistance, John Locke noted in his *Second Treatise on Government*, the people would be forced to obey their King even while he was extinguishing their freedom and making them slaves. Only then -- after they had been made slaves -- could the people "take care of their liberty" and "act like Freeman." Maier, *Freedom*, *supra* note 11, at 28 (quoting John Locke).

27. Throughout this paper, the term "extralegal" action is used to refer to conduct that is consistent with either the rights of resistance or revolution.

28. *See* Reid, *supra* note 15, at 1049-50.

the King abused his powers, he automatically forfeited these powers back to the people. Through acts of resistance and even revolution, the people merely reclaimed authority that was lawfully theirs. Resistance to an unlawful act, the reasoning went, was not, and could never be, illegal. Second, resistance and revolution were extralegal because they took place outside of the procedures established by the existing government for redress of grievances. Rather than seek redress through the courts or the legislature, those who engaged in extralegal action sought redress outside of the formal institutions set up by the government.[29]

Although extralegal action was not considered illegal, the colonists nonetheless sought to limit when and how it should be used. In 1775, Alexander Hamilton suggested that resistance to government was appropriate only "[w]hen the first principles of civil society are violated, and the rights of the whole people are invaded . . . ."[30] Nevertheless, it was often difficult to determine when these "first principles" were violated and when they were not. Accordingly, by the middle of the 1760's, the colonists formed clubs and committees to help them answer this difficult question and to regulate resistance to the British Crown. The Sons of Liberty was one such committee which organized the resistance during the 1760's and 1770's.[31] The Sons of Liberty insisted that colonial resistance be controlled, purposeful, and organized. It stressed to its members that before the people could turn to extralegal action, all lawful and peaceful means of redress must be exhausted.[32] Even when all legal means of redress had failed, resistance was only proper if it had "specific" and "limited" goals, aimed at remedying "immediate grievances which the regular processes of government seemed incapable of satisfying."[33] Active opposition to government was considered lawful and appropriate only if all these requirements were met.

By the middle of the 1770's, organized resistance committees like the Sons of Liberty underwent a dramatic transformation: they no longer merely opposed laws, but they began to make them.[34] By 1775, in fact,

---

29. *See* WOOD, CREATION, *supra* note 15, at 320-21. Wood defines extralegal as "outside the legal representative institutions." He also uses the term "people out of doors." *Id.*

30. ALEXANDER HAMILTON, THE FARMER REFUTED . . . . (1775), *reprinted in* 1 THE PAPERS OF ALEXANDER HAMILTON 81, 136 (Harold C. Syrett & Jacob E. Cooke eds., 1961).

31. *See* MAIER, RESISTANCE, *supra* note 14, at chs. 3-4.

32. *Id.* at 32-38.

33. 1 PAMPHLETS OF THE AMERICAN REVOLUTION, *supra* note 15, at 581-82.

34. *See* DONALD S. LUTZ, POPULAR CONSENT AND POPULAR CONTROL: WHIG POLITICAL THEORY IN THE EARLY STATE CONSTITUTIONS 77-78 (1980). When royal governors dissolved the colonial legislatures, the members of these legislatures estab-

"committees and conventions in nearly all of the colonies were controlling and regulating all aspects of American life . . . ."[35] So pervasive were these extralegal organizations that some Tories actually complained that committees were directing "what we shall eat, drink, wear, speak, and think."[36] This evolution from extralegal committee to legal government occurred throughout the colonies. In the states, *ad hoc* committees and clubs, organized for the purpose of resistance, became the first formal state and local governments. On the national level, the Continental Congress which organized the colonial resistance to the Crown and drafted the *Declaration of Independence*, became the governing body for the new nation. By 1776, extralegal resistance had turned into revolution which, in turn, had led to the creation of America's first formal governments.

Despite the creation of these new American governments, many Americans continued to assemble in committees, clubs, and conventions and openly resist acts of their federal and state legislatures.[37] Nowhere was this "spirit of resistance" more noticeable than in Massachusetts. Throughout the 1780's, Massachusetts citizens set up committees and organized extralegal conventions to protest high taxes, and to urge reform of their state constitution.[38] When these extralegal conventions failed to obtain redress from the General Court in Boston, many in western Massachusetts turned to mob action and closed down the state courts in what was later called Shays' Rebellion. Those who took part in Shays' Rebellion thought that their actions were consistent with their revolutionary tradition. In the 1770's, they argued, when Great Britain tried to act contrary to the best interests of the colonists, the people exercised their rights of resistance and revolution. Now in the 1780's, many in Massachusetts felt that their government was following the same path toward oppression. Popular protests, mobs, and extralegal committees, they reasoned, were the logical and required response.

When extralegal committees and mobs turned against American governments, however, many American leaders started to question the appropriateness of the old "spirit of resistance."[39] To many Americans, the old rights of resistance and revolution clashed with the need for stable and

---

lished legislative committees that met even though the legislature had been formally dissolved. Towns and local governments also set up similar committees to run local affairs. During the years before the Revolution, these committees not only organized the resistance movement in the colonies, but they also ran local and colonial government. *Id.*

35. WOOD, CREATION, *supra* note 15, at 322.

36. *Id.*

37. *Id.* at 323.

38. *See* Thompson, *supra* note 9, at 44-45.

39. *Id.* at 51-53.

strong governments in the new republic. American leaders knew that survival of the young republic depended upon respect for its laws and obedience to its authority. In his *Circular Letter to the States*, for instance, George Washington prayed that God "would incline the hearts of the citizens to cultivate a spirit of subordination and obedience to government."[40] Washington and others knew that uprisings, riots, and resistance to the enforcement of laws passed by American governments, would only diminish the strength and stability of the fragile new nation.[41] Thus, when Americans attempted to exercise their right of resistance in the 1780's--when extralegal committees, conventions, and even riots sprouted throughout the nation[42]--American leaders tried to tame this revolutionary spirit and preserve order and stability throughout the republic.

One way that American leaders sought to tame the popular action of the 1780's was to insist that acts of resistance, like Shays' Rebellion, were unacceptable in America after the Revolution. No one articulated this theory better than Samuel Adams. Adams acknowledged that extralegal committees and mobs had "served an excellent Purpose" when used against the British.[43] Nonetheless, he reasoned that "as we now have constitutional & regular Governments and all our Men in Authority depend upon the annual & free Elections of the People, we are safe without them."[44] To Adams, extralegal action was no longer legitimate because Americans had representative government. In the 1760's-70's, the colonists were not represented in the British Parliament and, therefore, were subject to laws made without their consent. These colonists had to turn to mobs and extralegal committees because it was the only way they could block the enforcement of unconstitutional laws. In 1786, in contrast, Americans did not have to turn to extralegal action because they could end the enforcement of oppressive laws simply by voting for new representa-

---

40. George Washington's Circular Letter to the States (June 8, 1783), *reprinted in* COLONIES TO NATION, 1763-1789: A DOCUMENTARY HISTORY OF THE AMERICAN REVOLUTION 436, 443 (Jack P. Greene ed., 1967) [hereinafter COLONIES TO NATION].

41. *See* Thompson, *supra* note 9, at 53.

42. *See* WOOD, CREATION, *supra* note 15, at 324. On the use of extralegal action after the Revolution, Wood notes that "[c]ommittees and associations of the people, given form and sanction by the experience of the Revolutionary movement, were spilling out everywhere to voice grievances or to realize political goals . . . . Serious rioting under the direction of the radical committees recurred in all major cities . . . ." *Id. See also* GILJE, *supra* note 15, at 51-59.

43. *See* Letter from Adams to Webster, *in* 4 WRITINGS OF ADAMS, *supra* note 8, at 305.

44. *Id.* On Samuel Adams' change in thinking about extralegal activity, see generally Thompson, *supra* note 9, at 51-52; Pencak, *supra* note 12; MAIER, THE OLD REVOLUTIONARIES, *supra* note 12, at 30-31.

tives.[45] For "[e]very complaint, or grievance," imaginable in the 1780's, there was a peaceful means of redress.[46] To those like Samuel Adams, formal representative institutions became "a peaceful surrogate for resistance and revolution."[47]

American leaders also sought to quell the "spirit of resistance" in a very pragmatic way: they drafted and ratified a new Constitution that created a stronger national government with power to suppress extralegal action.[48] Article IV, Section 4 of the new Constitution specifically granted the national government the power to protect the states "against domestic violence."[49] The Constitution also gave the national government the powers to tax[50] and raise armies,[51] both of which were seen as essential to bringing stability and order to the new republic. In fact, soon after the adoption of the Constitution, these new powers were used to crush extralegal activity within the states. In 1794 and again in 1799, federal troops were sent into western Pennsylvania to suppress the uprisings of disgruntled Pennsylvanians during the Whiskey[52] and Fries Rebellions.[53]

Although a theory of opposition to extralegal action was emerging, and a strong national government was in place to support this new theory, a

---

45. *See Circular Letter to the Several Towns in the Commonwealth* (Sept. 13, 1786), *in* MASS. CENTINEL, Sept. 13, 1786. This letter was written by Adams in response to Shays' Rebellion. *See* Pencak, *supra* note 12.

46. Speech of Governor James Bowdoin (Sept. 28, 1786), *in* MASS. CENTINEL, Sept. 30, 1786.

47. Maier, *Freedom*, *supra* note 11, at 33. After the Revolution, some Americans also began to question the lawfulness of extralegal activity because, with the King now gone, the new threat to freedom seemed to come from the people themselves. To James Madison, for instance, the people posed a new threat to liberty: the tyranny of the majority. *See* James Madison, *Vices of the Political System of the United States*, *in* 9 THE PAPERS OF JAMES MADISON, 1786-87, at 353 (Robert A. Rutland & William M.E. Rachal eds., 1975). Visions of individuals openly resisting acts of legitimate American governments made many American leaders, like Madison, fear the end of the new republic. Letter from James Madison to Edmund Pendelton (Feb. 24, 1787), *in* 9 THE PAPERS OF JAMES MADISON, 1786-87, at 295; *see also* Letter from David Humphreys to George Washington (Jan. 20, 1787), *in* 1 LIFE AND TIMES OF DAVID HUMPHREYS: SOLDIER, STATESMAN, POET 396 (Frank Landon Humphreys ed., 1917).

48. *See* Thompson, *supra* note 9, at 57-61. *See generally* WOOD, CREATION, *supra* note 15, at pt. 4.

49. U.S. CONST. art. IV, § 4.

50. *Id.* art. I, § 8, cl. 1.

51. *Id.* at cl. 12.

52. For an overview of the Whiskey Rebellion, see THOMAS P. SLAUGHTER, THE WHISKEY REBELLION: FRONTIER EPILOGUE TO THE AMERICAN REVOLUTION (1986); THE WHISKEY REBELLION: PAST AND PRESENT PERSPECTIVES (Steven R. Boyd ed., 1985).

53. On Fries Rebellion, see WILLIAM W.H. DAVIS, THE FRIES REBELLION, 1798-99 (1969).

few Americans still clung to the old "spirit of resistance." Those Americans who took part in the rebellions in Massachusetts and Pennsylvania, for instance, were familiar with their revolutionary tradition, and invoked it often.[54] The Shaysites wore sprigs of evergreen in their hats, which was the symbol of liberty worn by many in the Continental Army during the Revolution.[55] Those who took part in the Whiskey Rebellion mimicked the Stamp Act rioters by coercing excise tax collectors to resign, rather than collect the tax.[56] A number of Americans who watched these uprisings in Massachusetts and Pennsylvania also embraced the old "spirit of resistance." Thomas Jefferson was the most eloquent of these observers. In response to Shays', for instance, Jefferson wrote that the "spirit of resistance" was "so valuable on certain occasions, that I wish it to be always kept alive."[57] Throughout the 1780's and 90's, the "spirit of resistance" was kept alive, albeit in a limited form. By the end of the eighteenth-century, in fact, new and old attitudes about the legitimacy of extralegal action in America existed side by side and in constant tension.[58]

One point of this article is to demonstrate how this deep-seated tension -- between America's revolutionary tradition and its need for stability and order -- did not simply die out at the end of the eighteenth century. No doubt many Americans like George Washington and Samuel Adams believed that the old "spirit of resistance" had little purpose after the struggle with Britain had ended.[59] Others, however, disagreed.[60] This debate about the legality of extralegal action was not resolved in 1780 or even in 1800. Indeed, it was this same debate that reemerged in Rhode Island nearly fifty years later. Before analyzing exactly how that debate played out, we must first understand what happened in Rhode Island in 1842 and why.

---

54. *See* GILJE, *supra* note 15, at 53-56.

55. *See id.* at 53.

56. *See id.* at 54-55.

57. Letter from Thomas Jefferson to Abigail Adams (Feb. 22, 1787), *in* 11 JEFFERSON PAPERS, *supra* note 17, at 174.

58. *See* Thompson, *supra* note 9, at 39, 56.

59. *See e.g.*, Letter from George Washington to Henry Lee (Oct. 31, 1786), *in* 4 THE PAPERS OF GEORGE WASHINGTON, CONFEDERATION SERIES 319 (W.W. Abbot & Dorothy Twohig eds., 1995) [hereinafter WASHINGTON PAPERS].

60. Thomas Jefferson made numerous statements at the time of Shays' Rebellion that demonstrated his view that resistance and revolution were an important part of America's revolutionary legacy. *See* Thompson, *supra* note 9, at 56.

II. THE RHODE ISLAND QUESTION

By the middle of the 1770's, Rhode Islanders were swept up in the "spirit of resistance." Nearly a year before the colonies declared their independence, Rhode Islanders removed their Royal Governor, Joseph Wanton. On May 4, 1776, Rhode Island broke its allegiance to the crown,[61] and, on July 18, 1776, it became the first state to approve the *Declaration of Independence*.[62] Although Rhode Islanders appeared to be casting off all things English in 1776, they nonetheless held onto their royal Charter, granted by Charles II in 1663. To some Americans, this behavior seemed contradictory. On the one hand, the Rhode Islanders of 1776 were radicals: they were the first Americans to declare their independence from Great Britain. Yet, at the same time, they appeared quite conservative: they clung to their ancient English Charter.[63]

To the Rhode Islanders of 1776, however, there was nothing at all contradictory about overthrowing English rule while clinging to an English Charter. In fact, it was their desire to preserve the liberties guaranteed under that Charter that led them to declare their independence.[64] Under the Charter, Rhode Island was freer than any other colony. For instance, the Charter provided Rhode Island with self-government and created a General Assembly with extraordinary powers. The Charter also provided for "full liberty in religious concernments"[65] which made the colony the most tolerant of different religious faiths out of all the colonies. As nineteenth-century historian George Bancroft noted, "no where in the world were life, liberty and property, safer than in Rhode Island."[66] Even into the late eighteenth-century, most Rhode Islanders viewed their grant from Charles II as the "Munificent Charter."[67]

Rhode Islanders held onto their seventeenth-century English Charter well into the nineteenth-century, even though dramatic changes were sweeping their state. Over the course of the next few decades, Rhode Is-

---

61. *See* 7 RECORDS OF THE COLONY OF RHODE ISLAND AND PROVIDENCE PLANTATIONS IN NEW ENGLAND 522-26 (John R. Bartlett ed., 1968).

62. *See id.* at 581-82.

63. *See* PETER J. COLEMAN, THE TRANSFORMATION OF RHODE ISLAND, 1790-1860, at 218 (1963).

64. *See* CONLEY, *supra* note 1, at 62-63.

65. RHODE ISLAND CHARTER OF 1663, *reprinted in* ARTHUR MAY MOWRY, THE DORR WAR, app. A, at 308 (1901).

66. GEORGE BANCROFT, 2 HISTORY OF THE UNITED STATES FROM THE DISCOVERY OF THE AMERICAN CONTINENT 64 (1834-1875), *quoted in* CONLEY, *supra* note 1, at 27.

67. CONLEY, *supra* note 1, at ch. 2.

land changed from a maritime to an industrial-based economy.[68] Cotton and woolen manufacturing developed first, and by the 1830's, the manufacturing of textile machinery and equipment was also expanding rapidly. This industrialization had a profound effect on the population distribution within the tiny state.[69] In 1790, Newport was the largest town in Rhode Island with 6,716 inhabitants, with Providence close behind at 6,380. While Newport grew to 8,333 in 1840, Providence (due to rapid industrial growth) grew to 23,171 inhabitants--an increase of 295% since 1790.[70] Irish immigrants began trickling into the state's northern industrial towns, like Providence, in the late 1830's. The "foreign" population of Providence increased from 39 in 1820 to 1,005 by 1835.[71] By 1840 Rhode Island had seen its economy, population, and society transform in dramatic ways.

As a result of these increases and shifts in population, many Rhode Islanders began to think of their "Munificent Charter" as less than generous. In particular, the Charter restricted the ability of many to participate in their government. Under the Charter only freeman could vote. Who the "freeman" were was left to the discretion of the General Assembly, which defined the term throughout the seventeenth and eighteenth centuries in terms of property. By 1798, a white male needed $134 of real estate, or rental of such property for $7 a year, to assume the status of freeman and be able to vote.[72] This property qualification was almost liberal by eighteenth-century standards when land was abundant. By 1840, however, the $134 freehold qualification severely limited the number of Rhode Islanders who could vote. Although as many as seventy-five percent of the adult white males in Rhode Island could vote at the end of the eighteenth-century, by 1841 this figure was down to forty percent.[73] In Providence, which contained over twenty percent of the state's population, only six

68. *See* COLEMAN, *supra* note 63, at chs. 2-3.

69. *See* CONLEY, *supra* note 1, at 150-57. Conley classifies the Rhode Island towns into three categories: 1) expanding--those that had a much higher rate of population growth than the state as a whole; 2) static--those whose population had either remained stationery or increased at a rate below the state average; and 3) declining--those that showed a rate of population loss. *See id.* By 1840, the expanding towns in Rhode Island were all in the north and east with Providence expanding at the most rapid rate. The declining and static towns were nearly all in the southern part of the state. *See id.* at 153-54.

70. *Id.* at 155-56.

71. *Id.* at 274-75. By 1850, foreign immigrants would amount to approximately one-fifth of the state's population. *See* COLEMAN, *supra* note 63, at 229.

72. *See* CHILTON WILLIAMSON, AMERICAN SUFFRAGE FROM PROPERTY TO DEMOCRACY, 1760-1860, 178 (1960); *see also* MARVIN E. GETTLEMAN, THE DORR REBELLION: A STUDY IN AMERICAN RADICALISM 1833-1849, 6 (1973).

73. *See* GETTLEMAN, *supra* note 72, at 6-7.

percent were freeman in 1841.[74] These suffrage restrictions were considered so severe that it was "universally conceded that it would have been better if a change in the elective franchise had been made many years since."[75]

Even more significant than the Charter's limitation on suffrage was its fixed system of apportionment. By 1840, the apportionment of representatives in the Rhode Island legislature was "extremely unequal."[76] Under the Charter, Providence received four seats in the General Assembly, while Newport received six. After the Revolution, this apportionment had seemed fair. By 1840, however, Providence had one representative for every 5,793 residents, while Newport had one for every 1,338 of its residents. Many southern farming towns faired even better than Newport. Jamestown, for instance, had one representative for every 182 residents, and Charlestown, one for every 461. The result of this malapportionment was that a small number of towns, comprising less than half of the state's citizens, elected more than two-thirds of its representatives.[77] To many of the Rhode Islanders living in the developing regions of the state, the old Charter government now had all the trappings of oligarchy.[78]

Rhode Island reformers latched onto the issues of suffrage and apportionment and tried to do away with the old Charter and draft a new constitution for the state. But for nearly fifty years, these efforts at reform were unsuccessful. There were rumblings for reform in the late eighteenth and early nineteenth centuries, but not until 1824 did the General Assembly call a constitutional convention.[79] Although the constitution that emerged from this convention would have lessened the inequities in the state, it was voted down in a referendum of the freeholders.[80] Reformers

---

74. *See* COLEMAN, *supra* note 63, at 259 n.65.

75. DR. FRANCIS WAYLAND, THE AFFAIRS OF RHODE ISLAND, A DISCOURSE 14 (Providence, B. Cranston & Co. 1842) [hereinafter WAYLAND, AFFAIRS OF RHODE ISLAND].

76. THOMAS WILSON DORR, INAUGURAL ADDRESS (1842), *reprinted in* BURKE'S REPORT, *supra* note 2, at 722.

77. *See* COLEMAN, *supra* note 63, at 255-58. Even if representation in the legislature was based on property, Coleman suggests that the system was still unfair. *Id.* at 257-58.

78. *See* THOMAS W. DORR ET AL., AN ADDRESS TO THE PEOPLE OF RHODE ISLAND TO PROMOTE THE ESTABLISHMENT OF A STATE CONSTITUTION 23 (Providence, Cranston & Hammond 1834) [hereinafter DORR, 1834 ADDRESS].

79. *See* DORR, INAUGURAL ADDRESS, *reprinted in* BURKE'S REPORT, *supra* note 2, at 722-23; *see also* JACOB FRIEZE, A CONCISE HISTORY OF THE EFFORTS TO OBTAIN AN EXTENSION OF SUFFRAGE IN RHODE ISLAND 12-20 (Thomas S. Hammond 1912) (1842).

80. *See* CONLEY, *supra* note 1, at 202-13.

were disappointed again in 1829 when the General Assembly rejected a number of petitions to extend the suffrage. [81] The reform movement resurfaced again in 1834. This time, the reformers organized into a Constitutional party and focused on securing a new constitution that would expand the suffrage and eliminate the apportionment problem.[82] The result, however, was the same. The General Assembly called a constitutional convention in June 1834. After failing to secure a quorum between the fall of 1834 and the summer of 1835, the convention delegates never reconvened. By the end of 1837, the Constitutional Party had dissolved and efforts at reform in Rhode Island appeared to be at an end.[83]

Often, reformers faced a freehold population that was simply disinterested or apathetic. In the referendum of 1824, for instance, public reaction to the proposed Constitution "oscillated between apathy and mild discussion."[84] Although some reformers in Providence tried to gather support for the proposed Constitution, they had little success. When the final votes were tallied, the Constitution went down by a two to one margin.[85] Frequently, however, reformers faced a freehold population that was actually hostile to the reform agenda. *Hazard's Report*, written in 1829 in response to petitions for extension of the suffrage, was just one example of the hostile reception reformers often received. *Hazard's Report* stressed that a strong and stable government would survive in Rhode Island only if the state retained its freehold qualification.[86] This freehold qualification, Hazard wrote, prevented "spoliation of private property" and protected the

---

81. *See* BENJAMIN HAZARD, REPORT OF THE COMMITTEE ON THE SUBJECT OF AN EXTENSION OF SUFFRAGE 1 (1829), *reprinted in* BURKE'S REPORT, *supra* note 2, at 377-401 [hereinafter HAZARD'S REPORT].

82. *See* DORR, 1834 ADDRESS, *supra* note 78, at 10.

83. *See* FRIEZE, *supra* note 79, at 24-28; COLEMAN, *supra* note 63, at 269-72; CONLEY, *supra* note 1, at chs. 10-11.

84. CONLEY, *supra* note 1, at 210.

85. For the final vote in the referendum of 1824 see *id.* at 211.

86. HAZARD'S REPORT, *reprinted in* BURKE'S REPORT, *supra* note 2, at 382-85. For a complete discussion of *Hazard's Report* see William M. Wiecek, *Popular Sovereignty in the Dorr War--Conservative Counterblast*, in 32 R. I. HIST. 35, 38-39 (1973) [hereinafter Wiecek, *Popular Sovereignty in the Dorr War*]. Hazard noted that anyone other than a property holder belonged to one of three groups: 1) a native Rhode Islander who was without property because of his "own improvidence, extravagance, or vices;" 2) an American citizen from outside the state who, by definition, was not a substantial citizen in his own state or else he would not have moved; or 3) an immigrant who was in Rhode Island as a result of America's liberal immigration policy which Hazard described as a "national misfortune." *Id.* at 38; *see also* HAZARD'S REPORT, *reprinted in* BURKE'S REPORT, *supra* note 2, at 385-93.

state from falling into "anarchy" or "despotism."[87] Representatives from agrarian towns agreed with Hazard, and fiercely opposed any efforts to alter the freehold qualification or the apportionment of the legislature.[88] Whether apathetic or hostile to change, however, all opponents of the reform movement agreed that "those who own the country ought to govern the country."[89]

Although reformers confronted formidable opposition, they tried once again in 1840 to secure a constitution for the state. [90] In March 1840, reform leaders banded together to form the Rhode Island Suffrage Association. This time, they resolved "not to shrink from the task till it shall have been accomplished."[91] In fact, in its founding documents, the Association warned the Rhode Island General Assembly of the events to come. "WE KNOW OUR RIGHTS," the Association's *Preamble* concluded, "AND KNOWING, DARE MAINTAIN THEM."[92]

During the spring and summer of 1841, the Rhode Island Suffrage Association employed extraordinary measures to obtain a new constitution for the state. The Association held public meetings, discussions, and demonstrations in Providence throughout the spring. When the General Assembly succumbed to popular pressure and called a constitutional convention for the fall of 1841, the reformers refused to send delegates. Because the General Assembly did not give the northern towns equal representation in this convention, and refused to permit all white men to vote in the ratification referendum, reformers feared that "the whole affair will result precisely as did the last attempt" at reform.[93] In response, the reformers called their own convention for October 1841, and allowed all adult white males to elect delegates.[94] The delegates to this convention drafted a constitution, dubbed the People's Constitution, which extended the suffrage[95]

---

87. Wiecek, *Popular Sovereignty in the Dorr War*, *supra* note 86, at 38; *see also* HAZARD'S REPORT, *reprinted in* BURKE'S REPORT, *supra* note 2, at 401.

88. *See* CONLEY, *supra* note 1, at 226.

89. PROVIDENCE JOURNAL, Nov. 10, 1841, *quoted in* Wiecek, *Popular Sovereignty in the Dorr War*, *supra* note 86, at 38.

90. By 1840, every state in the union had abandoned its old colonial government and drafted a new written constitution. Every state had also expanded its suffrage to guarantee "universal white manhood suffrage." WILLIAMSON, *supra* note 72, at vii.

91. PREAMBLE AND CONSTITUTION OF THE RHODE ISLAND SUFFRAGE ASSOCIATION 9 (Providence, B.T. Albro 1840) [hereinafter RISA PREAMBLE AND CONSTITUTION].

92. *Id.*

93. THE NEW AGE, Feb. 12, 1841, *quoted in* MOWRY, *supra* note 65, at 60.

94. *See* COLEMAN, *supra* note 63, at 274.

95. The Constitution eliminated the statutory $134 freehold requirement and extended the suffrage to adult white males over the age of twenty-one with one year's

and adjusted the apportionment of the representatives in the General Assembly.[96] In December 1841, this constitution was ratified 14,000 to 52 in an election in which all white males were allowed to participate. In no single previous Rhode Island election had so many votes been cast. Although the People's Constitution was drafted by a convention not legally called by the legislature, nor ratified by citizens entitled to vote under the current Charter, reformers declared it the supreme law of the land.[97]

---

residence in the state. The People's Constitution also reserved to property owners the right to vote on questions of taxation and spending of public monies. *See* PROPOSED CONSTITUTION OF THE STATE OF RHODE ISLAND AND PROVIDENCE PLANTATIONS AS FINALLY ADOPTED BY THE PEOPLE'S CONVENTION ASSEMBLED AT PROVIDENCE ON THE 18TH DAY OF NOVEMBER, 1841, art. II, §1 (Providence, 1841) [hereinafter PEOPLE'S CONST.]. A copy of the People's Constitution is also reprinted in MOWRY, *supra* note 65, app. B, at 322-46.

96. The House under the People's Constitution would contain 80 members. Providence was given twelve representatives and Newport was given five. PEOPLE'S CONST. art. V, §§ 2-3. The Senate had twelve members which were equally divided between expanding and declining towns. *Id.* art. VI, §§ 1-2. The Constitution, however, did not contain any provision for mandatory reapportionment. *See* CONLEY, *supra* note 1, at 312.

97. The reformers defended their use of these extraordinary measures by invoking a doctrine that they called popular sovereignty. No one articulated the theory of popular sovereignty better than Thomas Wilson Dorr. "[T]he actual living majority of the day possess the true sovereignty of the country," Dorr reasoned, "and have a right to investigate, review and amend its political constitution, and to accommodate it to the just demands and necessities of the people." Letter from Thomas W. Dorr to William B. Adams (Mar. 7, 1831), *quoted in* GEORGE M. DENNISON, THE DORR WAR: REPUBLICANISM ON TRIAL, 1831-1861, at 16 (1976) [hereinafter DENNISON, THE DORR WAR]. Although all reformers agreed that the people were sovereign, and that a majority of the people could abolish their existing government and frame a new one at any time, and in any way, they were deeply divided over one significant issue: who were the people? *See generally* THE RIGHT OF THE PEOPLE OF RHODE ISLAND TO FORM A GOVERNMENT: THE NINE LAWYERS' OPINION (1842), *reprinted in* 11 R.I. HIST. TRACTS 63 (Sidney S. Rider ed., Providence, 1880) [hereinafter NINE LAWYERS' OPINION]. Reformers struggled with this question during their convention in the fall of 1841. None of the reformers argued for enfranchisement of women. Many, however, did argue that black men should get the vote under the People's Constitution. Thomas Dorr, for instance, introduced a resolution calling for the enfranchisement of black men. *See* BURKE'S REPORT, *supra* note 2, at 111-13. In this resolution, Dorr maintained that the reformers' invocation of popular sovereignty would appear hollow indeed if blacks were not given the vote. *See id.* at 112; *see also* GETTLEMAN, *supra* note 72, at 46-47; CONLEY, *supra* note 1, at 311-12. Others in the reform movement strongly disagreed. Duttee J. Pearce, for example, claimed that the People's Constitution would never be ratified if blacks were given the vote. *See* GETTLEMAN, *supra* note 72, at 47. Ultimately, popular sovereignty gave in to practical politics and the reformers voted down a measure which would have enfranchised blacks. *See* CONLEY, *supra* note 1, at 311. Primarily to appease Thomas Dorr, a clause was inserted in the Peo-

At the same time that reformers were working out the details of the People's Constitution, a convention called by the Charter government was also debating a new constitution. The constitution which emerged from this convention in February 1842 conceded many of the points raised by the reformers. This Freeholder's Constitution, as it was called, extended the suffrage to native-born white males over the age of twenty-one. It also reapportioned the legislature, but did not go as far as the People's Constitution in shifting the balance of power to the expanding towns in the north. This constitution came up for a vote in March 1842, and it was defeated by a coalition of conservatives and reformers -- those who wanted no constitution and those who wanted the People's Constitution -- 8,013 in favor and 8,689 opposed.[98]

After this defeat, the Charter government resolved to crush any efforts by reformers to establish a working government under the People's Constitution. First, the General Assembly passed a law, that reformers called the "Algerine Law," making it a crime for any citizen to participate in the upcoming People's elections scheduled for April 1842. In addition, the Charter government sent emissaries to President John Tyler to request military support from the United States.[99] Although Tyler did not promise military aid at that time, he clearly stated that if an "actual insurrection" should arise in the state, "it will be my duty . . . to respect the requisitions of that Government which has been recognized as the existing government of the state through all time past . . . ."[100] After Tyler's response was published and made known throughout the state, many Rhode Islanders felt that the President had come down on the side of the Charter government.

With the passage of the Algerine law, and the publication of Tyler's letter, support for the People's Constitution began to wane.[101] Due in part to fear of prosecution under the Algerine law, and in part to the realization that the national government had thrown its weight on the side of the Charter government, many Rhode Islanders defected from the People's movement. Prominent suffragists like William Sprague and Jacob Frieze left the movement. By April 18, 1842, the date of the People's elections, it

---

ple's Constitution which required a referendum on the issue of black suffrage. PEOPLE'S CONST. art. 14, § 22.

98. *See* MOWRY, *supra* note 65, at 124.

99. For the correspondence between the officials of the Charter government and President Tyler, see BURKE'S REPORT, *supra* note 2, at 656-705.

100. Letter from John Tyler to Samuel Ward King (Apr. 11, 1842), *reprinted in* BURKE'S REPORT, *supra* note 2, at 659.

101. *See* CONLEY, *supra* note 1, at 327; *see also* Thomas W. Dorr, Address to the People of Rhode Island (Aug. 1843), *reprinted in* BURKE'S REPORT, *supra* note 2, at 731, 749 [hereinafter DORR, AUG. 1843 ADDRESS].

was evident that the reform movement had lost much of its popular support as well. For instance, Thomas Wilson Dorr ran unopposed as the "People's Governor" and received only 6,539 votes; 7,000 fewer votes than the People's Constitution had received in late December.[102] Inauguration day for the new government, May 3, 1842, also proved that the People's movement had lost support. A fitting symbol of their status, the People's government held its first legislative session at an unfinished foundry building. Although this building was just steps from the empty Providence State House, the reformers rejected Dorr's suggestion that they take control of the State House. Many years later, Thomas Dorr reasoned that "it was [t]here that the cause was defeated, if not lost."[103] Dorr suggested that "by suffering the critical moment to pass by unimproved, [the People's government] lost the accession of all the doubtful and wavering in the community . . . . Having lost the tide, our enterprise was cast among shoals and quicksands . . . . You know the rest."[104]

The rest of what has been called Dorr's Rebellion lasted for another year, but the outcome was never in doubt. Afraid of the Algerine law and the threat of federal intervention, the People's government dispatched Dorr to Washington to lobby President Tyler and members of Congress. In Dorr's absence, the Charter government strengthened its hold in Rhode Island. It issued warrants for Dorr's arrest, and the arrest of most of the leaders of the People's government. After several prominent officials in the People's government were seized, many others, hoping to avert their own prosecution, simply resigned.[105] In fact, just days after the People's government was formed, the *Providence Journal* aptly described the status of Dorr's movement: "Their revolution is in a state of suspended animation. Governor Dorr has hid or run away . . . and their headquarters nobody knows of. Their General Assembly has evaporated."[106] By May of 1842, without a firing of a single shot, the Rebellion seemed to be at an end.

When Dorr returned to the state on May 16, 1842, the tone of the People's movement had changed. While away, Dorr had received assurances of military support from Pennsylvania and New York, if federal troops entered Rhode Island. Dorr and his remaining supporters were now prepared to defend the People's Constitution by force, if necessary. With this new goal in mind, Dorr and a small group of supporters marched on the

---

102. *See* CONLEY, *supra* note 1, at 326-27.

103. DORR, AUG. 1843 ADDRESS, *reprinted in* BURKE'S REPORT, *supra* note 2, at 738.

104. *Id.* at 739.

105. *See* MOWRY, *supra* note 65, at 162-63.

106. *See id.* at 163 (quoting PROVIDENCE JOURNAL, May 9, 1842).

State arsenal on Cranston Street shortly after midnight on May 18th. Dorr assumed that the troops guarding the arsenal would join the People's forces and help prepare for battle against the Charter government. He was wrong. The troops inside refused to surrender. Dorr ordered a cannon fired at the arsenal, but the cannon misfired, and the arsenal was undamaged. Rather than charge the arsenal, Dorr and his supporters simply ran away, and then fled the state to avoid arrest by the Charter government. The Rebellion officially continued until Thomas Dorr returned to Rhode Island in October 1843 and was tried and convicted for treason. Nonetheless, after the events of May 18th, there was little doubt that the Charter government would retain its hold on the government of Rhode Island.[107]

Although the Dorrites lost the battle for control of the state, many of their reforms were enshrined in a new Constitution, adopted by the people of Rhode Island in 1842. The Charter government held a convention in September 1842 as an effort to appease most of the reformers.[108] When the document that emerged from this convention was put to a vote in November, the Dorrites boycotted the voting, and the Constitution was ratified 7,024 to 51.[109] About half of the new Constitution was "identical, even in wording" to the People's Constitution of 1841.[110] The new Constitution extended the suffrage to white males over the age of twenty-one.[111] It also reapportioned the House of Representatives of the General Assembly in a way that was even more equitable than in the People's Constitution. No town, however, could have more than twelve members.[112] Not only did the new Constitution equal the reforms of the People's Constitution, but to some, it eclipsed them. In May 1843, this new Constitution went into effect.

Dorr's Rebellion had an effect on the political landscape outside of Rhode Island as well. In 1844, the United States Congress published *Burke's Report*, a 1,000-page investigation of President Tyler's "interfer-

---

107. *See* DENNISON, THE DORR WAR, *supra* note 97, at 81-82; *see also* CONLEY, *supra* note 1, at 339; MOWRY, *supra* note 65, at 253-54.

108. *See* CONLEY, *supra* note 1, at 345.

109. *See id.* at 351.

110. *Id.*

111. Limitations still existed, however. Foreign-born citizens, for instance, could only vote if they owned $134 worth of real property. Further, only citizens who paid taxes on $134 worth of property could vote in elections for the City Council in Providence or in propositions to impose taxes or spend public money. *See* CONSTITUTION OF THE STATE OF RHODE ISLAND AND PROVIDENCE PLANTATIONS, art. II, §§ 1-2 [hereinafter RHODE ISLAND CONST.]; *see also* COLEMAN, *supra* note 63, at 284-85.

112. *See* RHODE ISLAND CONST. art. V, § 1; *see also* MOWRY, *supra* note 65, at 288-89.

ence" with the affairs in Rhode Island. Written by New Hampshire Congressman Edmund Burke, the report sided with the Dorrites and endorsed the people's right to alter or abolish their government without the consent of their existing government.[113] The report also criticized President John Tyler for his promises of military aid to the state and condemned the Charter government for its opposition to the principle of popular sovereignty. After *Burke's Report* was printed, many Democrats throughout the nation began to call for Thomas Dorr's liberation. In fact, the Democrats ran their 1844 Presidential election under the campaign slogan of "Polk, Dallas, and Dorr."[114] This political pressure was felt back home in Rhode Island. In 1845, Charles Jackson was elected Governor of Rhode Island on a "Liberation" ticket. Soon after the election, Jackson signed a bill to release Dorr from prison. Nine years later, the General Assembly fully restored Dorr's political and civil rights.

Although many issues were settled in the years after Dorr's Rebellion, one question still remained. To contemporaries, it was called the "Rhode Island Question."[115] Could a majority of the people abolish their government and form a new one without the permission of the existing government? At its root, it was a question about the legality of extralegal action. Throughout the 1840s, Americans debated this question at all levels and in all branches of government.[116] Although the Rhode Island Question seemed simple enough, for nearly ten years Americans struggled to answer it.

---

113. *See* BURKE'S REPORT, *supra* note 2, at 83.

114. CONLEY, *supra* note 1, at 358.

115. *See generally* THE RHODE ISLAND QUESTION: ARGUMENTS OF MESSRS. WHIPPLE AND WEBSTER IN THE CASE OF MARTIN LUTHER PLAINTIFF IN ERROR, VERSUS LUTHER M. BORDEN AND OTHERS 1 (Charles Burnett ed., 1848) [hereinafter ARGUMENTS OF WHIPPLE AND WEBSTER]. "The most exciting topic now before the public is the Rhode Island question . . . ." Letter from John O. Bradford to James K. Polk (Oct. 5, 1844), *in* 8 CORRESPONDENCE OF JAMES POLK, 1844, at 154 (Wayne Cutler ed., 1993).

116. A case arising out of the Rebellion, *Luther v. Borden*, reached the United States Supreme Court in 1849. *See generally* Michael A. Conron, *Law, Politics, and Chief Justice Taney: A Reconsideration of the Luther v. Borden Decision*, 11 AM. J. LEGAL HIST. 377 (1967); George M. Dennison, *Thomas Wilson Dorr: Counsel of Record in Luther v. Borden*, 15 ST. LOUIS U. L.J. 398 (1971) [hereinafter Dennison, *Counsel of Record*]; Mahlon H. Hellerich, *The Luther Cases in the Lower Courts*, 11 RHODE ISLAND HIST. 33 (1952).

III. THE RESURRECTION OF EXTRALEGAL ACTION

To Thomas Dorr and his supporters, the answer to the Rhode Island Question was simple: a majority of the people had the right to amend or alter their form of government at any time and in any manner they saw fit. In answering this question, the Dorrites looked back to the eighteenth-century legacy of extralegal action and embraced this legacy in a number of ways. For one, they adopted a theory about the relationship between the people and the government that would have been familiar to those steeped in the legacy of extralegal action. In addition, the Dorrites not only thought like their revolutionary forebearers, they acted like them. They formed extralegal clubs and committees and organized extralegal conventions in order to form their new government. In both theory and practice -- word and deed -- the Dorrites embraced the eighteenth-century legacy of extralegal action and gave meaning to the "legal" in extralegal action.

The Dorrites' theory was one that eighteenth-century Americans would have found oddly familiar. While some of the terms were different, most of the ideas were essentially the same. For instance, the Dorrites talked about popular sovereignty.[117] The Dorrite doctrine of popular sovereignty was comprised of two elements. First, the people were the sole source of power in America. Second, the people could use this power to abolish their current government and create a new one, even when doing so amounted to defying the wishes of the current government. It was within this second element of popular sovereignty that the Dorrites embraced the legacy of extralegal action. "The greater power inherent in the People, by virtue of the sovereignty, to form a Constitution," Dorr noted, "involves the less power: that of proceeding in the *way* and *manner*, which the People deem proper."[118] As Dorr's statement shows, the Dorrites believed that the people could exercise their sovereignty in any way they pleased, even through extralegal action.

There was a striking similarity between the Dorrites' theory of popular sovereignty and the old right of resistance in particular. Eighteenth-century Americans understood that the people played an active role in monitoring their rulers. Extralegal committees, clubs, and mobs, for in-

---

117. *See* Christian G. Fritz, *Alternative Visions of American Constitutionalism: Popular Sovereignty and the Early American Constitutional Debate*, 24 HASTINGS CONST. L.Q. 287, 314 (1997). Fritz argues that one view of popular sovereignty in the 1780's was what he calls the "expansive" view. This expansive view of popular sovereignty would permit the people to amend their Constitution even when doing so amounted to bypassing or defying the amendment provisions of that Constitution. Fritz contends that it was this "expansive" view of popular sovereignty that the Dorrites appealed to during the crisis in Rhode Island. *Id.* at 356-57.

118. NINE LAWYERS' OPINION, *supra* note 97, at 78.

stance, were instruments used by the people to keep their rulers in line with popular sentiment. In fact, when rulers engaged in unlawful acts, eighteenth-century Americans believed that the people should step in and reclaim the reins of government. This is what the old right of resistance to authority was all about. The right of resistance allowed the people to protest individual acts of their rulers without having to engage in full-scale revolution. To the Americans of the 1770's, the right of resistance provided a valuable "middle ground" between complete obedience to oppressive rulers and full-scale revolution.[119]

To the Dorrites, popular sovereignty also provided the people with a "middle ground" between revolution and blind obedience to their government.[120] In fact, the Dorrites envisioned a role for the people that was similar to the role the people played under the eighteenth-century right of resistance. Thomas Dorr analogized the role of the people to an invisible hand, which could, at any time, become visible and reclaim control of government.[121]

> [T]he mighty hand of the people which created the state and its institutions, though now withdrawn and invisible, is still potent with its original energies; and let but a finger of that hand be pointed at the wrongs or corruptions *which resist reform through ordinary and convenient modes*, and bid defiance to the wishes of the majority and the miracle is wrought . . . . The will of the majority prevails, and the practical working of the State is made to conform to the standard of equality and justice. The hand is withdrawn. Its use is suspended. The certainty that it exists, and can be used in the last resort, prevents the necessity of using it at all.[122]

Like the right of resistance, the doctrine of popular sovereignty gave the people a direct role in their government.[123] To Dorr and his supporters, it

---

119. *See* Maier, *Freedom*, *supra* note 11, at 28. For discussion of the right of resistance see *supra* text accompanying notes 24-42.

120. *See* Fritz, *supra* note 117, at 314. Fritz argues that:

> Thomas Dorr and his adherents . . . justified their right to initiate constitutional revision outside the authority of the existing state government as consistent with the principle of popular sovereignty. That principle, they argued, provided a constitutional middle ground between revision authorized by existing governments and constitutions and the ultimate right of revolution based on sheer power.

Id. at 356. *See also* George M. Dennison, *An Empire of Liberty: Congressional Attitudes Toward Popular Sovereignty in the Territories 1787-1867*, 6 MD. HIST. 19, 26 (1975) [hereinafter Dennison, *An Empire of Liberty*].

121. *See* GETTLEMAN, *supra* note 72, at 69-71.

122. DORR, AUG. 1843 ADDRESS, *reprinted in* BURKE'S REPORT, *supra* note 2, at 745 (emphasis added).

123. Thomas Dorr reasoned that after independence from Britain the "right of

was the function of the people to monitor their rulers and investigate abuses.[124] If the government became corrupt, or if reforms were not obtainable "through ordinary and convenient modes," then the people could turn to extralegal action, reclaim the reins of government, and implement the "will of the majority."[125]

Despite the similarities between the Dorrites' theory of popular sovereignty and the old right of resistance, the Dorrites significantly changed the old right of revolution. The Dorrites believed that the American Revolution and the *Declaration of Independence* had transformed this old right of revolution into a new right of *peaceable* revolution.[126]

The difference between this new right and the eighteenth-century right of revolution was striking. Under the old right of revolution, the people could abolish their existing government only in extraordinary circumstances of extreme oppression. As John Locke noted, it was only after "a long train of Abuses, Prevarications, and Artifices, all tending the same way" that the people could "rouze themselves" and overthrow their rulers.[127] The right of peaceable revolution, in contrast, permitted the people to change their government even before full-scale revolution was necessary. In his *Majority Report*, Congressman Edmund Burke argued that it was absurd to read the *Declaration of Independence* as enshrining the old right of revolution. All people in all nations shared that right, Burke suggested.[128] Instead, the drafters of the *Declaration of Independence* were creating a new right just for Americans: the right of the people to change their government "whenever [they] thought it necessary for their security and happiness."[129] This new right of peaceable revolution fit in easily with the Dorrites' view that the people could monitor and change their government at any time.

---

ultimate sovereignty" passed from the King to the people. Thus, the people themselves now possessed the right to run their own government, and redress their own grievances. THOMAS W. DORR, Speech on the Right of the People of Rhode Island to Form a Constitution, Delivered in the People's Convention (Nov. 18, 1841), *reprinted in* BURKE'S REPORT, *supra* note 2, at 851, 855 [hereinafter DORR, NOV. 1841 SPEECH].

124. *See* Letter from Thomas W. Dorr to William B. Adams (Mar. 7, 1831), *quoted in* DENNISON, THE DORR WAR, *supra* note 97, at 16. Gettleman writes that "[the] doctrine of an Invisible Hand in politics implied a muted but continuously operating set of democratic safeguards, which, once the citizens were sufficiently enlightened, would operate almost automatically." GETTLEMAN, *supra* note 72, at 69-70 (citations omitted).

125. DORR, AUG. 1843 ADDRESS, *reprinted in* BURKE'S REPORT, *supra* note 2, at 745.

126. *See* Majority Report, *reprinted in* BURKE'S REPORT, *supra* note 2, at 27.

127. LOCKE, *supra* note 18, at 415.

128. *See* Majority Report, *reprinted in* BURKE'S REPORT, *supra* note 2, at 27.

129. *Id.*

At its root though, the Dorrites' theory of popular sovereignty was similar to the old rights of resistance and revolution in one significant way: it regarded resistance and revolution (whether peaceable or not) as *extralegal*, not illegal. For instance, Thomas Dorr insisted that once the people decided to change their government, their proceedings need not be "legally called; that is, called by the General Assembly . . . ."[130] Instead, the people could abolish their government whether or not their government consented.[131] Dorr also stated that the people were not bound by any amendment provisions found in an existing constitution.[132] The existence of an amendment provision within a constitution, Dorr wrote, cannot "impair the general unalienated [and] inalienable right of the People at large to make alterations in their organic laws in *any other* mode, which they may deem expedient . . . ."[133] Just as legislatures could not limit how the people exercised their sovereignty, "Constitutions and plans of government" were not "barriers against Popular Sovereignty."[134]

To support their view that the people were not acting illegally when they sought to change their existing government without its consent, the Dorrites pointed to the *Declaration of Independence*. In fact, no document was cited by Dorr and his supporters more often than this one. Dorr reasoned that the *Declaration of Independence* not only granted the people the right "to alter or abolish their government whenever they deem it expedient,"[135] but it also empowered the people to exercise that right without the consent of their current government. To support this reading of the *Declaration*, the Dorrites looked to the writings of one of its signers, James Wilson. In his *Lectures on Law*, Wilson noted that the people "*always retain* the right of abolishing, altering or amending their Constitution, *at whatever time*, and *in whatever manner*, they shall deem expedient."[136] Just as no single document was cited more than the *Declaration of Independence*, no single author was quoted more often by the Dorrites

---

130. *Id.*

131. *See* DORR, NOV. 1841 SPEECH, *reprinted in* BURKE'S REPORT, *supra* note 2, at 861.

132. *See* NINE LAWYERS' OPINION, *supra* note 97, at 78.

133. Letter from Thomas W. Dorr to Nathan Clifford (Jan. 24, 1848), *quoted in* Dennison, *Counsel of Record*, *supra* note 116, at 424; *see also* BENJAMIN HALLETT, ARGUMENT IN THE RHODE ISLAND CAUSES BEFORE THE SUPREME COURT OF THE UNITED STATES 10 (Boston, Beals & Greene 1848).

134. Letter from Thomas W. Dorr to Nathan Clifford (Jan. 24, 1848), *quoted in* Dennison, *Counsel of Record*, *supra* note 116, at 424.

135. DORR, INAUGURAL ADDRESS, *reprinted in* BURKE'S REPORT, *supra* note 2, at 727; *see also* Majority Report, *reprinted in* BURKE'S REPORT, *supra* note 2, at 26-27; NINE LAWYERS' OPINION, *supra* note 97, at 78-79.

136. JAMES WILSON, 1 LECTURES ON LAW 17, *quoted in* NINE LAWYERS' OPINION, *supra* note 97, at 84.

than Wilson.[137] Quotations from Wilson's work were often listed, without comment or analysis, in the pamphlets distributed by the Dorrites. Words and phrases such as "at whatever time," or "in whatever manner" which stressed the people's right to change their government through extralegal action, were always italicized.[138] The Dorrites argued that the *Declaration of Independence*, as interpreted by Wilson, proved that when the people changed their government through extralegal action, they were in complete compliance with the law.

If the *Declaration of Independence* was not proof enough that the Dorrites' actions were not illegal, then the actual practice of the founding fathers was.[139] The way in which the Constitution of 1787 was drafted and ratified, in fact, seemed to confirm the idea that it was wholly legal for the people to abolish their existing government and build a new one. Congress had originally called the Constitutional Convention in 1787 merely to revise the *Articles of Confederation*.[140] "Instead of attending to that work, solely, [the delegates] abandoned [the Articles] altogether, and proceeded to frame an entire new system of government."[141] Although they had defied the instructions of Congress and the amendment provisions in the Articles themselves, those who drafted the Constitution of 1787 argued that their actions were proper.[142] The Dorrites saw in the formation and ratification of the Constitution support for their own extralegal proceedings. If an "irregular and unauthorized proceeding" was the means in

---

137. *See, e.g.*, Majority Report, *reprinted in* BURKE'S REPORT, *supra* note 2, at 26-41; NINE LAWYERS' OPINION, *supra* note 97, at 78-91.

138. *See, e.g.*, NINE LAWYERS' OPINION, *supra* note 97, at 83-86. "[T]he people may change their Constitution *whenever* and *however* they please." *Id.* at 85.

139. *See, e.g.*, FACTS INVOLVED IN THE RHODE ISLAND CONTROVERSY WITH SOME VIEWS UPON THE RIGHTS OF BOTH PARTIES 36-37 (Boston, B. B. Mussey 1842); NINE LAWYERS' OPINION, *supra* note 97, at 78.

140. For the amendment procedures in the Articles see ARTICLES OF CONFEDERATION, § XIII (1781). The formation and ratification of the Constitution differed from the amendment procedures set down in the Articles in many respects. First, the Articles called for amendments to be confirmed by the state legislatures. Most states, however, ratified the Constitution in conventions of the people. Second, the Articles required the consent of every state in the union before an amendment could go into effect. The Constitution, however, went into effect after it was ratified by nine states. On the extralegal formation of the Constitution see BRUCE ACKERMAN, WE THE PEOPLE: FOUNDATIONS 173-75 (1991).

141. Majority Report, *reprinted in* BURKE'S REPORT, *supra* note 2, at 39.

142. In *Federalist No. 40*, for instance, Madison argued that "forms ought to give way to substance" because "rigid adherence" to forms would make the people's right to "abolish or alter their government's" a useless right indeed. THE FEDERALIST NO. 40 (James Madison), *reprinted in* THE FEDERALIST 258, 265 (Jacob E. Cooke ed., 1961) (quoting THE DECLARATION OF INDEPENDENCE para. 2 (U.S. 1776)).

which the "present constitution of the union [was] brought into being," then the way in which the Dorrites framed the People's Constitution was certainly lawful as well.[143]

Recent events also confirmed that extralegal action was wholly legal. In 1836, the people of the Michigan territory met in a convention without the consent of their existing territorial government.[144] When this convention petitioned the United States Congress to admit Michigan as a new state, the Congress became embroiled in an intense debate about the extralegal means used by the Michigan convention. Although opponents called the Michigan convention "a tumultuary assemblage," and "a riotous, unlawful mob," the "Congress took a different view."[145] In fact, the Congress expressly endorsed the people's use of extralegal means to frame a Constitution.

> The people of any State may, at any time, meet in convention, *without a law of their legislature, and* WITHOUT ANY PROVISION, OR AGAINST ANY PROVISION IN THEIR CONSTITUTION, and may alter or abolish the whole frame of government as they please. The sovereign power to govern themselves was in the majority, and they could not be divested of it.[146]

In the drafting and ratification of the U.S. Constitution and the formation of the State of Michigan, the Dorrites found ample evidence to support their own effort to abolish the Charter government and frame the People's Constitution. These two examples also proved, to the Dorrites at least, that when the people exercised their sovereignty, extralegal action was not only legal, but it was at times required.

When the Dorrites put their theory of popular sovereignty into practice, they also acted in ways that eighteenth-century Americans would have understood. For instance, those engaged in the suffrage movement knew that extralegal action was only proper after all legal means of redress had failed.[147] Thus, during the 1830's, Dorr insisted that reformers seek redress of their grievances through the "slow process of legislation."[148] An-

---

143. Majority Report, *reprinted in* BURKE'S REPORT, *supra* note 2, at 39; *see also* WILLIAM GOODELL, THE RIGHTS AND THE WRONGS OF RHODE ISLAND 29 (Press of the Oneida Inst., 1842); HENRY WILLIAMS, SPEECH IN VINDICATION OF THE RIGHTS OF THE PEOPLE OF RHODE ISLAND TO AMEND THEIR FORM OF GOVERNMENT 4 (1845).

144. *See* FACTS INVOLVED IN THE RHODE ISLAND CONTROVERSY, *supra* note 139, at 36-37.

145. *Id.* at 36.

146. Majority Report, *reprinted in* BURKE'S REPORT, *supra* note 2, at 39 (quoting Senator Benton during the debates over the admission of Michigan) (emphasis added).

147. *See* Robert Ciaburri, *The Dorr Rebellion in Rhode Island: The Moderate Phase*, 26 RHODE ISLAND HIST. 73, 74 (1967).

148. Letter from Thomas W. Dorr to William B. Adams (May 28, 1832), *reprinted in* DENNISON, THE DORR WAR, *supra* note 97, at 16.

other suffragist, Jacob Frieze, suggested that the people could turn to resistance and revolution only "after the proper application, through all the usual modes of approaching the legislative body," had failed to achieve the desired reforms.[149] By the end of 1840, however, many reformers believed that the time had come for the people to act outside of the legislative process and "[take] the matter into their own hands."[150] Speaking in traditional eighteenth-century fashion, *The New Age* cautioned that "if [the people] cannot obtain redress of their grievances in the ordinary way, they will take extraordinary measures to obtain it."[151]

In their letters and speeches, the Dorrites took pains to prove that all legal means of redress had failed before they moved to extraordinary measures. For example, John Bolles noted that "for more than forty years," the people had been "quietly endeavoring, by petition and appeal to the legislature" to get a constitution of their own making.[152] Likewise, in his *Inaugural Address*, Thomas Dorr described in detail how a generation of Rhode Islanders had asserted their right "to govern themselves" to no avail.[153] He noted the movements for reform in 1811, 1819, 1824, 1829, 1832 and 1834. "A whole generation had thus passed away" without obtaining any reforms at all.[154] "The legislature had been repeatedly approached in terms of respectful petition," but it "had turned a deaf ear to the reiterated and most earnest remonstrances of a long injured and oppressed majority."[155] After recounting this recent history, Dorr then asked, "[i]s there no remedy for these manifest grievances?"[156] To Dorr, the answer was obvious. The people had to move to extralegal action in order to obtain a new constitution for the state.[157]

---

149. FRIEZE, *supra* note 79, at 42.

150. *Id.*

151. THE NEW AGE, Feb. 12, 1841, *quoted in* MOWRY, *supra* note 65, at 60.

152. JOHN A. BOLLES, "THE AFFAIRS OF RHODE ISLAND," BEING A REVIEW OF PRESIDENT WAYLAND'S "DISCOURSE;" A VINDICATION OF THE SOVEREIGNTY OF THE PEOPLE, AND A REFUTATION OF THE DOCTRINES AND DOCTORS OF DESPOTISM 3 (1842).

153. DORR, INAUGURAL ADDRESS, *reprinted in* BURKE'S REPORT, *supra* note 2, at 724.

154. *Id.* at 722-23.

155. *Id.* at 723.

156. *Id.*

157. *See id.* at 724. In July of 1841, the Rhode Island Suffrage Association argued that all lawful means of redress had been exhausted. In its *Address*, RISA noted that the people cannot find

> redress through the ballot box, from which, by law, they are excluded. Nor is it much more likely that they will derive it from legislative aid; the members of the legislature being exclusively the representatives of the minority . . . . The laws of the state are against them; the legislative authority is against

When the Dorrites turned to extralegal means, they did so in a way that would have been familiar to those steeped in the eighteenth-century rights of resistance and revolution. Public meetings were held in which participants discussed possible resistance to the current government.[158] The Dorrites held parades and rallies which were aimed at maintaining current supporters and gaining others.[159] They set up town committees for the purpose of correspondence throughout the state.[160] They established a state-wide committee to organize and plan an extralegal convention to draft a new constitution.[161] Like their eighteenth-century counterparts in many ways, the Dorrites were creating an extralegal structure to coordinate their resistance. This structure would last throughout their convention in the Fall of 1841, and into the Spring of 1842, when the Dorrites attempted to turn their extralegal committees, convention, and Constitution into the legal government of Rhode Island.[162]

In both theory and practice then, the Dorrites embraced the eighteenth-century legacy of extralegal action. Although they updated and altered the eighteenth-century rights of resistance and revolution, they espoused a theory of the relationship between the people and the government that was consistent with eighteenth-century beliefs. Like their eighteenth-century forebearers, the Dorrites viewed extralegal action as entirely legal and at times necessary. The right to extralegal action was enshrined in the *Declaration of Independence*, confirmed by the words and practices of the founding fathers, and validated by recent events in the United States. It was a firm belief in the legality of extralegal action that influenced the Dorrites when they answered the Rhode Island Question. Of course the people could abolish their existing government and frame a new one at

---

> them; the custom of more than half a century is against them; and, no doubt, the opinions, interests, political aspirations, and the prejudices and prepossessions of a majority of the landed class are against them.

ADDRESS OF THE STATE SUFFRAGE COMMITTEE (1841), *reprinted in* BURKE'S REPORT, *supra* note 2, at 261; *see also* GETTLEMAN, *supra* note 72, at 42-43.

158. *See* MOWRY, *supra* note 65, at 61-62.

159. *Cf.* EDMUND S. MORGAN, INVENTING THE PEOPLE: THE RISE OF POPULAR SOVEREIGNTY IN ENGLAND AND AMERICA chs. 8-9 (1988).

160. *See* HOW THE PEOPLE'S CONSTITUTION WAS MADE FOR RHODE ISLAND WITHOUT THE AID OF THE LAW OR OF THE LEGISLATURE 2 (1842).

161. *See id.*

162. *Cf. supra* text accompanying notes 34-38. On colonial committees of correspondence, see generally RICHARD D. BROWN, REVOLUTIONARY POLITICS IN MASSACHUSETTS: THE BOSTON COMMITTEE OF CORRESPONDENCE AND THE TOWNS, 1772-1774 (1970). Many Dorrites actually believed that they were reliving the American War for Independence. For instance, they often referred to the members of the Charter government as "Tories." *See* DORR, NOV. 1841 SPEECH, *reprinted in* BURKE'S REPORT, *supra* note 2, at 852.

any time, they answered. To Thomas Dorr and his supporters, this answer was so obvious that it was almost instinctive. Yet, as simple as the answer may have seemed to Thomas Dorr, the conservative reaction to the Rebellion reveals that this seemingly simple question did not yield an easy answer.

IV. THE REJECTION OF EXTRALEGAL ACTION

Conservatives also believed that the answer to the Rhode Island Question was simple: a majority of the people could not abolish their government and form a new one without the permission of their existing government. In answering this question, conservatives espoused a view of extralegal action that was very different than the one held by the Dorrites. For instance, opponents of Dorr's argued that the people had no right to openly resist their existing government. Rather, the people had to obey their existing laws and institutions. Opponents boasted that nineteenth-century American institutions--such as the legislature and the courts--provided for the redress of all grievances and made extralegal action unnecessary and illegal. If the people wanted to change their government, conservatives reasoned, they had to do one of two things: either work through the procedures for change established by their existing government, or engage in a full-scale violent revolution -- and win. In 1842, the legacy of extralegal action had to give way to the need for strong, stable, and orderly government.[163]

To conservatives, the debate over Dorr's Rebellion turned on the answer to a narrow question: how can the people exercise their sovereignty?[164] The way that they phrased the question revealed just how many points they were willing to concede. First, those opposed to Dorr's acknowledged that the people were sovereign. Rhode Island Congressman Elisha Potter griped that "[t]hese Dorrites met together and resolved . . .

---

163. Conservatives rejected any "middle ground" for the people between revolution and obedience. *See* Fritz, *supra* note 117, at 356. Fritz argues that opponents of Dorr's Rebellion "rejected such a constitutional middle ground. Their position denied that the people had inherent rights of revision, and held that constitutional revision could legitimately take place only as authorized by existing governmental authorities and in compliance with established procedures." *Id.*

164. *See* Wiecek, *Popular Sovereignty in the Dorr War*, *supra* note 86, at 42; *see also* WIECEK, GUARANTEE CLAUSE, *supra* note 2, at 96. Wiecek calls the opponents of Dorr's Rebellion "conservatives." Others have called the opponents of Dorr's "landholders" or "freeholders." All of these terms will be used interchangeably in this section to represent those who spoke out against the rebellion in Rhode Island.

that the people are sovereign, as if anybody had ever denied it." [165] William Goddard also admitted that the people were sovereign. "Never have we denied the right of the people to make and alter their constitutions of government, a right which constitutes the basis of our political systems."[166] Second, conservatives conceded the substance of the debate to the Dorrites: they expanded the suffrage and reapportioned the legislature in the Charter Constitution of 1842.[167] In a letter to President Tyler written in early April, 1842, conservative legislator John Whipple admitted that "free suffrage" would "prevail" in Rhode Island because the "freeholders of Rhode Island have yielded that point."[168]

The one point that conservatives did not concede, however, was that the people could exercise their sovereignty through extralegal action. Those hostile to the suffrage movement insisted that the Dorrites' theory of popular sovereignty, with its endorsement of the old spirit of resistance, was not acceptable in nineteenth-century America. The Rhode Island Charter could only be amended or abolished with the consent of the existing Rhode Island legislature.[169] Conservatives and freeholders claimed that this approach conformed with logic and reason. "The fact that their Constitutions do not provide for their amendment," South Carolina Senator John C. Calhoun wrote, "cannot, by any force of reason, imply the right of the majority, to alter or abolish their Constitutions, without the consent of their respective Governments." [170] This approach to amending or abolishing existing constitutions was also confirmed by the history of

---

165. Congressman Elisha R. Potter, *quoted in* William M. Wiecek, *"A Peculiar Conservatism" and the Dorr Rebellion: Constitutional Clash in Jacksonian America*, 22 AM. J. L. HIST. 237, 248 (1978) [hereinafter Wiecek, *Peculiar Conservatism*].

166. WILLIAM GODDARD, AN ADDRESS TO THE PEOPLE OF RHODE ISLAND, ON THE OCCASION OF THE CHANGE OF CIVIL GOVERNMENT 40 (Providence, Knowles and Vose 1843) [hereinafter GODDARD, ADDRESS]; *see also* ARGUMENTS OF WHIPPLE AND WEBSTER, *supra* note 115, at 38.

167. *See* WIECEK, GUARANTEE CLAUSE, *supra* note 2, at 95-96. Wiecek notes that "[t]he *sine non qua* of the Freeholder conservatives . . . was legality: reforms were conceded, however, reluctantly, but they could be brought about only in a legitimate way." *Id.* at 96.

168. Letter from John Whipple to John Tyler (April 9, 1842), *reprinted in* BURKE'S REPORT, *supra* note 2, at 669.

169. *See* ELISHA R. POTTER, SPEECH ON THE MEMORIAL OF THE DEMOCRATIC MEMBERS OF THE LEGISLATURE OF RHODE ISLAND 8-9 (Washington, 1844).

170. Letter from John C. Calhoun to William Smith (July 3, 1843), *in* 17 THE PAPERS OF JOHN C. CALHOUN, 1843-1844, at 280 (Clyde N. Wilson ed., 1986) [hereinafter CALHOUN PAPERS]. Rhode Islander William Goddard also argued that "where a Constitution provides no mode of amending itself, the people must effect the desired reform, through the agency of the Legislature, the representatives and agents of the people." GODDARD, ADDRESS, *supra* note 166, at 41.

the new nation.[171] The U.S. Constitution was drafted by a convention called by the existing legislature. The Constitutions of all the states had been drafted "by conventions called by the Legislature . . . ."[172] This history seemed so conclusive, in fact, that Daniel Webster asked, "Now what state ever altered its Constitution in any other mode?"[173] With their theory supported by reason and backed by history, the opponents of Dorr's insisted that the Dorrites' use of an extralegal convention to draft an extralegal constitution was simply improper.[174]

Conservatives and freeholders argued that the Dorrites' action was not merely improper, it was illegal. Throughout the crisis in Rhode Island, the Dorrites insisted that the people could rightly turn to extralegal action to frame a new constitution for the state. Although the Dorrites admitted that they acted "without law," they nonetheless claimed that their conduct was legal because they did not act "against law."[175] Opponents of the Rebellion did not accept this distinction. Instead, these critics argued that "without law" and "against law" meant the same thing.[176] Thus, when Dorr and his followers defied the General Assembly and framed the People's Constitu-

---

171. *See* GEORGE TICKNOR CURTIS, THE MERITS OF T.W. DORR AND GEORGE BANCROFT, AS THEY ARE POLITICALLY CONNECTED 16 (Boston, John Eastburn 1844). Curtis argued that "the whole course of American history establishes the doctrine that a constitution can only be made binding on the whole people, when the existing public organs have superintended its adoption." *Id.*

172. ARGUMENTS OF WHIPPLE AND WEBSTER, *supra* note 115, at 40.

173. *Id. See also* GETTLEMAN, *supra* note 72, at 191-92.

174. Although the heart of the conservative attack was that extralegal action was improper and illegal, some opponents of Dorr's still accepted the legitimacy of extralegal action. Jacob Frieze, for instance, wrote that the people could turn to extralegal action when all "legal and peaceable means" of redress had been exhausted. Nonetheless, he argued that the Dorrites had not exhausted all of these means. First, Frieze thought that the Dorrites failed to appeal to the people to change their government through the ballot box. Second, Frieze maintained that the Dorrites never appealed directly to the legislature to redress their grievances, "for no petition, memorial, or remonstrance . . . was ever presented to the General Assembly. . . ." Frieze was convinced that the Dorrites were wrong, however, after the Charter government endorsed many of the Dorrites' reforms in the Constitution of 1842. Frieze felt that the Dorrites rejected another possible peaceful mode of redress when they refused to accept this Constitution. FRIEZE, *supra* note 79, at 42-43.

175. REPORT ON THE SUBJECT OF RHODE ISLAND THE CONSTITUTION, GENERAL ASSEMBLY 7 (1842). In an editorial, *The New Age* responded to accusations that the Dorrites had acted "without law" by admitting that the Dorrites had acted without "positive law," but nonetheless denying that they disregarded all law. *See The New Age*, March 4, 1842, *cited in* DENNISON, THE DORR WAR, *supra* note 97, at 65.

176. Letter from Joseph Story to Daniel Webster (Apr. 26, 1842), *in* 5 THE PAPERS OF DANIEL WEBSTER, CORRESPONDENCE, 202 (Harold D. Moser ed., 1982) [hereinafter DANIEL WEBSTER PAPERS].

tion, they acted "not only *without law* but *against law*"[177] -- not solely extralegally, but also illegally. Even if the suffrage movement was "against no law in particular," Rhode Island Chief Justice Job Durfee explained, "it [was], nevertheless, a movement against all law . . . ."[178] The sole purpose of the Dorrites' movement was to overthrow the Charter government and "terminate the existence of the State," Durfee claimed.[179] Therefore, no matter what term one attached to it, such a movement was treason and could only be illegal.

It was the judges in Rhode Island, like Durfee, who were the first to clearly argue that the Dorrites' action was illegal. In March 1842, three Justices of the Rhode Island Supreme Court--Durfee, Levi Hallie, and William Staples--issued a public letter in the *Providence Journal* condemning the Dorrites and declaring the People's Constitution void.[180] These Justices wrote that "the convention which formed the 'People's Constitution' assembled without law; that in forming it they proceeded without law; that the votes, given in favor of it, were given without law, and . . . that said constitution . . . is of no binding force whatever . . . ."[181] In this letter, these Justices warned Dorr that any attempt to erect a government "will be treason against this state . . . ."[182] Chief Justice Durfee reinforced this strong message in a series of speeches that he delivered throughout the state during the Rebellion.[183] Even Federal District Court Judge John Pitman engaged in the public debate about the legality of the Dorrites' action. In a January 1842 pamphlet entitled *To the Members of the General Assembly of Rhode Island*, Pitman also charged the Dorrites with acting "without law and against law."[184] What all these judges agreed

---

177. REPORT ON THE SUBJECT OF RHODE ISLAND THE CONSTITUTION, GENERAL ASSEMBLY, *supra* note 175, at 7-8.

178. CHARGE OF THE HONORABLE JUSTICE DURFEE, DELIVERED TO THE GRAND JURY AT THE MARCH TERM OF THE SUPREME JUDICIAL COURT, AT BRISTOL, RHODE ISLAND (1842), *reprinted in* BURKE'S REPORT, *supra* note 2, at 707 [hereinafter DURFEE, GRAND JURY CHARGE].

179. *Id.*

180. *See* PROVIDENCE JOURNAL, Mar. 3, 1842, *quoted in* MOWRY, *supra* note 65, at 128-29.

181. *Id.*

182. *Id.*

183. *See* John Schuchman, *The Political Background of the Political-Question Doctrine: The Judges and the Dorr War*, 16 AM. J. LEGAL HIST. 111, 116 (1972).

184. JOHN PITMAN, TO THE MEMBERS OF THE GENERAL ASSEMBLY OF RHODE ISLAND 10, 23 (1842) [hereinafter PITMAN, GENERAL ASSEMBLY]. Even United States Supreme Court Justice Joseph Story made his opinion known that the Dorrites were acting "without law and against law" in his private correspondence. Letter from Joseph Story to Daniel Webster (Apr. 26, 1842), *in* 5 DANIEL WEBSTER PAPERS, *supra* note 176, at 202.

on--and most of Dorr's opponents accepted with ease--was the idea that the Dorrites, in drafting and ratifying the People's Constitution, had broken the law.

That judges played a major role in articulating a theory of opposition to Dorr's Rebellion is fitting, because the conservative attack on extralegal action was rooted in a desire for order, stability, and respect for the law.[185] No one expressed this need for stability better than Brown University President Dr. Francis Wayland. Although Wayland believed that reforms were needed in Rhode Island, he opposed the way in which the Dorrites sought these changes. "If an established government may be overturned on the principles which have been advocated," Wayland explained, "no constitution in the land is worth the parchment it is written on." [186] When taken to its logical conclusion, Wayland reasoned, Dorr's theory of popular sovereignty would permit the people to ignore all laws passed by their legislature and defy the amendment provisions of their constitutions. In short, constitutions and laws would bind no one.[187] Chaos would consume Rhode Island, "spread over all New England,"[188] and "all government [would come] to an end."[189] To prevent this turn to anarchy, Wayland insisted that the people "live under a government of law."[190] Under this government of law, the will of the people "must be expressed, not by mass meetings and party resolutions, but according to the forms prescribed by the constitution."[191] Echoing the arguments of John Calhoun and Daniel Webster, Wayland believed that obedience to existing laws and institutions was the only way to preserve republican government in Rhode Island.

Many other conservatives agreed with Wayland's view that extralegal action was a direct threat to order, stability, and a government under law. In fact, fears of anarchy and licentiousness filled the pages of pamphlets produced by the opponents of Dorr's.[192] "Once [you] admit the right of

---

185. *See* Wiecek, *Peculiar Conservatism*, *supra* note 165, at 249; Wiecek, *Popular Sovereignty in the Dorr War*, *supra* note 86, at 42. Elisha Potter noted that respect for existing laws was the "cardinal principle, especially in a republican government . . . ." POTTER, *supra* note 169, at 9.

186. WAYLAND, AFFAIRS OF RHODE ISLAND, *supra* note 75, at 7.

187. *See id.*

188. *Id.* at 8.

189. DR. FRANCIS WAYLAND, A DISCOURSE ON THE DAY OF PUBLIC THANKSGIVING, JULY 21, 1842, at 16-17 (Providence, H. H. Brown 1842) [hereinafter WAYLAND, THANKSGIVING DISCOURSE].

190. WAYLAND, AFFAIRS OF RHODE ISLAND, *supra* note 75, at 11.

191. *Id.*

192. *See, e.g.*, JOHN PITMAN, REPLY TO THE LETTER OF THE HONORABLE MARCUS MORTON, LATE GOVERNOR OF MASSACHUSETTS ON THE RHODE ISLAND QUESTION 21

one portion of the people to act directly upon the rights of another," John Whipple stated, "without the intervention of judges and legislators, bound by their oaths to proceed according to fixed rules, and you abolish all law and government!"[193] Indiana Congressman Caleb Smith agreed with Whipple, and argued that Dorr's theory was "but a substitution of anarchy for government -- of lawless violence and faction, for law and order."[194] "Law and Order," in fact, was the name that conservatives adopted for their coalition, which opposed the Rebellion and ruled Rhode Island during the crisis. This name revealed the conservative fixation on the need for stability and obedience to existing laws.[195] All members of the "Law and Order" party agreed that there was only one way to prevent anarchy, chaos, and "Dorrism" from taking over in Rhode Island: "Loyalty," as Francis Bowen put it.[196] "To cherish this feeling is our best safeguard against anarchy and licentiousness . . . ."[197]

Loyalty, in fact, was the cornerstone of the conservative theory about the relationship between the people and their government. There were sound reasons, conservatives stressed, why Rhode Islanders should be loyal to their existing laws and institutions. The primary reason was that under a republican government, all grievances could be redressed legally through the legislature. This argument, in fact, was the same one that Samuel Adams had made when he opposed Shays' Rebellion.[198] The opponents of Dorr's argued that representation was the genius of the American system of government. Representation made loyalty to government

---

(Providence, Knowles and Vose 1842) [hereinafter PITMAN, REPLY TO MORTON]; REPORT ON THE SUBJECT OF RHODE ISLAND THE CONSTITUTION, GENERAL ASSEMBLY, *supra* note 175, at 11; FRANCIS BOWEN, THE RECENT CONTEST IN RHODE ISLAND 69 (Boston, Otis, Broaders & Co. 1844) [hereinafter BOWEN, RECENT CONTEST]. The conservative argument on the need for order and stability is similar to the argument made by the opponents of extralegal action in the late eighteenth century. For a discussion of the argument of eighteenth-century leaders see Thompson, *supra* note 9, at 49-50.

193. JOHN WHIPPLE, ADDRESS TO THE PEOPLE OF RHODE ISLAND, ON THE APPROACHING ELECTION 9 (Providence, Knowles and Vose 1843).

194. CALEB C. SMITH, SPEECH ON THE MEMORIAL OF THE DEMOCRATIC MEMBERS OF THE LEGISLATURE OF RHODE ISLAND 8 (Washington, 1844); *see also* Resolutions Prepared To Be Offered Had The Rhode Island Controversy Been Brought Up For Discussion (May 1842), *in* 16 CALHOUN PAPERS, *supra* note 170, at 269.

195. *See* Wiecek, *Peculiar Conservatism*, *supra* note 165, at 253.

196. BOWEN, RECENT CONTEST, *supra* note 192, at 69.

197. *Id.* Chief Justice Durfee used a slightly different word than loyalty. "Allegiance," he argued, was the first principle of government. DURFEE, GRAND JURY CHARGE, *reprinted in* BURKE'S REPORT, *supra* note 2, at 706; *see also* Daniel Webster, Notes for Oral Argument (Jan. 1848), *in* 3 THE PAPERS OF DANIEL WEBSTER, LEGAL PAPERS 751 (Andrew J. King, ed., 1989).

198. *See supra* text accompanying notes 43-47.

essential and extralegal action unnecessary. "[T]he whole system of American institutions," Daniel Webster maintained, "do not contemplate a case in which resort will be necessary to proceedings, *aliunde*, or outside of the law and the Constitution, for the purpose of amending the frame of government."[199] Instead, Webster submitted, the entire American system assumed that the state legislature could resolve all grievances against the state. "[T]hese popular governments in each State . . .," Webster suggested, will work "to bring about changes which the people may judge necessary in their constitutions."[200] If, for some reason, the legislature did not respond to the people's requests, the people could vote their representatives out of office.[201] Thus, for "[e]very complaint, or grievance," there was a peaceful means of redress,[202] and representation was "a peaceful surrogate" for extralegal action.[203]

Representation, therefore, limited how the people could exercise their sovereignty.[204] In contrast to the Dorrites, conservatives did not believe that there was any "middle ground" between obedience and revolution in which the people could monitor, investigate, or change their government. Once the people set up their government, conservatives contended, they played no role in it whatsoever, other than to express their will through their legislature. "The spirit of our institutions," Henry Dorr wrote to his brother Thomas, "requires that the people . . . do everything by deputy -- that they should not assume to themselves the administration of affairs."[205] Henry Dorr reasoned that because representation was the cure for all the possible diseases found in the body politic, there was "very little for the

---

199. Charles O. Lerche, *The Dorr Rebellion and the Federal Constitution*, 9 RHODE ISLAND HIST. 1, 8 n.22 (1950) (quoting Daniel Webster).

200. *Id.*

201. *See* MINORITY REPORT ON RHODE ISLAND MEMORIAL, 28TH CONG., H.R. REP. NO. 581, at 22 (1844) [hereinafter MINORITY REPORT]. The *Minority Report* was written by Congressman John M. S. Causin.

202. Speech of Governor James Bowdoin (Sept. 28, 1786), *in* MASS. CENTINEL, Sept. 30, 1786.

203. Maier, *Freedom*, *supra* note 11, at 33.

204. *See* Wiecek, *Peculiar Conservatism*, *supra* note 165, at 250. *See also* MINORITY REPORT, *supra* note 201, at 22. The *Minority Report* quoted Thomas Paine for the idea that sovereignty could only be exercised through the ballot box. "This sovereignty is exercised in electing and deputing a certain number of persons to represent and act for the whole, and who, if they do not act right, may be displaced by the same power that placed them there, and others elected or deputied in their stead . . . . Therefore, the republican form and principle leaves no room for insurrection, because it provides and establishes a rightful means in its stead." MINORITY REPORT, *supra* note 201, at 21.

205. Letter from Henry C. Dorr to Thomas W. Dorr (Jan. 1842), *quoted in* DENNISON, THE DORR WAR, *supra* note 97, at 56.

people to do, in their original capacity. Having elected their officers, their work is done."[206] John Whipple agreed. To Whipple, the people – "as a people [possessed] no power whatever except to appoint their rulers[.]"[207] Therefore, Whipple hypothesized that "if nine-tenths of the people should meet and repeal a law, it would continue still to be a law."[208] To opponents of Dorr's Rebellion, popular sovereignty meant only one thing: the right of the people to choose their representatives. "Every man's part in the exercise of sovereign power," Webster concluded, was "the right to choose representatives,"[209] and nothing more.[210]

---

206. *Id.*

207. WHIPPLE, *supra* note 193, at 8. *See also* Wiecek, *Popular Sovereignty in the Dorr War*, *supra* note 86, at 44.

208. WHIPPLE, *supra* note 193, at 8.

209. ARGUMENTS OF WHIPPLE AND WEBSTER, *supra* note 115, at 40.

210. To those Rhode Islanders who were not allowed to vote, however, this conservative reliance on representation seemed almost un-American. To many, in fact, the conservative argument appeared remarkably similar to the eighteenth-century British concept of virtual representation. *See* Wiecek, *Popular Sovereignty in the Dorr War*, *supra* note 86, at 44-45. During the Stamp Act crisis, the British articulated the notion of virtual representation to respond to American claims that the colonists were not represented in Parliament. The British admitted that the colonists were not "actually" represented in Parliament, but neither were "Nine-Tenths of the People of Britain." THOMAS WHATELY, THE REGULATION LATELY MADE CONCERNING THE COLONIES AND THE TAXES IMPOSED UPON THEM CONSIDERED (1765), *reprinted in* COLONIES TO NATION, *supra* note 40, at 46, 49. "None are actually, all are virtually represented," Thomas Whately explained, "for every Member of Parliament sits in the House, not as a Representative of his own Constituents, but as one of that august Assembly by which all the Commons of *Great Britain* are represented." *Id.*; *but see* DANIEL DULANY, CONSIDERATIONS ON THE PROPRIETY OF IMPOSING TAXES IN THE BRITISH COLONIES (1765), *reprinted in id.* at 51, 55 (for the colonial response to virtual representation). Because all English people were a "unitary homogenous order with a fundamental common interest," WOOD, CREATION, *supra* note 15, at 174, the interests of representatives were so "interwoven" with the disenfranchised "that if they act for themselves, they must act for the common interest of the whole . . . ." Response of the Worcester Committee of Correspondence, Oct. 8, 1778, *in* POPULAR SOURCES OF POLITICAL AUTHORITY: DOCUMENTS ON THE MASSACHUSETTS CONSTITUTION OF 1780, at 369, 372 (Oscar & Mary Handlin eds., 1966). The British argued that even those who were without the vote were fairly, adequately, and virtually represented in Parliament.

Some of the opponents of Dorr's Rebellion actually adopted this idea of virtual representation. *See* Wiecek, *Popular Sovereignty in the Dorr War*, *supra* note 86, at 45. Congressman John Causin, for instance, argued that even those citizens who could not vote were represented in their legislature. Echoing the sentiments of Thomas Whately, Causin noted that "[a] majority of those legally entitled to vote, by necessity of society, represent the whole." MINORITY REPORT, *supra* note 201, at 29; *see also* Wiecek, *Popular Sovereignty in the Dorr War*, *supra* note 86, at 45. Other conservatives, however, sought to avoid the analogy to the British theory. One way that they side-stepped the issue was to attack the Dorrites' own theory of popular sovereignty.

Even if the legislature was unresponsive to the demands of the people, conservatives argued that the courts would provide redress. Thus, American government provided a second institution that the people could turn to, to seek their reforms. The mere presence of the courts, in fact, made loyalty to government essential, and extralegal action unnecessary. When a person became a citizen of Rhode Island, Francis Wayland claimed, he bound himself "first to the judicial tribunal of the State . . . and ultimately to the highest judicial tribunal of the United States."[211] The courts protected the people from encroachments on their rights by the legislature, and provided a legal avenue to redress any grievances that the legislature might fail to remedy. Resistance to government, therefore, was "scarcely intelligible," Congressman Causin stated, when courts were available to redress grievances and control legislatures.[212]

To some conservatives, this reliance on the courts was linked with an understanding of judicial review.[213] When Samuel Adams spoke out against the "spirit of resistance" in the late eighteenth century, he did not discuss the courts. During the late eighteenth and early nineteenth centuries, however, most of the states had passed written constitutions. Constitutional questions, therefore, started to be adjudicated in the courts. "Judges, then, replaced the people in deciding when laws were unconstitutional and of no force, [and] formalized procedures replaced the traditional actions of the people 'out of doors.'"[214] Congressman John Causin most clearly recognized this new role for the courts. To Causin, the people no longer had to engage in extralegal action to keep their rulers in line, because "the power of rulers is not only limited by elective supervision, but the enactments of their will are subject to the restraint of judicial scrutiny and control."[215] Whereas the people once played a role in monitoring, investigating, and reviewing the actions of their rulers, the courts now

---

The People's Constitution, conservatives insisted, excluded blacks and women from the suffrage. With such notable exceptions imbedded in their own constitution, conservatives reasoned, the Dorrites had no grounds to complain that the current government of Rhode Island excluded nonfreeholders. *See* SMITH, *supra* note 194, at 7; *see also* Wiecek, *Popular Sovereignty in the Dorr War*, *supra* note 86, at 42. Whether the voting line was drawn between whites and blacks, men and women, or freeholders and nonfreeholders, was merely a question of degree. Under either the People's Constitution or the Charter Constitution, conservatives explained, those who could vote had to exercise their votes on behalf of all other members of society. For those who could not vote, therefore, popular sovereignty meant nothing more than obedience to their legislature. *See generally* DENNISON, THE DORR WAR, *supra* note 97, at 164.

211. WAYLAND, THANKSGIVING DISCOURSE, *supra* note 189, at 16-17.

212. *See* MINORITY REPORT, *supra* note 201, at 22.

213. *See* DENNISON, THE DORR WAR, *supra* note 97, at 141.

214. Maier, *Freedom*, *supra* note 11, at 34.

215. MINORITY REPORT, *supra* note 201, at 22.

filled that function. Thus, judicial review became a peaceful and formalized surrogate for extralegal action.[216]

Although most conservatives believed that Dorrism threatened to undermine respect for American institutions, such as the legislature and the courts, some were also concerned about the effect that it would have on a more "peculiar institution": slavery.[217] By 1842, many southern politicians feared that the Dorrite theory of popular sovereignty would encourage slaves to stand up to their masters and revolt.[218] "You can readily comprehend and feel what would be the effects and consequences of Dorrism here at the South, if Dorrism were predominant,"[219] Henry Clay told a crowd in North Carolina. "Any unprincipled adventurer" could merely "collect around him a mosaic majority black and white, aliens and citizens . . . male and female" and "overturn existing governments and set up new ones at his pleasure or caprice."[220] Even northerners recognized the threat to the institution of slavery posed by Dorrism. William Goodell, a supporter of the Dorrites, reasoned that if Rhode Islanders could "form a constitution without leave of their masters, the disenfranchised majority of S[outh] Carolina might do the same, and the 'peculiar institution' would be overthrown."[221] Goodell then explained what the conservative solution to this problem would be--a solution with which most southern politicians agreed. "The northern laborer must therefore be put down, lest the south-

---

216. The conservative argument about the courts was also subject to criticism for reviving the eighteenth-century doctrine of virtual representation. Up until the adoption of the new Rhode Island Constitution in 1843, all judges in Rhode Island were subject to annual appointment by the legislature. Further, until 1843, the legislature had the power to hear and determine petitions praying relief from court decisions. Even into the middle of the nineteenth century, the Rhode Island legislature had appellate powers that were similar to those of the British House of Lords. *See* CONLEY, *supra* note 1, at 43. The judicial system in Rhode Island excluded nonfreeholders in two other ways. First, only freeholders could serve on juries. Second, a nonfreeholder could not bring suit to collect a debt, unless he had the assistance of a freeholder. *See* MOWRY, *supra* note 65, at 80; *see also* BOLLES, *supra* note 152, at 16. For those Rhode Islanders who were without the vote, appealing to the courts to redress grievances was not much different than petitioning their legislature.

217. GOODELL, *supra* note 143, at 47.

218. *See, e.g.*, Elisha Potter, *Memorandum May 6, 1842*, *quoted in* Wiecek, *Popular Sovereignty in the Dorr War*, *supra* note 86, at 47. In this memorandum, Potter recounts a meeting he had with President Tyler concerning the rebellion in Rhode Island. "I mentioned once we leave the laws what is there . . . to prevent negroes revolutionising south? He agreed." *Id.*

219. Henry Clay, Speech in Raleigh North Carolina (Apr. 13, 1844), *in* 10 THE PAPERS OF HENRY CLAY 24 (Melba P. Hay et al. eds., 1991).

220. *Id.*

221. GOODELL, *supra* note 143, at 51.

ern laborer should rise."[222]

Legislatures, courts, and to some, even slavery, were American institutions that demanded respect, loyalty, and obedience, and made extralegal action unnecessary, illegal, and potentially dangerous.[223] To conservatives, however, there was one other reason that the people could not engage in extralegal action: constitutions, by their nature, limited the power of the people. Both the Dorrites and their opponents agreed that a constitution harnessed the powers of government "so that those who exercise its powers, shall not oppress or do violence to the rest of the community."[224] Yet, conservatives also believed that constitutions limited and restrained the sovereignty of the people. As John Calhoun argued:

> [I]t would be the death-blow of constitutional democracy, to admit the right of the numerical majority, to alter or abolish constitutions at pleasure, regardless of the consent of the Government, or the forms prescribed for their amendment. It would be to admit, that it had the right to set aside, at pleasure, that which was intended to restrain it, and which would make it just no restraint at all.[225]

Thus, when a government functioned under a constitution--and to conservatives the Charter of 1663 was a constitution--the people could only change that constitution in two ways. First, they could comply with the amendment procedures in the constitution. Second, if there were no amendment provisions, as in the Charter, the people needed the consent of the existing government.[226] Through a constitution, Daniel Webster rea-

---

222. *Id.*

223. On the growth of American institutions see generally DENNISON, THE DORR WAR, *supra* note 97.

224. Letter from John C. Calhoun to William Smith (July 3, 1843), *in* 17 CALHOUN PAPERS, *supra* note 170, at 284.

225. *Id.* Calhoun's opposition to Dorr's Rebellion was consistent with his theory of nullification. First, the nullification theory was based on the fear of majoritarian tyranny. Calhoun was concerned mainly with protecting minority rights under a popular government. "Nullification was therefore for [Calhoun] a constitutional means of forcing majorities to respect the minority interests represented by the various states . . . ." Maier, *The Road Not Taken*, *supra* note 20, at 16. To Calhoun, nothing could have been more threatening to minority rights than the Dorrite theory that a majority of the people could alter or abolish their government at will. Second, Calhoun's theory of nullification required that "the people act only through regular, constitutional procedures and within established institutions." *Id.* at 10. To nullify a federal law, the state had to act as the state, and through a legally called convention. *Id.* at 11. Unlike the nullifiers of 1832, however, the Dorrites disregarded the institutions of the state, ignored the legally called constitutional convention, and acted extralegally. Calhoun's opposition to Dorr's Rebellion, therefore, is understandable.

226. *See* GODDARD, ADDRESS, *supra* note 166, at 40-41.

soned, the people agreed to "set bounds to their own power."[227]

If the people went beyond the bounds of their own power--if they tried to change their government extralegally--that was seen as nothing less than revolution. Other than compliance with existing laws and institutions, "[n]o other mode of changing [constitutions] can be rescued from the reproach of being revolutionary in its character," William Goddard noted.[228] The Dorrites never claimed to be exercising the old right of revolution. Instead, they argued that they were exercising the new right of peaceable revolution. Nonetheless, many conservatives analyzed the Dorrites' movement in terms of that old right and found it wanting. First, revolution was only justifiable when the people were subject to absolute tyranny.[229] Although some conservatives agreed that changes were needed in the state, they did not think that life had become oppressive enough to justify revolution. In fact, after the Charter government drafted the March 1842 Constitution which extended the suffrage and reapportioned the legislature, conservatives argued that there was no way that the government was tyrannical. Francis Bowen noted that the right of revolution "is not to be brought out of its shrine on any mean or ordinary occasion. It must not be used as a cloak for ambitious usurpation, for reckless love of change, or for treacherous revolt."[230] Only on extremely rare occasions was revolution considered proper. To most conservatives, Dorr's Rebellion was not one of those occasions.

Second, even if life in Rhode Island had become so oppressive that revolution was necessary, the right of revolution meant nothing at all, unless the rebels succeeded - and success depended on power. "Revolution is not to be regulated by law," Judge Pitman noted, "[t]his is an appeal to force against the law and the government."[231] If the Dorrites were indeed engaging in revolution, conservatives argued, then they needed to back up their attempt to overthrow their government with force. In his majority opinion in *Luther v. Borden*, U.S. Supreme Court Chief Justice Taney acknowledged that the people "may alter and change their form of government at their own pleasure."[232] Nonetheless, he stated that "the right of revolution," means "a change by force," and "is nowhere sanctioned . . . in the Constitution of the United States . . . ."[233] Thus, conservatives ac-

---

227. ARGUMENTS OF WHIPPLE AND WEBSTER, *supra* note 115, at 40.

228. GODDARD, ADDRESS, *supra* note 166, at 40.

229. *See* PITMAN, GENERAL ASSEMBLY, *supra* note 184, at 15; *see also* SMITH, *supra* note 194, at 7.

230. BOWEN, RECENT CONTEST, *supra* note 192, at 68.

231. PITMAN, REPLY TO MORTON, *supra* note 192, at 21.

232. *Luther v. Borden*, 48 U.S. (7 How.) 1, 47 (1849).

233. *Id.* at 25. Even Justice Woodbury in dissent noted that revolutions, such as the Revolution of 1776, succeed only "by a union of physical with moral strength . . .

knowledged that the old right of revolution was an acceptable theory. Yet, in practice, if the revolution did not succeed, it was nothing more than a rebellion or insurrection, and "it [was] the duty of the government to put it down."[234]

Rebellion was exactly how most conservatives described the Dorrites' attempt at extralegal action. It is not surprising, therefore, that they compared Dorr's Rebellion to an insurrection that occurred in Massachusetts more than fifty years earlier.[235] Both Judge John Pitman and Francis Bowen, two of the most articulate opponents of the rebellion in Rhode Island, saw parallels between Dorr's Rebellion and Shays' Rebellion.[236] In drawing these analogies, Pitman and Bowen implicitly rejected the idea that extralegal action could be legal. Like the western Massachusetts farmers who took part in Shays', the Dorrites "held unlicensed conventions and 'voted their own constitutionality.'"[237] The participants in both rebellions "assumed the name of the people, demanded a revision of the constitution, arrayed themselves against the legislature, and demanded redress of grievances with arms in their hands."[238] The Dorrites, much like the Shaysites, espoused "lax notions concerning public and private faith, and erroneous opinions which confound liberty with an exemption from legal control . . . ."[239] Noting these similarities between the two insurrections, many conservatives easily concluded that Dorr's Rebellion, like Shays', was improper.

Some conservatives even argued that the Dorrites were less justified in appealing to extralegal action than the Shaysites had been. Dorr's Rebellion, they argued, was "the most flagitious" of all the rebellions in American history because "it was unprovoked by any practical grievances[.]"[240] Unlike Shays' Rebellion, Dorr's was not "incited by the weight of heavy taxation," it did not grow "out of commercial distress," and it was not "nurtured by the destitution and misery of the disaffected classes."[241]

." *Id.* at 55.

234. PITMAN, REPLY TO MORTON, *supra* note 192, at 21.

235. For the reaction to Shays' Rebellion see Thompson, *supra* note 9, at 48-57.

236. *See* PITMAN, GENERAL ASSEMBLY, *supra* note 184, at 4-7; BOWEN, RECENT CONTEST, *supra* note 192, at 65-81.

237. BOWEN, RECENT CONTEST, *supra* note 192, at 65.

238. *Id.* Pitman quoted at length from John Marshall's attack on Shays' Rebellion in Marshall's *Life of Washington.* Marshall had argued that the Shaysites possessed a "disorderly spirit" which was "cherished by *unlicensed conventions* . . . and *assemblies in the name of the people.* . . ." PITMAN, GENERAL ASSEMBLY, *supra* note 184, at 5 (quoting John Marshall).

239. PITMAN, GENERAL ASSEMBLY, *supra* note 184, at 4 (quoting John Marshall).

240. BOWEN, RECENT CONTEST, *supra* note 192, at 65.

241. *Id.*

Rather, Dorr's Rebellion was simply "a causeless revolt."[242] In fact, the change sought by the Dorrites, "was offered to them in the incipient stages of their revolt, *and was refused.*"[243] The Dorrites had been offered a new constitution which extended the suffrage and reapportioned the legislature in March 1842, and the Dorrites had rejected this constitution. Thus, conservatives concluded that if Shays' Rebellion was condemned by the founding fathers when no redress was forthcoming from the legislature, then Dorr's Rebellion should be condemned as well.[244]

Although the comparison to Shays' convinced most conservatives that extralegal action was no longer proper, some still insisted on finding tighter analogies. To do so, they looked beyond America's borders and compared Dorr's to the French Revolution. Many conservatives suggested that when the Dorrites' theory of popular sovereignty was taken to its logical limit, the result would be the same chaos and anarchy that consumed France after its Revolution. All Americans knew about the horrors of the French Revolution. Chief Justice Durfee played on the general American understanding of the events in France when he instructed the Bristol Grand Jury that the Dorrites wanted to relive in Rhode Island the "bloody drama" of the French Revolution.[245] Everyone knew that after its Revolution, France "groaned under the tyranny of the mob[.]"[246] Anarchy was substituted for government, and violence for law and order.[247] The French drafted, ratified, and then abrogated constitutions "as if they were the playthings of the hour[.]"[248] They formed "illegal assemblies" almost

---

242. *Id.*

243. *Id.* at 67.

244. Both the opponents of Dorr's and Shays' spoke of their fears of anarchy and licentiousness. Both also believed that representation made extralegal action no longer proper in America. On Shays' see Thompson, *supra* note 9, at 48-57. The opponents of Dorr's also drew on the writings of Washington and others who opposed Shays' Rebellion. Conservatives, like Judge Pitman, quoted Washington extensively and saw a parallel between Washington's response to Shays' Rebellion in 1786 and the conservative response to Dorr's Rebellion in 1842. Many conservatives suggested that what Washington said of Shays' was just as true of Dorr's Rebellion. "It is but the other day that we were shedding our blood to obtain the Constitutions under which we now live -- Constitutions of our own choice and framing--and now we are unsheathing the Sword to overturn them!" Letter from George Washington to David Humphreys (Dec. 26, 1786), *in* 4 WASHINGTON PAPERS, *supra* note 59, at 478 (W.W. Abbott & Dorothy Twohig eds., 1995). For an example of the use of this quotation see PITMAN, GENERAL ASSEMBLY, *supra* note 184, at 5.

245. *See* DURFEE, GRAND JURY CHARGE, *reprinted in* BURKE'S REPORT, *supra* note 2, at 708.

246. BOWEN, RECENT CONTEST, *supra* note 192, at 64.

247. *See* SMITH, *supra* note 194, at 9.

248. BOWEN, RECENT CONTEST, *supra* note 192, at 64.

daily which "usurped" the powers of their existing legislatures.[249] All of these horrors, Durfee and others warned, would come to pass in America, unless Dorrism--with its endorsement of extralegal action--was defeated.

Not every conservative compared Dorr's Rebellion to the French Revolution. Nonetheless, all conservatives expressed fears of anarchy and chaos, and insisted that the Dorrites respect existing laws and institutions. The cumulative effect of the conservative argument, in fact, was the development of a coherent theory about the relationship between the people and their government and the continuing vitality of extralegal action. John Whipple succinctly expressed this conservative theory when he said that the people possessed "no power whatever except to appoint their rulers."[250] To conservatives, the people could only exercise their sovereignty through existing laws and institutions, making the old eighteenth-century rights obsolete and even illegal. Thus, if the people wanted to change their government, they could do so, as Elisha Potter noted, in only "two ways--by a revolution . . . or by the forms of law."[251] This narrow conception of the people's role in their government was not only a rejection of the Dorrites' theories of popular sovereignty and peaceable revolution, it was also a rejection of the use of extralegal action. To conservatives, there was no longer a "middle ground" between revolution and obedience to government; only the extremes of revolution and obedience survived. Thus, extralegal committees, "combinations and [a]ssociations," which were so valuable at the time of the American Revolution, were of a "fatal tendency," and no longer acceptable in nineteenth-century America.[252]

---

249. *Id.*

250. WHIPPLE, *supra* note 193, at 8. *See also* Wiecek, *Popular Sovereignty in the Dorr War*, *supra* note 86, at 44.

251. POTTER, *supra* note 169, at 8-9.

252. Conservatives also quoted Washington's *Farewell Address*.

> The basis of our political systems is the right of the people to make and to alter their Constitutions of Government. But the Constitution which at any time exists, 'till changed by an explicit and authentic act of the whole People, is sacredly obligatory upon all. The very idea of the power and the right of the People to establish Government presupposes the duty of every Individual to obey the established Government.
>
> All obstructions to the execution of the Laws, all combinations and Associations, under whatever plausible character, with the real design to direct, controul, counteract, or awe the regular deliberation and action of the Constituted authorities are distructive of this fundamental principle and of fatal tendency.

George Washington, *Farewell Address*, *reprinted in* GEORGE WASHINGTON WRITINGS

V. THE SIGNIFICANCE OF DORR'S REBELLION

The purpose of this article was to identify a tension that arose soon after the American Revolution and persisted throughout the middle of the nineteenth-century. As the debate over Dorr's Rebellion demonstrates, throughout the nineteenth-century, old and new attitudes about the legality of extralegal action existed side by side and in constant tension. I recognize that my subject runs the risk of evoking the dreaded "so what?" response. Before Dorr's Rebellion, Americans were divided about the legality of extralegal action, and after the Rebellion, they continued to be at odds about the same issue. So what? If there was really so little change in American attitudes, what is the significance of Dorr's Rebellion? And why, exactly, is it worthy of study?

For one, Dorr's Rebellion sheds light on how nineteenth-century Americans viewed the founding moments of the American Republic. To many Americans who lived at the time of Dorr's Rebellion, the American Revolution meant something far different than it means to us today.[253] The American Revolution can be viewed as preserving two distinct rights: 1) the right to free republican governments in America; and 2) the right to invoke whatever means necessary, even extralegal action, to preserve the first right.[254] To those who fought in the Revolution, these two rights were intimately related. Soon after the end of the war with Britain, however,

---

962, 968 (John Rhodehamel ed., 1997).

253. Viewing the past through twenty-first century eyes often causes us to place significance on certain events or practices and ignore others as unimportant based primarily on our view of how things have come to pass. Recent scholarship has sought to undercover some of the "truths" about the founding moment by focusing on how ideas, practices, and events were understood by those who experienced them at the time. *See, e.g.*, Larry D. Kramer, *Madison's Audience*, 112 HARV. L. REV. 611 (1999); Christine A. Desan, *The Constitutional Commitment to Legislative Adjudication in the Early American Tradition*, 111 HARV. L. REV. 1382 (1998).

254. *See* MAIER, AMERICAN SCRIPTURE, *supra* note 23, at 135-37, 209-15. The *Declaration of Independence* succinctly stated these two rights as follows:

> all Men are created equal, that they are endowed by their Creator with certain unalienable Rights, that among these rights are Life, Liberty, and the Pursuit of Happiness--That to secure these Rights, Governments are instituted among Men, deriving their just Powers from the Consent of the Governed, *that whenever any Form of Government becomes destructive of these Ends, it is the Right of the People to alter or abolish it, and to institute new Government*. . . .

THE DECLARATION OF INDEPENDENCE para. 1 (U.S. 1776) (emphasis added); *see* MAIER, AMERICAN SCRIPTURE, *supra* note 23, at 135-37, 209-15.

Americans had difficulty reconciling the tension inherent in these two, very different, rights. And today, few people would recognize the second right as more than just an eighteenth-century anomaly.[255] Accordingly, most scholars have disregarded the tradition of extralegal action, viewing it as having little significance to the understanding of the legal or political origins of the American Republic.[256] At a minimum, Dorr's Rebellion proves that, even in the middle of the nineteenth century, many Americans still viewed extralegal action as one of the legacies of their founding moment. Standing alone, this is a meaningful historical observation.

Second, Dorr's is significant because it reveals exactly what happened to the legacy of extralegal action and why that legacy has fallen into relative obscurity today. Strong institutions, such as legislatures and the courts, were beginning to develop in America by the middle of the nineteenth century.[257] These institutions had more authority and a wider influence over the lives of Americans than any institutions that existed in America at the time of the Revolution. By the time of Dorr's Rebellion, Americans could file a lawsuit, rather than take to the streets to protest an unconstitutional law.[258] In addition, many nineteenth-century Americans could vote.[259] Americans in the 1840's, unlike those in the 1770's and 1780's, had institutions through which they could channel their political and legal activity. By 1842, many Americans were turning away from riots, mob action, and other forms of extralegal action to seek legal change, and were looking to courts, legislatures, and other institutions to resolve the problems posed by everyday life.

In 1842, Rhode Island was an exception to this trend. In fact, the Rebellion occurred because Rhode Island's formal legal institutions failed to effectively redress the grievances of many Rhode Islanders. In 1840, the Rhode Island courts lacked independence and were still beholden to the General Assembly. At the same time, the suffrage in the state was severely limited, leaving few Rhode Islanders with any means to seek change in their government. Without access to these institutions, Rhode Islanders had little choice but to turn to extralegal action. The Dorrites acted through extralegal means to establish a government with formal institutions that would be more responsive and more accessible to the people of Rhode Island. What is ironic is that the Dorrites won their battle for re-

---

255. *See* MAIER, AMERICAN SCRIPTURE, *supra* note 23, at 135-37, 209-15.

256. One notable exception is the work of Professor Pauline Maier. *See infra* notes 11, 14, 20, 23.

257. On the growth of American institutions see generally DENNISON, THE DORR WAR, *supra* note 97.

258. *See* Maier, *Freedom*, *supra* note 11, at 34.

259. *See* WILLIAMSON, *supra* note 72, at vii.

form, but did so at the expense of their revolutionary ideology. The result of Dorr's Rebellion was to put institutions in place that would make resort to extralegal action increasingly obsolete.

This brings us to another reason that Dorr's Rebellion is a notable event, worthy of study. The Rebellion says much about the process of institutional change in America. The story of Dorr's Rebellion is really a story about how people bring about change in their legal institutions. Throughout American history, there are numerous examples of Americans turning to extralegal action to effect change. Shays' Rebellion in 1786-87 is a notable example. The Shaysites lost their battle, but won the war for reform of their state government.[260] Dorr's Rebellion is merely another example. In both cases, individuals who found their formal legal institutions unresponsive and inaccessible turned to the only form of legal action that they had available to them: extralegal activity. And, in both cases, the formal legal institutions of the state changed and evolved to meet the needs of the protesters and satisfy the calls for reform. Understanding extralegal action, therefore, helps us to recognize the limits of our formal legal institutions and to see precisely how these institutions evolve.

Which brings us to the question posed in the title of this article: is there anything "legal" about extralegal action? To some people, at some time, the answer is certainly yes. To those who engaged in the mob activity of the 1760's and 1770's, extralegal action was as legal as a bill passed by the House of Commons or an edict from the House of Lords. To those western Massachusetts farmers who closed down the courts in 1786 and 1787, extralegal action was as legal as the taxes that had been imposed by the Massachusetts legislature. And to the Dorrites, who were shut out of the formal legal institutions of their state, extralegal action was as legal as any act of the formal government that purported to represent them. In today's highly structured world of formal legal institutions, few people would recognize the riots, mob actions, or extralegal conventions of the eighteenth and nineteenth centuries as "legal." Yet, when our formal legal institutions fail to adequately meet our needs, extralegal action will resurface and, like an invisible hand, will make our institutions more responsive to those they were meant to serve. At that time, the question will reemerge: is there anything "legal" about extralegal action? And the debate will continue.

260. *See* Thompson, *supra* note 9, at 47-48.

COMMENTS

John v. Baker & Child Custody: A Battle Between Parents and Courts

INTRODUCTION

> America, separated from Europe by a wide ocean, was inhabited by a distinct people, divided into separate nations, independent of each other and of the rest of the world, having institutions of their own, and governing themselves by their own laws. It is difficult to comprehend the proposition, that the inhabitants of either quarter of the globe could have rightful original claims of dominion over the inhabitants of the other, or over the lands they occupied; or that the discovery of either by the other should give the discoverer rights in the country discovered, which annulled the pre-existing rights of its ancient possessors.[1]

Native Americans, including Native Alaskans, have been subject to the will of a conquering people since the "discovery" of America.[2] Russian explorers in the 1700s began the decimation of the human population of Indians and later the United States government took over this role of slow destruction of the Indian culture and government, while claiming to be the protector of Native Americans.[3] Until the decision reached by the Supreme Court of Alaska in *John v. Baker*,[4] the annihilation of the inherent rights of Native American tribes seemed imminent.[5]

---

1. Worcester v. Georgia, 31 U.S. (6 Pet.) 515, 542-43, 8 L.Ed. 83 (1832) (Marshall, C.J.).

2. *See* FRANK POMMERSHEIM, BRAID OF FEATHERS 37 (1995). Pommersheim notes that the reason for a "ruthless pursuit of advantage" by the United States government was the colonial fur trade which rendered Indian tribes dependant on the Europeans for foreign goods. *Id.*

3. *See* Ben Summit, *The Alaska Native Claims Settlement Act (ANCSA): Friend or Foe in the Struggle to Recover Alaska Native Heritage*, 14 T.M. COOLEY L. REV. 607, 609 (1997). The Russians hunted fur-bearing sea mammals to near extinction and enslaved the Aleut people. *See id.*

4. 982 P.2d 738 (Alaska 1999), *cert. denied*, 528 U.S. 1182 (2000).

5. *See* Benjamin W. Thompson, *The De Facto Termination of Alaska Native Sovereignty: An Anomaly in an Era of Self-Determination*, 24 AM. INDIAN L. REV. 421, 422 (2000) (discussing the deleterious effects of the Alaska Native Claims Settlement Act resulting in what is, in fact, a de facto end to Native Alas-

The recognition of jurisdiction for child custody determinations by the *Baker* court is a step in the right direction for the Native Alaskan and Native American legal system.[6] The Native Alaskan and Native American tribes were allowed broad discretion by the United States federal government in the early days of modern American jurisprudence,[7] but the courts and Congress slowly eroded the tribes' powers in favor of the national and state authorities.[8] A return to Native American tribal sovereignty as advocated by the *Baker* court and the executive and legislative branches is desperately needed. Native American lands have been enveloped by states over the years, and there has been a slow progression toward complete assimilation into society.[9] The people of these tribes own a special place in American history as they were forced to cede their territory under the will of a conquering people. As a consequence, Native Americans are not viewed as the powerful, independent, sovereign nations that were so eloquently described by Justice Marshall in the Cherokee cases of the 1830s.[10]

This Comment argues that in *John v. Baker*,[11] the Supreme Court of Alaska recognized the right of a tribe to adjudicate child custody cases as an inherent sovereign power while still allowing concurrent jurisdiction with the state, thereby striking a balance between tribal and state government. This was not an extension of jurisdiction; it was a return to the original idea of sovereign power as was evidenced by the decisions in the 1800's. This decision does not close the doors of the American courts on native people.

Part I explores the history of tribal relations with the federal and state governments and shows the varying policies that the government has used to weaken the power of Indian tribes in America.[12] The relevant federal

---

kan tribal sovereignty).

6. Tribal courts in their modern form have existed since the Indian Reorganization Act in 1934, which allowed Indian tribes to write their own constitutions and charter tribal courts. *See infra* Part I.C. These courts are entirely separate from the state and federal courts and may exercise concurrent jurisdiction. *See* Cynthia Ford, *Integrating Indian Law Into a Traditional Civil Procedure Course*, 46 SYRACUSE L. REV. 1243, 1251 (1996).

7. *See infra* Part I.A for further discussion of the view of the Supreme Court towards Native Americans and the idea of inherent sovereignty.

8. *See infra* Part I.B-G for the enactment of statutes and the development of case law utilized to strip Indian tribes of power and increase the power of the state courts.

9. *See infra* Part I.B.

10. Cherokee Nation v. Georgia, 30 U.S. (5 Pet.) 1, 15 (1831); *Worcester*, 31 U.S. (6 Pet.) at 515.

11. 982 P.2d 738 (Alaska 1999), *cert. denied*, 528 U.S. 1182 (2000).

12. *See infra* notes 26-119 and accompanying text.

statutes and case law that affect *John v. Baker* are also discussed.[13] Part II describes *John v. Baker*, including the facts, procedural history and appellate arguments offered by the parties to the Alaskan Supreme Court.[14] Part III discusses the court's holding and its reasoning in regard to Public Law 280[15] and the Indian Child Welfare Act.[16] The Supreme Court of Alaska explored these two statutes, finding that they did not apply, and thus the Alaskan Supreme Court's previous cases based on an interpretation of them were not relevant.[17] Part IV is an analysis of the bases for deciding that an expansion on the current view of jurisdiction is permissible and even needed to restore the original sovereignty of Native Alaskans and American Indians in the eyes of the government.[18] First, the tribal court retains the inherent power to decide domestic relation cases.[19] This power is based on Native Villages as federally recognized tribes; tribal law governs unless Congress divests them of power, domestic relations as a sovereign power of the tribes, as well as congressional intent that evidences support of inherent sovereignty in child custody cases.[20] Second, the nonexistence of "Indian country" does not affect jurisdiction in this case.[21] Third, the tribal court has power over a non-member Indian and finally, both the executive and legislative branches advocated this extension.[22] Part V details the policy arguments that support tribal authority to decide domestic relations disputes among Indian members.[23] Part VI is a review of the holding of the case and subsequent rulings.[24] Finally, Part VII concludes that the balance between tribal and state governments used by the Alaska Supreme Court is a step in the right direction for Native Alaskan tribes.[25]

---

13. *See id.*

14. *See infra* notes 120-47 and accompanying text.

15. Act of Aug. 15, 1953, ch. 505, 67 Stat. 588 (codified as amended at 18 U.S.C. § 1162, 25 U.S.C. §§ 1321-1326, 28 U.S.C. § 1360 (1994)).

16. Indian Child Welfare Act of 1978, Pub. L. No. 95-608, 92 Stat. 3069 (codified in pertinent part at 25 U.S.C. §§ 1901-1963 (1994)).

17. *See infra* notes 148-99 and accompanying text.

18. *See infra* notes 200-349 and accompanying text.

19. *See infra* notes 205-99 and accompanying text.

20. *See id.*

21. *See infra* notes 300-16 and accompanying text.

22. *See infra* notes 317-49 and accompanying text.

23. *See infra* notes 350-71 and accompanying text.

24. *See infra* notes 372-74 and accompanying text.

25. *See infra* notes 375-77 and accompanying text.

I. A SHORT HISTORY OF FEDERAL, STATE AND TRIBAL LAW RELATIONS AS IT RELATES TO NATIVE ALASKANS

The situation in Alaskan tribal sovereignty is a unique one. With the passage of certain federal laws, particularly Public Law 280 (PL 280)[26] and Alaska Native Claims Settlement Act (ANCSA)[27] in the twentieth century, the future of tribal sovereignty seemed dim. However, *John v. Baker*,[28] a recent holding from the Supreme Court of Alaska, reversed the current trend toward societal assimilation and the move toward state court rule over tribal members. *Baker* re-instituted the power of tribal sovereignty by granting the tribal court wider jurisdiction than ever before.[29] In order to fully understand the implications of this far-reaching case, one must understand Native Alaska's unique jurisdictional history, especially as applied to domestic issues such as child custody.

A. Early Federal View of Indian Tribes: The Cherokee Cases

The legal status of Indian tribes in America has been changing since the time of the Cherokee cases in the early 1830s.[30] Chief Justice Marshall of

---

26. Pub. L. 280, 28 U.S.C. § 1360(a) (1988). The statute states:

> Each of the States listed in the following table shall have jurisdiction over civil causes of action between Indians or to which Indians are parties which arise in the areas of Indian country listed opposite the name of the State to the same extent that such State has jurisdiction over other causes of action, and those civil laws of such State that are of general application to private persons or private property shall have the same force and effect within such Indian country as they have elsewhere within the State.

Id.

27. Pub. L. No. 92-203, 85 Stat. 668 (1971) (codified as amended at 43 U.S.C. § 1601-1628 (1988)). In 1971, Congress bought almost forty million acres from Native Alaskans and compensated them with one billion dollars as a means of settling all claims to aboriginal lands in Alaska. *See* John F. Walsh, *Settling the Alaska Native Claims Settlement Act*, 38 STAN. L. REV. 227, 227 (1985). The Act's effect, however, was to cause some to believe that Alaska Natives' sovereignty was extinguished. *See* Joseph D. Matal, *A Revisionist History of Indian Country*, 14 ALASKA L. REV. 283, 339 (1997); *see also infra* Part I.G.

28. 982 P.2d at 738.

29. *See id.* at 765 (holding, as a matter of first impression, Native Alaskan villages have the inherent right to resolve domestic disputes between members).

30. *See* Cherokee Nation v. Georgia, 30 U.S. (5 Pet.) 1, 15 (1831); *Worcester*, 31 U.S. (6 Pet.) at 515. These Supreme Court cases established that Indian tribes in America were states with which the federal government had a special relationship and which had a right to self-government. *See Cherokee Nation*, 30 U.S.

the United States Supreme Court spoke of these tribes in a sympathetic tone in *Cherokee Nation v. Georgia*.[31] Marshall said that the Cherokee Nation was not a foreign state as defined in the United States Constitution,[32] and thus the judicial power granted over cases involving foreign governments did not extend to Indian tribes.[33] Instead, the Cherokee Nation was a "distinct political society, separated from others, capable of managing its own affairs and governing itself."[34] According to the Court, the United States recognized the tribe as a state and other states are bound to recognize them as the same.[35] The Court stated that because the tribes resided within the state, they were probably more appropriately called "domestic dependant nations" rather than foreign governments.[36] The Cherokees had an "unquestioned right to the lands they occupy, until that right shall be extinguished by a voluntary cessation to our government."[37] This statement holds import for the connection between the land and sovereign tribal power to be analyzed later.[38]

The following year, the Supreme Court again had an opportunity to address the issue of state jurisdiction over Indian tribes in *Worcester v. Georgia*.[39] In this case, Chief Justice Marshall declared a Georgia law void as it violated the tribes' sovereignty, the United States Constitution and the laws and treaties of the United States by infringing upon the United States' relationship with the Cherokees.[40] This relationship, according to "settled

---

(5 Pet.) at 16-17, 19; *Worcester*, 31 U.S. (6 Pet.) at 559.

31. *See Cherokee Nation*, 30 U.S. (5 Pet.) at 15. The Court spoke of the Cherokee Nation as

> A people once numerous, powerful, and truly independent, found by our ancestors in the quiet and uncontrolled possession of ample domain, gradually sinking beneath our superior policy, which contains a solemn guarantee of the residue, until they retain no more of their formerly extensive territory than is deemed necessary to their comfortable subsistence.

Id.

32. U.S. CONST. art. III, § 2, cl. 1.

33. *See Cherokee Nation*, 30 U.S. (5 Pet.) at 20.

34. *Id.* at 16.

35. *See id.*

36. *Id.* at 17. As tribes are "domestic dependant nation[s]," the federal government holds title to the land that the tribes occupy and may take possession of it when the tribe's right to possess the land terminates. *See id.* The Supreme Court compares the relationship between the government and the tribe as that of a guardian to a ward. *See id.*

37. *Id.*

38. *See infra* Part IV.B and accompanying text.

39. 31 U.S. (6 Pet.) 515 (1832) (Marshall, C.J.).

40. *See id.* at 516. This Georgia law made the act of living with the Cherokee Indians by a non-Indian punishable if the governor did not give his permission.

principles of our constitution," belonged exclusively to the federal government, which recognized the tribe as a sovereign state.[41] Chief Justice Marshall stated that the Indian nations should be treated as separate political communities with authority within their territorial boundaries, and that the United States federal government guaranteed this authority.[42] The Court described the Cherokee Nation as:

> [A] distinct community, occupying its own territory, with boundaries accurately described, in which the laws of Georgia can have no force, and which the citizens of Georgia have no right to enter, but with the assent of the Cherokees themselves, or in conformity with treaties, and with acts of Congress.[43]

Hence, the Court articulated its view that Indian tribes in the United States derive their power from their authority over the land that they possess as well as over the members of their tribe.[44] In this manner, the tribes were able to exercise power over tribal members, Indians of other tribes and non-member Indians as long as the dispute arose on Indian country.[45] Tribal affairs within Indian country were not subject to state laws.[46] The federal government still kept legislative authority over the tribes, but during this time they seldom employed it.[47]

Cherokee Nation was also important because it was the first Supreme Court case to recognize the trust doctrine.[48] The Court characterized Indian

---

See id. at 534. The Act was enacted to "prevent white persons from residing within that part of the chartered limits of Georgia, occupied by the Cherokee Indians." *Id.* at 515.

41. *Id.* at 561.

42. *See id.* at 557.

43. *Worcester*, 31 U.S. (6 Pet.) at 561.

44. *See id.* Interestingly, the Native Americans themselves did not believe that land was "owned" by any one person. *See* Anthony Peirson Xavier Bothwell, *We Live on Their Land: Implications of Long-Ago Takings of Native American Indian Property*, 6 ANN. SURV. INT'L & COMP. L. 175, 183-84 (2000). Land ownership was not what was important to the Native Americans; rather, the land was a source of sustenance for them. *See* Summit, *supra* note 3, at 608. Appreciation of the land and a close relationship to it has been passed down and is a vital part of the Indian culture. *See id.*

45. *See* Allison M. Dussias, *Geographically-Based and Membership-Based Views of Indian Tribal Sovereignty: The Supreme Court's Changing Vision*, 55 U. PITT. L. REV. 1, 4 (1993).

46. *See* Matal, *supra* note 27, at 291.

47. *See id.*

48. *See* Vanessa J. Jimenez & Soo C. Song, *Concurrent Tribal and State Jurisdiction Under Public Law 280*, 47 AM. U. L. REV. 1627, 1647 (1998).

tribes as "in a state of pupilage" where their relationship to the United States "resembles that of a ward to his guardian."[49] The United States government was entrusted to protect the interests of Native Americans by, in theory, keeping Congress from enacting laws that were detrimental to Indians and stopping the executive branch from acting in opposition to the best interests of Native Americans.[50] The Department of the Interior currently works on behalf of tribal governments to fulfill these trust responsibilities.[51]

B. Narrowing of Tribal Sovereignty by the Federal Government

Deference to the tribal sovereignty did not last long, as the federal government began to narrow the rights of Indians and their tribes.[52] The tribes became enveloped by states and state laws began to trespass on Indian sovereignty.[53] The federal government began to heavily regulate criminal affairs on the reservations and replaced tribal jurisdiction with federal jurisdiction.[54] As non-Indians swept west, encroaching on Indian lands, these Indians were forced to move further west.[55]

---

49. *Cherokee Nation*, 30 U.S. (5 Pet.) at 17.

50. *See* Jimenez & Song, *supra* note 48 at 1648.

51. *See* DEPARTMENT OF THE INTERIOR, STRATEGIC PLAN *at* FY 2000-2005, 7 (2000). The DOI's mission is "to protect and provide access to our nation's natural and cultural heritage and honor our trust responsibilities to Indian tribes and our commitments to island communities." *Id.*

52. *See* FELIX COHEN, HANDBOOK OF FEDERAL INDIAN LAW 78 (3d ed. 1942).

53. *See* Matal, *supra* note 27, at 292. For a detailed history of the development of federal and state law on tribal governments, see *id.* at 292-315.

54. Indians who commit any other crimes specified in the Indian Major Crimes Act "shall be subject to the same law and penalties as all other persons committing any of the above offenses, within the exclusive jurisdiction of the United States" and if federal law does not define or punish any of the offenses, then state law applies. Indian Major Crimes Act, ch. 341, 23 Stat. 362, 385 (1885), (codified as amended at 18 U.S.C. § 1153 (1994)); *see also* COHEN, *supra* note 52, at 146-47. This created a paradox wherein the federal government maintained power over tribal criminal jurisdiction when only states have the power to exercise general police power within their domain. *See* Matal, *supra* note 27, at 292. The United States Congress solved this problem by creating federal enclaves with exclusive federal jurisdiction. *See id.* at 301.

55. *See* POMMERSHEIM, *supra* note 2, at 18. When American Indians could no longer be moved any further west after the admission of California following the Mexican-American War of 1846-48, the federal government forced these Natives to settle on reservations within state boundaries. *See* Matal, *supra* note 27, at 294-95. Congress made these new reservations federal enclaves with exclusive federal jurisdiction and began to make all Indian country into federal enclaves with

Decisions from the United States Supreme Court towards the end of the 1800s evidenced a movement toward the assimilation of tribes into American society.[56] The General Allotment Act of 1887[57] allowed the Bureau of Indian Affairs to give each head of household 160 acres and 40 to each minor.[58] The Act also gave "surplus" reservation lands to white settlers.[59] This assimilationist policy had a destructive effect on the Indian way of life.[60] The homelands of the Indians

> were cut open. The bright line separating Indians from non-Indians was obliterated. Much land was lost as many non-Indian settlers came into Indian country. Cultural ways were strained, and traditional tribal institutions were undermined and weakened. For many, this was the most devastating historical blow to tribalism and Indian life.[61]

In addition, in order to punish the criminal defendant in *Donnelly v. United States*,[62] the Supreme Court decided to redefine the term "Indian country" in an attempt to divorce the term from the idea of sovereignty.[63] The Supreme Court in *Donnelly* specifically separated the term "Indian country" from the idea of inherent Indian sovereignty.[64] Thus, non-sovereign Indian country was created.[65] This case later formed the basis for 18 U.S.C. § 1151(a), enacted in 1948, "which brings in all federal Indian reservations within the scope of Indian country."[66]

---

the permission of the states. *See id.* at 295; *see also* U.S. CONST. art I., § 8, cl. 17.

56. *See* COHEN, *supra* note 52, at 78. The government justified this policy as benefiting American Indians because if they were to assimilate into American society and adopt a civilized lifestyle, the Indians would not need as much land and there would be more available for the white settlers. *See id.*

57. Act of Feb. 8, 1887, ch. 119, 24 Stat. 388 (1887) (codified as amended at 25 U.S.C. §§ 331-34, 339, 341-42, 348-49, 354, 381 (1994)).

58. *See* POMMERSHEIM, *supra* note 2, at 19.

59. *See id.* at 20.

60. *See id.* at 19.

61. *Id.* As a result of the Act, Indian land was reduced in value from $138 million to $52 million and $60 million of that was lost because of the "surplus" land provisions. *See id.* at 20.

62. 228 U.S. 243 (1913). Donnelly, a white man, was convicted of murdering an Indian within the extension of a reservation. He argued, among other things, that Indian country contained only the land that a tribe held under their original right of possession. *See id.* at 268.

63. *See* Matal, *supra* note 27, at 309.

64. *See id.* at 311.

65. *See id.*

66. 18 U.S.C. § 1151(a) (1994). The statute defines "Indian country" as:

> (a) all land within the limits of any Indian reservation under the juris-

C. Indian Reorganization Act

The destructive practices of the federal government towards American Indians were reversed in 1934 with the enactment of the Indian Reorganization Act (IRA).[67] In 1936, Congress extended the IRA to include Alaska in an effort to retreat from previous harmful policies towards Alaskan Natives.[68] This Act created a federal recognition of tribes and was designed to help tribes gain self-sufficiency after years of forced assimilation by creating loans for Indian economic development.[69] The IRA ended the policy of the General Allotment Act,[70] which had given certain lands to Indians to be used for reservations and gave the rest to non-Indians.[71] The IRA also allowed the tribes to form a tribal court system and write a constitution, and most tribal courts today are the result of the IRA.[72] Although the IRA

> diction of the United States Government, notwithstanding the issuance of any patent, and, including rights-of-way running through the reservation, (b) all dependant Indian communities within the borders of the United States whether within the original or subsequently acquired territory thereof, and whether within or without the limits of a state, and (c) all Indian allotments, the Indian titles to which have not been extinguished, including rights-of-way running through the same.

Id.

67. 25 U.S.C. §§ 461-497 (1994); *see also* POMMERSHEIM, *supra* note 2, at 22. Tribal culture was scarred by the assimilation policy because Indian land was opened up to non-Indians, weakening the authority of tribal governments. *See id.* While the IRA was intended to support tribal governments, many Native Americans saw the Act as a further disturbance to the tribal way of life, as the governments are tainted by the influence of "the 'white man's way' of elections, the use of English, and the written word." *Id.* Thus, tribal governments originating from the IRA are sometimes seen as less legitimate. *See id.* at 23.

68. *See* 25 U.S.C. § 473(a) (1994).

69. *See* 25 U.S.C. § 470 (1994); *see also* F. Henry Ellis, III, *Indian Tribal Sovereignty and the Tribal Courts: The Myth and the Reality*, 13 SUFFOLK TRANSNAT'L L.J. 714, 720 (1990).

70. General Allotment Act, ch. 119, 24 Stat. 388 (codified as amended at 25 U.S.C. §§ 331-496) (1994) (repealed Oct. 21, 1976).

71. *See id.*; *see also* Patricia Thompson, *Recognizing Sovereignty in Alaska Native Villages After the Passage of ANCSA*, 68 WASH. L. REV. 373, 378 (1993).

72. *See* Fredric Brandfon, *Tradition and Judicial Review in the American Indian Tribal Court System*, 38 UCLA L. REV. 991, 998 (1991). Modern tribal councils and courts in Alaska work on ICWA and adoption cases, enforce community laws such as alcohol violations and generally maintain order within the Village. *See* Susanne Di Pietro, *Tribal Court Jurisdiction and Public Law 280: What Role for Tribal Courts in Alaska?*, 10 ALASKA L. REV. 335, 337 (1993). These courts are low-cost, staffed by volunteers and able to resolve the disputes of the tribal members in domestic relations and minor criminal violations successfully. *See id.* at 338. In 1993, more than a hundred Native Alaskan Villagers took their

shows that Congress treated Native Alaskans the same as American Indian tribes, at least until 1971 when the Alaska Native Claim Settlement Act was enacted, the Act also showed that Congress saw the right to self-government as a right granted by the United States government and not an inherent right of the tribal government.[73]

D. Public Law 280 and the Build-up of State Powers Over Indians

Congress' encroachment on Indian powers continued with the ratification of Public Law 280.[74] Congress enacted Public Law 280[75] in 1953, which transferred to five specified states jurisdiction "over offenses committed by or against Indians" within Indian country.[76] Alaska was added to the list of mandatory Public Law 280 states in 1958.[77] This law was the first general federal law that gave states jurisdiction over Indians.[78] In 1968, the law was modified so that state law could not extend to Indian country unless the tribe consented.[79] No tribe has done so since 1971.[80] The law did not end tribal jurisdiction but tribal courts were now forced to exercise concurrent jurisdiction with the state in some areas and relinquish jurisdiction in others.[81] The state courts can choose to apply tribal or state

---

disputes to the tribal court or council. *See id.*

73. *See* Thompson, *supra* note 71, at 378-79; *see also* Ellis, *supra* note 69, at 720-21. Sovereignty is not a right granted by the United States government as it predates the United States Constitution and is an inherent right. *See* Thompson, *supra* note 5, at 428.

74. 28 U.S.C. § 1360(a) (1994). Congress used its authority to delegate its jurisdictional power to the states broadly with Public Law 280. *See* Jimenez & Song, *supra* note 48, at 1656-57. This law upsets the traditional balance of power between the tribe and the federal government as it unilaterally grants jurisdiction to the states over the tribes. *See id.* at 1657.

75. 28 U.S.C. § 1360(a) (1994).

76. *See* Jimenez & Song, *supra* note 48, at 1657.

77. *See* Act of Aug. 8, 1958, P.L. No. 85-615, § 2, 72 Stat. 545 (1958). Congress enacted Public Law 280 to address three points: "lawlessness on reservations, the desire to assimilate Indian tribes into the population at large, and a shrinking federal budget for Indian affairs." Jimenez & Song, *supra* note 48, at 1659.

78. *See* Julie A. Pace, *Enforcement of Tribal Law In Federal Court: Affirmation of Indian Sovereignty or a Step Backward Towards Assimilation?*, 24 ARIZ. ST. L.J. 435, 449 (1992). The Law was enacted to restrain federal spending and improve local law enforcement and was enacted in a period of pervasive assimilationist attitudes. *See* Jimenez & Song, *supra* note 48, at 1658-59.

79. *See* Pace, *supra* note 78, at 449.

80. *See id.* at 450.

81. *See id.* at 450-51. Congress did not explicitly state that Public Law 280 extinguished tribal sovereignty; thus, tribal jurisdiction continues unless Congress

law in cases that arise out of Indian country and in civil cases, the state is required to utilize tribal law unless inconsistent with state law.[82] Tribal law is now rarely exercised in state courts.[83]

The Supreme Court of Alaska has interpreted this statute several times. First, in *Native Village of Nenana v. State of Alaska*,[84] the Supreme Court of Alaska concluded that the reassumption requirement in the Indian Child Welfare Act[85] meant that Congress' intent in enacting Public Law 280 was to grant the chosen states exclusive jurisdiction in child custody disputes.[86] Thus, as the Alaskan Supreme Court interpreted Public Law 280, chosen states retain exclusive jurisdiction over the dispute unless the tribe petitions the Secretary of the Interior to request tribal jurisdiction. Later, in *In re F.P.*,[87] the Alaskan Supreme Court reaffirmed its finding that Public Law 280 grants states exclusive jurisdiction in child custody disputes.[88] This will be important for a discussion of the dissent in *John v. Baker*.[89]

E. "Indian Country" and a Return to its Connection with Tribal Sovereignty

The Supreme Court began to again put the idea of sovereignty back into the term "Indian country" by assuming that whenever the criteria of 18 U.S.C. § 1151 were met, state law was pre-empted in the mid-1970s.[90] Thus, if a tribe's land were to fall within the definition of "Indian country,"[91] state courts would not have jurisdiction.[92] The Supreme Court in *DeCoteau v. District County Court*[93] stated that, according to section 1151, if the reservation was found to be a continuing one, tribal law and exclu-

---

explicitly divests it. *See id.* at 450.

82. *See id.* at 451.

83. *See id.*

84. 722 P.2d 219 (Alaska 1986).

85. 25 U.S.C. § 1918(a) (1994). The tribe may "reassume jurisdiction over Indian child custody proceedings" by submitting a petition to the Secretary of the Interior. *Id.*

86. *See Nenana*, 722 P.2d at 221.

87. 843 P.2d 1214 (Alaska 1992).

88. *See id.* at 1215. The court quoted its holding in *Nenana* that ICWA and Public Law 280 combine to give the state exclusive jurisdiction over the child custody determinations of Indian children. *See id.* at 1215-16.

89. *See infra* notes 202-04 and accompanying text.

90. *See* Matal, *supra* note 27, at 314.

91. *See* 18 U.S.C. § 1151 (1994).

92. *See* Matal, *supra* note 27, at 316-17; *see also* DeCoteau v. District County Court, 420 U.S. 425, 427 (1975) (stating that both of the parties stipulated that state courts do not have jurisdiction over a dispute if the conflict arose on "Indian country" as defined by section 1151).

93. 420 U.S. at 425.

sive federal law would apply.[94] However, the Court in *Donnelly* (from which the codification of section 1151 had emerged) said that the definition of "Indian country" was not meant to give the United States "sole and exclusive jurisdiction over Indian country."[95] Since the *DeCoteau* case did not find that the land in question fell within the definition of "Indian country,"[96] this distinction was not relevant, but is an example of the change in the Court's thinking which has been applied in subsequent decisions.[97]

F. Alaska's Unique Jurisdictional Development

In the 1970s, the situation in Alaska took a turn in a different direction from standard Indian jurisprudence in America. Alaska decided to adopt a unique method of handling Indian claims to land. In 1971, the Alaska Native Claim Settlement Act (ANCSA)[98] was enacted to settle claims by Native Alaskans over aboriginal lands without creating reservations or other means of "lengthy wardship or trusteeship."[99] In order to achieve this goal,

---

94. *See DeCoteau*, 420 U.S. at 427 n.2. The United States Supreme Court found the reservation at issue in *DeCoteau* had been extinguished. *See id.* at 427-28. Thus, as the Native American tribal land did not have reservation status, state law governs. *See id.* at 428.

95. 228 U.S. at 268.

96. 420 U.S. at 445.

97. *See* Matal, *supra* note 27. Some lower courts have not applied the reasoning of *DeCoteau*, refusing to find that section 1151 preempts state jurisdiction, while other lower courts have felt bound to follow *DeCoteau*. *See id.*

98. 43 U.S.C. §§ 1601-1628 (1988).

99. *Id.* Native Americans have been in a trust relationship with the federal government since Chief Justice Marshall described their relationship as that of a guardian to a ward. *See* Cherokee Nation v. Georgia, 30 U.S. (5 Pet.) 1, 17 (1831). Although the stated purpose of ANCSA was to settle land claims without affecting the relationship of the tribes to the government, some have interpreted its language to mean that the drafters were supporting a "complete 'termination' of the trust status of Alaska Natives." Walsh, *supra* note 27, at 244. Hence, the sovereignty of Alaskan Native villages would not be recognized. However, although ANCSA did change the relationship between the villages and the government with respect to the creation of corporations, it did not explicitly sever the trust relationship between the two. *See id.* at 246. The Supreme Court has ruled that in order to break the federal/Indian trust relationship, a bill must "explicitly and clearly" terminate the relationship. *See id.* ANCSA did not do so. *See id.* The relationship of the federal government toward individuals was not changed by ANCSA and subsequent legislation affirms the Congress' intent to maintain that relationship by recognizing Native villages as federally recognized tribes and by continuing to provide services to them. *See id.* The corporations are also in a trust relationship with the government, as evidenced in part by ANCSA itself, which gives the corporation special

ANCSA created village and regional corporations with the Native Alaskans as corporate stock recipients.[100] Two hundred village corporations were created as well as twelve regional ones.[101] Members of the villages, who were at least one-quarter Native Alaskan, were entitled to become shareholders in either a village or regional corporation, or both.[102] The government then gave the corporations $962.5 million in cash and 44 million acres of land.[103] All of the reservations in Alaska were extinguished with the exception of Metlakatla.[104] Interpretations of the effect of this language have differed.[105]

G. Congress' Recent Legislation Regarding the Custody of Indian Children

Congress began to move back towards a benevolent view of tribal sovereignty in regards to Indian children. In 1978, Congress moved to correct the growing problem of large numbers of Indian children being removed

---

protection. *See* Walsh, *supra* note 27, at 248. For example, the corporation's land is not subject to local or federal taxes and the lands are only alienable with the consent of the Secretary of the Interior. *See id.* at 249.

100. *See* 43 U.S.C. §§ 1605-1607 (1988). The money was given to corporations chartered by the State of Alaska rather than directly to the tribal governments themselves. *See* Summit, *supra* note 3, at 615. Senator Henry Jackson, principal drafter of ANCSA, intended the Act to be an assimilation vehicle with the anticipated effect of pushing Native Alaskans into the business world and requiring them to administer their land. *See id.* at 620.

101. *See* Summit, *supra* note 3, at 615.

102. *See id.*

103. *See* 43 U.S.C. §§ 1605-1607 (1988). The land given to the Villages was approximately ten percent of the land in Alaska. *See* Summit, *supra* note 3, at 614. The title of the land was given to the village corporation but subsurface rights were not included. *See id.* at 616.

104. *See* 43 U.S.C. § 1618 (1988).

105. One commentator states that the effect of this language is unambiguous: tribal sovereignty is extinguished. *See* Matal, *supra* note 27, at 339-40. However, he notes that this conclusion is not shared by everyone and gives references to other law review articles that conclude that Native Alaska tribes are sovereign nations. *See id.* at 341 n.459. State courts in Alaska historically found that Native Villages did not have sovereignty since they did not occupy "Indian country" after the enactment of ANCSA. *See* Thompson, *supra* note 71, at 380-81. The federal courts have directly contradicted the Alaska state courts by consistently holding that Native Alaskans are entitled to the same benefits of all American Indians and therefore have inherent sovereignty to decide certain issues on their own. *See id.* at 382.

from their homes by non-tribal agencies, which then placed Indian children with non-Indian families.[106] Congress enacted the Indian Child Welfare Act (ICWA)[107] in order to undertake the responsibility through "statutes, treaties and the general course of dealing with Indian tribes"[108] of protecting Indian tribes and their resources.[109] According to Congress, the most important resource to an Indian tribe, is the children of the tribe and thus Congress passed legislation to protect those children who are members or who are eligible for membership.[110] The ICWA gives tribal courts exclusive jurisdiction over any child welfare case where an Indian child resides or is domiciled on a reservation that is not subject to state jurisdiction by federal law.[111] There are three exceptions to the exclusive jurisdiction of the tribal court: an objection of a parent to the tribe's jurisdiction; good cause by the state court to not allow tribal jurisdiction; and a declination by the tribe of jurisdiction.[112] According to ICWA, the tribal courts perform a number of functions, including "receiving notices of and intervening in children's proceedings in state court and adjudicating such disputes as where a child has been declared a 'ward' of the court."[113] The tribal court can also have exclusive jurisdiction over child custody and adoption cases when it petitions the Secretary of the Interior.[114]

As the above history indicates, Alaska historically denied tribes the right to adjudicate disputes internally unless the conflict arose in Indian country, and even then Alaskan courts removed jurisdiction when they found it

---

106. *See* 25 U.S.C. § 1902 (1994).

107. 25 U.S.C. § 1901 (1994).

108. This power was granted to them by the Constitution. *See* U.S. CONST. art. I, § 8, cl. 3.

109. 25 U.S.C. § 1901(2) (1994).

110. *See* 25 U.S.C. § 1901(3) (1994).

111. *See* Barbara Ann Atwood, *Fighting Over Indian Children: The Uses and Abuses of Jurisdictional Ambiguity*, 36 UCLA L. REV. 1051, 1059 (1989). However, this "exclusive" jurisdiction is limited by Public Law 280, which gives the states power to decide criminal and civil disputes. 28 U.S.C. § 1360 (1988).

112. *See* 25 U.S.C. § 1911(a) (1994). The text reads:

An Indian tribe shall have jurisdiction exclusive as to any State over any child custody proceeding involving an Indian child who resides or is domiciled within the reservation of such tribe, except where such jurisdiction is otherwise vested in the State by existing Federal law. Where an Indian child is a ward of a tribal court, the Indian tribe shall retain exclusive jurisdiction, notwithstanding the residence or domicile of the child.

Id.

113. Eric Smith & Mary Kancewick, *The Tribal Status of Alaska Natives*, 61 U. COLO. L. REV. 455, 459 (1990).

114. *See* 25 U.S.C. § 1918(a) (1994).

wise to do so.[115] In addition, decisions of tribal courts were always reviewable in federal court.[116]

As the next section will show, in *John v. Baker*,[117] the Supreme Court of Alaska pushed the idea of the inherent sovereignty of tribes even further from its original roots in the Cherokee cases.[118] The court held that a Native village outside Indian country does "possess the inherent sovereign power to adjudicate child custody disputes between tribal members in their own courts."[119]

II: *JOHN V. BAKER*: AN ADVENTURE IN A TRIBAL LABYRINTH

A. Case Background

John Baker, a member of Northway Village, and Anita John, a member of Mentasta Village,[120] had two children together: John Jr. and Emmanuel.[121] The family lived together unwed in Ms. Johns' village until the couple separated in 1993. At that time John Jr. was two years old and Emmanuel was one.[122] Mr. Baker and Ms. John shared custody over the children until 1995, at which time Mr. Baker kept the children, refusing to give them to Ms. John.[123] At that time, Mr. Baker filed a custody petition for sole custody with the Northway Village.[124] Both parents attended the custody hearing on August 29, 1995, and the tribal court ordered temporary shared custody on an alternating monthly basis.[125] Four members of that tribal court were relatives of Ms. John including her sister, grandmother and two cousins.[126] In addition, the tribal court judge had previ-

---

115. *See supra* Part I.A.-G.

116. *See* Nathan R. Margold, *Introduction to* FELIX COHEN, HANDBOOK OF FEDERAL INDIAN LAW x (3d ed., 1942).

117. 982 P.2d at 738.

118. *See supra* Part II.

119. *Baker*, 982 P.2d at 743.

120. Congress and the executive branch determine what Native American groups are recognized as sovereign tribes. *See infra* note 213 and accompanying text. Most of the Villages in Alaska are recognized as "tribes." *See infra* note 214 and accompanying text.

121. *See Baker*, 982 P.2d at 743. Both Northway Village and Mentasta Village are Native Alaskan villages. *See id.*

122. *See id.*

123. *See id.*

124. *See id.*

125. *See Baker*, 982 P.2d at 743.

126. *See* Baker v. John, No. 4FA-95-3103 (Alaska Sup. Ct. Sept. 2000) (order denying comity) *at* http://www.alaska.net/~akctlib/4fa95310.txt (last visited Nov. 17, 2000).

ously spoken with Ms. John's sister "[o]ut of common respect and traditional custom," as Ms. John's sister was the First Chief of the Mentasta tribe.[127] Tribal Court Judge Lorraine Titus told Mr. Baker and Ms. John that she would revisit the issue one year later when John Jr. was to enter school.[128]

Mr. Baker appealed the decision of the tribal judge in tribal court but his request was denied.[129] The custody arrangement continued on an alternating schedule from September until December except for the short time when Ms. John had sole custody while Mr. Baker was in jail on a DWI charge.[130] In December, Mr. Baker filed another custody petition, this time to the Alaska state court.[131] In this complaint, Mr. Baker stated that he was unaware of any other custody dispute in regard to his children.[132] Ms. John filed a motion to dismiss based on the tribal court proceedings, but her dismissal was denied.[133]

The Alaskan state court ruled that ICWA did not apply to custody disputes between parents and thus the state court had subject matter jurisdiction.[134] The judge also ruled that even if its jurisdiction was concurrent with the tribal court's jurisdiction, the facts of the case dictate that the state court should assert jurisdiction. The court cited as justification that both parents came from different villages and the state court's ability to access to a child custody investigator.[135]

The Alaska state court ordered Mr. Baker and Ms. John to continue sharing custody on an alternating monthly schedule which they did until April 1996.[136] The state court at that time altered its decision and granted Mr. Baker primary custody with Ms. John receiving visitation every other weekend and at least eight weeks during the summer.[137] The court decided that, although both Mr. Baker and Ms. John previously had problems with substance abuse, Mr. Baker was in better control of his problems, as Ms.

---

127. *Id.* The two agreed that the Northway and Mentasta villages should decide the case jointly. *See id.* The tribal court was made up of four or five members of the Northway Village, four blood relatives of Ms. John from the Mentasta Village and two child care providers of the two children while they were in Ms. John's custody. *See id.*

128. *See* John v. Baker, 982 P.2d 738, 743 (Alaska 1999).

129. *See id.*

130. *See id.*

131. *See id.*

132. *See id.*

133. *See* John v. Baker, 982 P.2d 738, 743 (Alaska 1999).

134. *See id.*

135. *See id.* at 743-44.

136. *See id.* at 744.

137. *See id.*

John was still dealing with severe depression.[138] The Supreme Court of Alaska granted Ms. John's appeal.[139]

B. The Appellate Arguments to Alaska Supreme Court

Ms. John argued that Northway Village had the inherent sovereign power to adjudicate custody disputes.[140] Since the tribal court had the authority to decide the domestic relations issue, and this authority was concurrent with the jurisdiction of the state, Ms. John argued that the state should have deferred to the tribal court's decision.[141] First, she argued that the prior decisions of the Supreme Court of Alaska in *Native Village of Nenana v. Dept. of Health and Social Services*[142] and *In re F.P.*,[143] do not need to be overturned because those decisions did not apply to this case since they were decided under ICWA.[144] Second, she argued that even if these cases did apply, the court should reconsider their holdings.[145]

Mr. Baker argued that Northway Village did not have jurisdiction and even if they did, there were problems with the tribal court's ruling which violated his due process rights.[146] Also, Mr. Baker asserted that the state

---

138. *See* John v. Baker, 982 P.2d 738, 744 (Alaska 1999).

139. *See id.*

140. *See id.* Alaskan state courts have constantly denied that Native Villages possess sovereign powers while federal courts have found the opposite, even though they both relied on the same federal law. *See* Thompson, *supra* note 71, at 373.

141. *See Baker*, 982 P.2d at 744.

142. 722 P.2d 219 (Alaska 1986) (holding that Congress intended Public Law 280 to grant states the exclusive jurisdiction over child custody matters).

143. 843 P.2d 1214 (Alaska 1992) (holding that if the tribe does not successfully petition the Secretary of the Interior, the tribe has no jurisdiction over a child custody case).

144. *Baker*, 982 P.2d at 744.

145. *Id.*

146. *Id.* Mr. Baker argued that his due process rights were violated because the tribal courts do not have an appellate process. *See id.* at 764. The Supreme Court of Alaska did not decide whether the absence of tribal appellate review violated due process in this instance. *See id.* at 764. The court did note that while appellate review is required in order for foreign courts to comport with due process, state courts are not required to have an appellate system. *See id.* On remand the superior court held "that Mr. Baker did not receive the benefit of guaranteed due process protection . . . ." Baker v. John, No. 4FA-95-3103 (Alaska Super. Ct. Sept. 2000) (order denying comity) *at* http://www.alaska.net/~akctlib/4fa95310.txt (last visited Nov. 17, 2000). Specifically, the court found that the Northway Village did have an appellate process in place that Mr. Baker did not take advantage of, therefore, the court did not need to decide if the absence of an appellate system violated due process. *See id.* Next, the court found that Mr. Baker did receive proper and timely notice of matters which would have been determinative in the

court should have jurisdiction because his children were not members of Northway Village.[147]

III. EXAMINATION OF ALASKAN SUPREME COURT PRECEDENTS

A. *Nenana* and *In Re F.P.* Are Not Controlling

The majority of the Supreme Court of Alaska first found that their previous decisions in *Nenana*[148] and *In re F.P.*[149] did not apply to *John v. Baker*.[150] Although these decisions did involve the custody of children, they were based on interpretations of ICWA[151] and Public Law 280[152] in combination.[153] In *Nenana*, the court found that "Congress intended that Public Law 280 give certain states, including Alaska, exclusive jurisdiction over matters involving the custody of Indian children, and that those states exercise such jurisdiction until a particular tribe petitions to reassume jurisdiction over such matters, and the Secretary of the Interior approves [the] tribe's petition."[154] However, Supreme Court of Alaska in Nenana

---

custody of the children. *See id.* However, Mr. Baker's due process rights were violated since his father was not allowed to testify completely, Judge Titus' interactions with Ms. John's relatives raised concerns and Ms. John's relatives all played a significant role in the hearing. *See id.*

147. *See Baker*, 982 P.2d at 744. The tribal court has subject matter jurisdiction only over those internal disputes that arise between tribal members. *See id.* at 759. The court was not given any information about the tribal membership of the children, an issue which could take jurisdiction away from the tribal court. *See id.* at 764. The superior court on remand found that John Jr. and Emmanuel were not members of Northway Village but they were eligible for membership. *See* Baker v. John, No. 4FA-95-3103 (Alaska Super. Ct. Sept. 2000) (order denying comity) *at* http://www.alaska.net/~akctlib/4fa95310.txt (last visited Nov. 17, 2000). Thus, the tribe has adequate subject matter jurisdiction over them. *See id.*

148. 722 P.2d 219 (Alaska 1986).

149. 843 P.2d 1214 (Alaska 1992).

150. 982 P.2d at 748. The dissent argued that state law governed outside of Indian country and that Public Law 280 requires that state law governs when a state law conflicts with a tribal law. *See id.* at 801 (Matthews, J., dissenting). By concluding that tribal law trumps state law in this case, the Supreme Court of Alaska, according to the dissent, is allowing forum-shopping, a practice the Supreme Court tried to avoid in *Erie v. Tompkins*. *See id.* at 803 (Matthews, J., dissenting) (citing *Erie*, 304 U.S. 64 (1938)).

151. 25 U.S.C. §§ 1901-1963 (1994).

152. 28 U.S.C. § 1360 (1988).

153. *See Baker*, 982 P.2d at 745.

154. *Nenana*, 722 P.2d at 221. Thus the state court, not the tribal court, has exclusive jurisdiction over the custody of Indian children in states governed by Public Law 280 and the tribal court can only try to assume jurisdiction under

admitted that other commentators have interpreted Public Law 280 to give the state concurrent jurisdiction with the tribal court.[155]

The Ninth Circuit's decision in *Native Village of Venetie I.R.A. Council v. Alaska*[156] failed to persuade the Alaskan Supreme Court to overrule its prior decision, *In re F.P.*[157] The Supreme Court of Alaska merely stated in *In re F.P.* that it continued to believe that its interpretation of Public Law 280 was correct: Public Law 280 granted states exclusive jurisdiction in child custody cases.[158] However, Chief Justice Rabinowitz dissented, asserting that he was persuaded by the Ninth Circuit's analysis and believed that "it is inconsistent with the doctrine of inherent tribal sovereignty to conclude that § 1918 of the ICWA and Public Law 280, taken together, divest tribes of even concurrent jurisdiction over child custody matters."[159] As *Nenana*[160] and *In re F.P.*[161] were based on both ICWA and Public Law 280, the Alaskan Supreme Court held in *Baker* that those decisions were not controlling.[162]

The Supreme Court of Alaska was not, at the time of *In re F.P.*, bound to accept the decision of the Ninth Circuit.[163] As there was a federal question involved (the interpretation of ICWA, a federal statute), the state courts in Alaska are only bound by decisions of the United States Supreme Court.[164] Similarly, the federal courts in Alaska are not bound by state court decisions.[165] Although *Venetie II* has since overturned *Venetie I.R.A.*,

---

ICWA if the Secretary approves. *See id.*

155. *See id.* (citing F. COHEN, HANDBOOK OF FEDERAL INDIAN LAW, 344-45 (1982 ed.) and D. CASE, ALASKA NATIVES AND AMERICAN LAWS, 490 n. 119 (1978)).

156. 944 F.2d 548 (9th Cir. 1991).

157. 843 P.2d at 1215.

158. *See id.* at 1216.

159. *Id.* at 1217-18 (Rabinowitz, C.J., dissenting). *Venetie I.R.A.* was later overruled by the Supreme Court. *See* Alaska v. Native Village of Venetie Tribal Gov't., 522 U.S. 520 (1998) [hereinafter *Venetie II*] (holding that Public Law 280 does not apply to Alaska territory after the enactment of ANCSA). The Ninth Circuit decided the issue utilizing a two-part analysis. *See Baker*, 982 P.2d at 745. The court looked to whether the Native Village had inherent sovereignty in the area of domestic relations and whether Congress had divested the Village of this aspect of their sovereignty. *See id.* 745-46 (citing Native Village of Venetie I.R.A. Council v. Alaska, 944 F.2d 548, 556, 558 (9th Cir. 1991)).

160. 722 P.2d 219 (Alaska 1986).

161. 843 P.2d 1214 (Alaska 1992).

162. *See Baker*, 982 P.2d at 748.

163. *See In re F.P.*, 843 P.2d at 1215 n.1.

164. *See id.*

165. *See id.*

its reasoning was persuasive.[166]

In *Venetie I.R.A.*, the United States Court of Appeals for the Ninth Circuit decided, in the five years since *Nenana*, that a Native Village could be considered sovereign and that Public Law 280 granted the states concurrent jurisdiction with tribes only in child custody matters.[167] The court also stated earlier in the opinion that Congress distrusted the states in deciding child custody disputes because they have "often failed to recognize the essential tribal relations of Indian people and the cultural and social standards prevailing in Indian communities and families."[168] Thus, the court found that Congress would not have given the state exclusive jurisdiction under ICWA.[169] As ICWA was enacted for the benefit of American Indians, it is disingenuous to believe that this statute deprived American Indians of jurisdiction over their members.[170]

B. The Indian Child Welfare Act and Public Law 280

The Supreme Court of Alaska decided in *John v. Baker* that their previous decisions were not controlling precedent because ICWA[171] and Public Law 280,[172] the statutes interpreted in *Nenana* and *In re F.P.*, did not apply.[173]

1. ICWA Does Not Apply To Interparental Custody Disputes

First, ICWA relates only to a "child custody proceeding" which did not include "a placement based upon . . . an award, in a divorce proceeding, of

---

166. *See supra* note 159 and accompanying text.

167. *See Venetie I.R.A.*, 944 F.2d 548 at 558-59, 562. The issue in *Venetie I.R.A.* was whether a state court in Alaska must give full faith and credit to a determination of child custody made by a tribal court. *See id.* at 550. In *Venetie I.R.A.*, two women adopted babies through a tribal court but were refused AFDC benefits as the state would not recognize the purported adoptions. *See id.* at 550-51.

168. *Id.* at 553 (citation omitted). The Ninth Circuit found that Congress would not permit state courts alone to have jurisdiction over child custody because it never believed that state courts protected tribal interests. *See id.* at 553-54. The court also presumed that Congress was knowledgeable of prior law and understood that statutes passed for the benefit of Native Americans would be construed liberally in favor of such tribes. *See id.* at 554.

169. *See Venetie I.R.A.*, 944 F.2d at 554. The court also stated that the state court must defer to the tribal court unless there is good cause. *See id.* at 555.

170. *See* Atwood, *supra* note 113, at 1058-59.

171. 25 U.S.C. §§ 1901-1963 (1994).

172. 28 U.S.C. § 1360 (1988).

173. *See Baker*, 982 P.2d at 746-48.

custody to one of the parents."[174] Although Mr. Baker and Ms. John were never married, the court found that the divorce exception was meant to apply to all parental custody disputes.[175] The policy goals behind the enactment of ICWA (to end the removal of Indian children from Indian households and into non-Indian homes) were not implicated during a divorce proceeding nor are they implicated in this case, where both parents are Indian and the state is not removing the children.[176]

Although ICWA did not apply to *John v. Baker*, the intent of Congress in enacting the statute can be utilized as evidence that the court should allow Indians to adjudicate interparental custody disputes. The spirit in which ICWA was enacted was for the benefit of tribes and the future of their cultural identity.[177] It is congruent with an emerging federal policy of native self-determination.[178] The current trend in executive advocacy and federal statutes is tribal self-determination and this decision is consistent with that policy.[179] Also, this decision granting concurrent jurisdiction to state and tribal courts is helpful to Indian children who will not be disadvantaged by being cut off from their cultural lineage.

2. Public Law 280 Does Not Apply to Alaskan Lands After the Enactment of ANCSA

The Supreme Court's interpretation of Public Law 280 in *Alaska ex. rel. Yukon Flats School District v. Native Village of Venetie Tribal Government (Venetie II)*[180] suggested that Public Law 280 did not apply to Native Villages in Alaska after the enactment of the Alaska Native Claims Settlement Act.[181] The issue in *Venetie II* was whether the land at issue in

---

174. 25 U.S.C. § 1903 (1994). This provision is called the "divorce exception." *Baker*, 982 P.2d at 746.

175. *See Baker*, 982 P.2d at 746-47.

176. *See* 25 U.S.C. § 1902 (1994) (stating Congress' policy).

177. *See* Atwood, *supra* note 111, at 1058.

178. *See* Thompson, *supra* note 5, at 437. Since Public Law 280 was enacted in the 1950s, Congress has enacted legislation granting tribes the right to sue in federal court and ending the policy undertaken in Public Law 280 (namely, the unilateral granting of jurisdiction to state courts over tribal courts). *See id.* In addition, "Congress passed statutes restoring certain previously terminated tribes to federal recognition, strengthening tribal control over federal programs, and maximizing tribal jurisdiction in child custody and adoption proceedings involving Indian children." *Id.* at 437-38.

179. *See id.* at 438.

180. 522 U.S. 520 (1998).

181. *See Baker*, 982 P.2d at 747. The Supreme Court in *Venetie II* held that the tribe at issue did not occupy "Indian country" in Alaska because they did not satisfy the requirements set out in section 1151. *See also* 18 U.S.C. § 1151 (1994).

Northern Alaska was considered "Indian country" under 43 U.S.C. § 1151 after the passage of the Alaska Native Claims Settlement Act.[182] Since the language in ANCSA did not state whether the land is considered "Indian country," the Native Village of Venetie argued to the United States Supreme Court that this silence should be interpreted as leaving the jurisdiction of the Village intact.[183] The State argued the opposite, citing the Department of the Interior Solicitor's Opinion which said that ANCSA terminated all Indian country in Alaska.[184] The Court held that the land in question was not Indian country and thus that ANCSA had terminated Indian country in Alaska.[185]

The Supreme Court articulated a two-part test in order to determine whether land constitutes a "dependent Indian community" in *Venetie II*, but in so doing ignored two important principles of federal Indian law.[186]

Indian country is defined as (1) a reservation, (2) a dependent Indian community, or (3) an Indian allotment. *See id.* ANCSA revoked all reservations in Alaska with the exception of the Annette Island Reserve. *See Venetie II*, 522 U.S. at 527 n.2; *see also* Annette Island Information Page, *at* http://www.ptialaska.net/~aisd/#LOCATION (last visited Feb. 17, 2001) (explaining that these islands are located at the southern tip of the Alexander Archipelago of southeastern Alaska); Native Village of Stevens v. Alaska Mgmt. & Planning, 757 P.2d 32, 35 (Alaska 1988) (finding that the Metlakatlans, located on the Annette Island Reservation, were distinct from other native Alaskan groups as they lived on a reservation set up by Congress and maintained a solid central tribal organization unlike other Native Alaskan groups). Thus no Native Villages in Alaska satisfied the first requirement of section 1151. In addition, no Indian allotment was at issue in that case. *See id.*; *see also supra* Part I.C. (explaining Indian allotments). Regarding the dependent Indian community category, the Supreme Court construed the "dependent Indian communit[y]" as requiring two elements: (1) the land must have been set aside by the government for the Indians to use as their own land and (2) the land must be under federal superintendence. *See Venetie II*, 522 U.S. at 527. Land in Alaska after ANCSA did not satisfy either of these requirements. *See id.* Congress did not set aside land in Alaska under ANCSA for the Villages' use; rather, ANCSA revoked all existing reservations that had been set aside in the past. *See id.* at 532. Congress intended that the land would not be under federal superintendence as evidenced specifically in the statute. The statute was enacted to avoid "lengthy wardship or trusteeship." *Id.* *See* 43 U.S.C. § 1601(b) (1988); *see also* 522 U.S. at 532.

182. *Venetie II*, 522 U.S. at 523.

183. *See* Kristen A. Carpenter, *Interpreting Indian Country in State of Alaska v. Native Village of Venetie*, 35 TULSA L.J. 73, 74-75 (1999).

184. *See id.* at 75.

185. *See Venetie II*, 522 U.S. at 523.

186. *See* Stacy Belisle, Note, *American Indian Law – Tribal Lands - An Indian Tribe That Holds Title to Land by Transfer Under the Alaska Native Claims Settlement Act May Not Impose Business Tax When Its Land is Not Indian Country Within the Statutory Definition.* Alaska v. Native Village of Venetie Tribal Gov-

First, as a result of the trust relationship between the federal government and Native Americans, statutes that affect Indians must be construed liberally.[187] As ANCSA is ambiguous on whether Congress intended to extinguish Indian land and since it was enacted for the benefit of Native Alaskans, as a matter of federal Indian law, the statute should have been interpreted liberally in favor of the Native Alaskans.[188]

In addition, if Congress wants to terminate the trust relationship they enjoy with Native Americans, it must do so with a clear and plain assertion.[189] The Ninth Circuit asserted that Congress did not specifically state so in ANCSA.[190] Moreover, Congress demonstrated its intent to continue its trust relationship by declaring in ANCSA that the Act did not "replace, or diminish any obligation of the United States or of the State or [sic] Alaska to protect and promote the rights or welfare of Natives as citizens of the United States or of Alaska."[191] Another reason the Ninth Circuit believed that Congress did not intend ANCSA to terminate "Indian country" in Alaska was that Congress did not stop federal assistance to these tribes.[192]

Relying on the Supreme Court's interpretation in *Venetie II*, the Supreme Court of Alaska in *John v. Baker*[193] found that since the land of the Northway Village in Alaska did not qualify as "Indian country" under 18 U.S.C. § 1151, Public Law 280 did not apply.[194] Even if the Court in Ve-

---

ernment, 188 S. Ct. 948 (1998) 76 U. DET. MERCY L. REV. 903, 924-26 (1999). The Supreme Court asked: (1) whether the land was federally set aside for Indians and (2) whether the Indian community was under the federal government's supervision. *See id.* at 924. The land in Alaska did not meet this test because ANCSA revoked the federal set-aside condition when ANCSA transferred the land to corporations. *See id.* Also, the federal government no longer supervised the land after the corporations owned the land. *See id.*

187. *See* Alaska ex. rel. Yukon Flats Sch. Dist. v. Native Village of Venetie, 101 F.3d 1286, 1294 (Alaska 1996).

188. *See* Carpenter, *supra* note 183, at 75.

189. *See Yukon Flats*, 101 F.3d at 1294; *see also* Belisle, *supra* note 186, at 925-26.

190. *See Yukon Flats*, 101 F.3d at 1297. The court in *Yukon Flats* stated that Congress did not say that they wanted to terminate their relationship with Native Americans in ANCSA and in fact, the court made clear that federal superintendence was not extinguished. *See id.* at 1297-98.

191. 43 U.S.C. § 1601(c) (1994); *see also Yukon Flats*, 101 F.3d at 1297.

192. *See Yukon Flats*, 101 F.3d at 1297-98.

193. 982 P.2d at 743.

194. *See id.* at 748. *Venetie II* weakened claims of Native Alaskans by finding that Native Villagers living on ANCSA lands did not qualify as living in "Indian country." *See id.* If Northway Village was one of these such Villages, Public Law 280 did not apply as it applies only within "Indian country." *See id.*; *see also* 28 U.S.C. § 1360(a) (1988).

netie II[195] had found that some Villages in Alaska do occupy "Indian country," the majority in this case still would have come to the conclusion that *Nenana*[196] and *In Re F.P.*[197] were not relevant as the Supreme Court in Alaska had limited its previous holdings to cases where both Public Law 280[198] and ICWA[199] applied.

IV. INHERENT TRIBAL SOVEREIGNTY ENCOMPASSES THE TRIBE'S RIGHT TO ADJUDICATE CHILD CUSTODY DISPUTES

The crux of the dispute in *John v. Baker* is whether a tribe has the authority to adjudicate disputes covered by their sovereignty outside of "Indian country."[200] The majority of the Supreme Court of Alaska believed that the tribe does have this authority based on their inherent sovereignty.[201] The dissent, however, led by Chief Justice Matthews, strongly disagreed, as evidenced by his forty-eight page dissenting opinion.[202] However, as sovereignty is an attribute of an Indian tribe that predates the birth of the United States and which has not been limited by Congress,[203] the tribes should be allowed to regulate their own members even in the absence of Indian land.[204]

---

195. 522 U.S. 520 (1998).

196. 722 P.2d at 219. The Supreme Court of Alaska in *Baker* did not decide that *Nenana* and *In Re F.P.* were wrongly decided. *See Baker*, 982 P.2d at 748. Since neither ICWA nor Public Law 280 applied, it was not necessary. *See id.* The court noted that since *Nenana* and *In Re F.P.* were interpretations of ICWA and Public Law 280, and neither applied, the rationale behind the Supreme Court of Alaska's previous decisions was not specifically questioned by *Baker*. *See id.*

197. 843 P.2d at 1214.

198. 28 U.S.C. § 1360 (1988).

199. 25 U.S.C. §§ 1901-1963 (1994).

200. *See Baker*, 982 P.2d at 766. Alaska no longer has Indian country since ANCSA has settled all claims to aboriginal lands. *See id.* at 748.

201. *See id.* at 748-49.

202. *See id.* at 765-813. Matthews stated that the majority's opinion was "a long step away from the Alaska constitutional goal of equal rights under the law and is contrary to federal law[.]" *Id.* at 767. He disputed that tribal courts had the jurisdiction to settle custody disputes, that potential membership of Native American children alone was a sufficient basis for an exercise of jurisdiction, that tribal law controls even when conflicting with state law, and that Public Law 280 did not apply to *Baker*. *See id.*

203. *See* JEANETTE WOLFLEY & SUSAN JOHNSON, TRIBAL SOVEREIGNTY 1 (1996).

204. Indian land has not always been a prerequisite to finding tribal sovereignty exists. *See infra* Part IV.B.

A. Tribal Courts Should Have the Inherent Power to Adjudicate Child Custody Disputes

The Alaskan Supreme Court decided in *John v. Baker*, as a matter of first impression, that tribal courts in Alaska have the power, based on their inherent sovereign authority, to adjudicate child custody disputes.[205] The court expressed several bases for its decision.

1. Native Alaskan Villages Are "Tribes" Under the Federal Definition

In order for the United States government to recognize the existence of tribal sovereignty, the government must first recognize the group asserting sovereignty as a "tribe."[206] It is now generally accepted that Alaska Natives maintain the same relationship with the federal government as Native Americans located in the rest of the United States, but this was never unequivocally stated.[207]

Historically, it was uncertain whether Native Alaskan villages were considered "tribes."[208] Some of this uncertainty was quelled in 1936 when the Indian Reorganization Act was extended to Alaska.[209] The goal of this act was to promote tribal governments, attempting to secure the idea that Congress viewed Native Alaskan villages similarly to Native Americans in other states.[210]

However, with the enactment of ANCSA, the tribal status of Native Alaskans was again put into uncertainty.[211] ANCSA was specifically

---

205. *See Baker*, 982 P.2d at 749.

206. *See* Paul A. Matteoni, *Alaskan Native Indian Villages: The Question of Sovereign Rights*, 28 SANTA CLARA L. REV. 875, 880 (1988). The definition of tribe used by the courts originated from *Montoya v. United States*, 180 U.S. 261 (1901). The court defines "tribe" as "a body of Indians of the same or a similar race, united in a community under one leadership or government, and inhabiting a particular though sometimes ill-defined territory." *Id.* at 266.

207. *See* COHEN, *supra* note 52, at 404. The legal status of tribes has been "substantially established": the Native Alaskans are in the same relation to the federal government as other Native Americans. *See id.*

208. *See* Matteoni, *supra* note 206, at 879.

209. *See id.* at 880-81; *see also supra* Part I.C. Congress allowed tribal groups in Alaska to organize by adopting constitutions and creating reservations; *see* Smith & Kancewick, *supra* note 113, at 507.

210. *See id.* Compare with Native Village of Stevens v. Alaska Management and Planning, 757 P.2d 32, 40 (Alaska 1988) (finding that the IRA was an "express Congressional statement applicable to most Native groups in Alaska that they had not been recognized as tribes").

211. *See* Smith & Kancewick, *supra* note 113, at 509.

adopted to settle all aboriginal title and claim to Alaska lands.[212] Although some have argued that ANCSA terminated Native Alaskan tribes,[213] since this statute did not address the status of Native Alaskan villages, it could not have terminated the villages' tribal status.[214] Ambiguities must be construed in favor of Indians according to Indian canons.[215]

After ANCSA, however, Congress definitively included Native Alaskans as enjoying the same attributes as tribes.[216] In order to end confusion over the issue, "in 1993 . . . the Department of the Interior issued a list of federally recognized tribes."[217] One year later, Congress included Native Alaskans in the Federally Recognized Tribe List Act of 1994.[218] A list has been published every year since 1995 and has always included Native Alaskan Villages.[219] This is further evidence that Congress intended to recognize Alaskan Native Villages as sovereign entities.

2. Tribal Law Governs Unless Congress Explicitly Divests

The United States government recognized the inherent sovereignty of Native American tribes.[220] This sovereignty did not come from the U.S. government, did not depend on the grace of Congress, and predated the

---

212. *See id.*; 43 U.S.C. §§ 1601-28 (1988).

213. *See supra* note 107 and accompanying text (discussing the different interpretations of ANCSA).

214. *See* Smith & Kancewick, *supra* note 113, at 509. However, the Supreme Court of Alaska found in *Native Village of Stevens v. Alaska Management & Planning* that Native Alaskan villages had never been deemed as sovereign entities by the United States government. 757 P.2d 32, 34 (holding that most Native Alaskan villages "[are] not self-governing or in any meaningful sense sovereign").

215. *See Yukon Flats*, 101 F.3d at 1294.

216. *See Baker*, 982 P.2d at 749-50.

217. *Id.* at 749. The Department, in addition to stating that Native Alaskan villages are the equivalent of American Indian tribes, acknowledged that as tribes, these groups held the attributes of governments with their own authority. *See id.*

218. *See id.* at 750; 25 U.S.C. § 479(a) (1994). The dissent in *Baker* notes that although the Tribe List Act recognizes Native Alaskan villages as tribes, it does not extend any power or authority to them by virtue of this definition. 25 U.S.C. § 479 (1994); *see Baker*, 982 P.2d at 792. It merely grants them the same rights as other tribes in America and other tribes do not have the inherent power to decide child custody cases outside "Indian country." *See Baker*, 982 P.2d at 792. The same analysis applies to the Tribal Justice Act also relied upon by the majority. 25 U.S.C. §§ 3601-631 (1994); *see also Baker*, 982 P.2d at 793.

219. *See Baker*, 982 P.2d at 750.

220. *See* WOLFLEY & JOHNSON, *supra* note 203, at 6. The government has recognized tribal sovereignty through treaties, executive actions, legislation and judicial opinions but inconsistently according to the social, political and moral views of the day. *See id.*

United States Constitution and the nation itself.[221] The most basic principle in Indian law today is that:

> those powers which are lawfully vested in an Indian tribe are not, in general, delegated powers granted by express acts of Congress, but rather inherent powers of a limited sovereignty which has never been extinguished . . . . The statutes of Congress, then, must be examined to determine the limitations of tribal sovereignty rather than to determine its sources or its positive content. *What is not expressly limited remains with the domain of tribal sovereignty.*[222]

The decision in *John v. Baker* is in harmony with the federal government's view of inherent tribal sovereignty: tribal law governs unless Congress explicitly removes it.[223] The Supreme Court recognized this principle by stating that "until Congress acts, the tribes retain their existing sovereign powers. [] Indian tribes still possess those aspects of sovereignty not withdrawn by treaty or statute, or by implication as a necessary result of their dependent status."[224]

The state of Alaska also recognized this principle. In the 1988 case of *Native Village of Stevens v. Alaska Mgmt. & Planning*, the Supreme Court of Alaska found that "there is no need for the federal government expressly to recognize the sovereignty of a tribe for the tribe to retain most attributes of its sovereignty."[225] The court noted that silence by Congress on the issue of a particular aspect of sovereignty leads to the inference that the authority of the tribe remains unharmed.[226] In an earlier case, cited as an example by the majority in *Baker*, the Alaskan Supreme Court found that only a provision by Congress can subject tribes to state law.[227]

---

221. *See id.* at 1. Before European contact, there were more than 600 independent tribes and more than 5 million Indians living in the United States. *See id.* Each of these tribes used its sovereign powers and recognized those of other tribes. *See* WOLFLEY & JOHNSON, *supra* note 203, at 1. This sovereignty still exists today. *See id.*

222. COHEN, *supra* note 52, at 122 (emphasis added). Thus tribal law governs unless Congress divests.

223. *See* 982 P.2d at 751.

224. United States v. Wheeler, 435 U.S. 313, 323 (1978); *see also* Iowa Mut. Ins. Co. v. LaPlante, 480 U.S. 9, 18 (1987); Merrion v. Jicarilla Apache Tribe, 455 U.S. 130, 149 (1982); Santa Clara Pueblo v. Martinez, 436 U.S. 49, 60 (1978).

225. 757 P.2d at 47.

226. *See id.* If the Village has been recognized as a sovereign tribe and Congress neither waives nor recognizes the Village, the court is bound to recognize the inherent sovereignty. *See id.*

227. *See Baker*, 982 P.2d at 751.

Thus the sovereign power of tribes existed unless Congress explicitly removed their authority to act, and they did not do so in *Baker*.[228] Congress did not act to divest these tribes of any of their authority in the child custody area. In fact, in interpreting federal law, the Supreme Court stressed the importance of allowing Indian tribes to regulate their own internal affairs, including domestic affairs.[229] In *Fisher v. District Court*, the Supreme Court held that the exclusive jurisdiction of the tribal court in the custody dispute resulted from the tribe's inherent quasi-sovereignty rather than the race of the plaintiff.[230]

In addition, Congress' intent indicates that in the area of child custody, tribal court jurisdiction has not been removed and, in fact, has been supported. Although ICWA does not encompass interparental custody disputes, the goal of the federal policy was aimed at allowing tribal courts to exercise their powers over their young members as a way to keep their cultural heritage intact and to raise their children as they see fit.[231] As ICWA is federal policy on the treatment of Indian children, it serves as an inherent restriction on state courts in assuming jurisdiction over these cases.[232] Since tribal law governs unless Congress explicitly takes it away, the result is elementary: the tribe has the inherent sovereignty to adjudicate a child custody dispute regardless of territory.

Dissenting Justice Matthews offered what he referred to as the "allocative principle" to explain federal Indian jurisprudence on the issue of jurisdiction.[233] According to his dissent, this principle was central in the United States Supreme Court's discussion of the power struggle between states and tribes.[234] The principle contained the two questions the Court always asks to determine which law applies.[235] These were: "(1) whether the litigants and the activity giving rise to the dispute [were] based in Indian

---

228. *See id.* at 751.

229. *See* Montana v. United States, 450 U.S. 544, 564 (1981) (finding that tribes preserve their inherent authority to, among other powers, control domestic affairs).

230. 424 U.S. 382, 390 (1976). The court also found that even if an Indian plaintiff is deprived of a forum in which a non-Indian has access, the denial of access is permissible because it is in furtherance of a congressional policy of self-determination. *See id.* at 390-91.

231. *See* Atwood, *supra* note 111, at 1105.

232. *See id.* at 1082.

233. *See Baker*, 982 P.2d at 774. Matthews stated that most cases discussing tribal sovereignty and jurisdiction issues employ the allocative principle either expressly or implicitly. *See id.* He also pointed to secondary sources that recognized the allocative principle as central to Indian jurisprudence. *See id.* at 775.

234. *See id.* at 774.

235. *See id.*

country, and (2) whether there [was] an act of Congress which expressly supports the particular exercise of state or tribal power."[236] When applied to *John v. Baker*,[237] Matthews believed it was clear that the dispute was not based in "Indian country" according to the Supreme Court's decision in *Venetie II*.[238] The question then became whether the tribe had the power to exercise its inherent powers outside Indian country.[239] The dissent believed that since state law governs unless Congress grants tribes the right to exert their power, inherent sovereignty of the tribes could not be carried off Indian country.[240] The majority had asserted the opposite: tribal law governs unless Congress divests the tribes of this authority.[241]

The dissent argued that the tribal court is not allowed jurisdiction since no disputes arise in Indian country and because Congress has never explicitly said that tribes can exercise their power over their members.[242] In the absence of Indian country, the tribal court loses all realistic ability to govern.[243] Thus it would lose its sovereign existence and become merely an association without any of the attributes of a separate government.[244] This is impermissible according to the trust doctrine whereby Congress can terminate its relationship with a tribe only through a "clear and plain statement."[245]

In addition, this statement ignores the principle that has been a fundamental part of American Indian jurisprudence since the 1800s: tribes retain inherent sovereignty to practice over their own members in certain areas such as domestic relations.[246] The connection to land was a judicially created doctrine that occurred later and has only a tenuous relationship with the fundamental ideas behind the government to government relations between Indian tribes and the United States government.[247] The federal government is granted power over American Indians in the Commerce Clause of the United States Constitution.[248] Using the trust doctrine, Con-

---

236. *Baker*, 982 P.2d at 774.

237. *Id.* at 738.

238. 522 U.S. at 520.

239. *See Baker*, 982 P.2d at 776.

240. *See id.* Dissenting Justice Matthews stated as a general rule that off Indian land, tribal law does not govern unless Congress has explicitly granted tribes that authority. *See id.* at 774.

241. *See id.* at 776.

242. *See id.* at 774.

243. *See* Thompson, *supra* note 5, at 449.

244. *See id.*

245. *See Yukon Flats*, 101 F.3d at 1294.

246. *See Merrion*, 455 U.S. at 169.

247. *See* Thompson, *supra* note 5, at 450-53.

248. U.S. CONST. Art. I, § 8, cl. 3. The Constitution gives authority to Con-

gress was able to enact laws that would benefit American Indian tribes, including section 1151, which linked "Indian country" with the idea of sovereignty.[249] Although the only clause in the United States Constitution relating to American Indians is the Commerce Clause, Congress assumed power over Native American affairs which allowed it to practice federal jurisdiction over Indian lands.[250] The sovereignty of tribes is retained, not granted or protected by the Constitution.[251] No statute enacted by Congress can be the basis for denying the tribe the authority to practice its sovereignty over members of its tribe.[252]

Further, if the allocative principle was applied to *John v. Baker*, the court's result would not violate the doctrine.[253] The court's holding did not affect the exercise of state court authority over interparental child custody disputes; it merely extended jurisdiction by allowing concurrent jurisdiction with tribal courts.[254] The state courts will still be able to adjudicate disputes through one of two means: either by exercising their concurrent jurisdiction or through the use of comity.[255]

Since Northway Village was recognized as a tribe after the enactment of ANCSA, the court decided that the Alaska Native Claim Settlement Act did not divest tribes of their inherent power by extinguishing the existence

---

gress "to regulate commerce with foreign nations, and among the several States, and with the Indian tribes." *Id.*

249. *See supra* Part I.C.

250. *See* Thompson, *supra* note 5, at 451. This authority "far exceed[ed] the scope of power granted in the Commerce Clause" and caused a shift in focus from viewing Native American tribes as possessing inherent sovereign powers to "entities that are an appropriate focus of plenary power." *See id.*

251. *See Merrion*, 455 U.S. at 168.

252. *See* Thompson, *supra* note 5, at 450. The Supreme Court's decision in *Venetie II* was based on the scope of tribal authority interpreting section 1151. *See id.* at 447. The Court reasoned that if the activity did not occur on lands described in section 1151, the tribe did not have authority to adjudicate the dispute. *See id.* The Court determined the scope of the tribes' authority based on a statute that defined criminal jurisdiction rather than determining the scope of the tribes' authority and then defining the jurisdiction. *See id.*

253. *See Baker*, 982 P.2d at 761.

254. *See id.*

255. *See id.* The comity doctrine is used when full faith and credit is not available and allows the courts to give effect to the decision of another court even though they are not required to do so. *See id.* at 762. Since no state or federal court gives full faith and credit to a tribal court, comity must be utilized. *See id.* at 763. Although as a general rule state courts try to respect tribal court decisions, the state court may refuse to recognize the decision of a tribal court if the state court feels it was inappropriate, particularly if issues of due process are raised. *See id.* On remand in *Baker*, the court denied comity stating Mr. Baker's necessary due process protection was not given. *See supra* note 138 and accompanying text.

of Indian country.[256] Accordingly, since Northway Village's "status as a federally recognized tribe is undisputed and its adjudication of child custody disputes over member children is necessary 'to protect self-government or to control internal relations,' its tribal courts require no express congressional delegation of the right to determine custody of tribal children."[257]

3. Domestic Relations Has Always Been a Sovereign Power of American Indian Tribes According to Judicial Interpretation

Along with the right to determine tribal membership and to prescribe rules of inheritance for members, tribal courts retained the power to adjudicate domestic relations disorders.[258] When the United States government conquered the Indian tribes, they lost all external powers of sovereignty (e.g. the power to make treaties), but their internal sovereignty was untouched (e.g. domestic relations powers).[259] The power of the tribe included authority over internal and social relations as well as sovereign immunity from suit, and the power to prescribe and enforce laws among its own members.[260] Tribes have an interest in regulating the domestic affairs of their members as a purpose of their inherent sovereignty, concurrent with the state.[261] Even though the state may also have strong interest in the child custody proceeding as it would want to ensure the safety and best interests of children living within state boundaries, regardless of the chil-

---

256. *See Baker*, 982 P.2d at 752.

257. *Id.* at 752. The majority asserts that the Supreme Court has always examined the nature of the power that the tribe is trying to assert, not only those events that took place in Indian country. *See id.* The Supreme Court, according to the majority, follows the principle set out in *Montana* that Congress need only expressly delegate power to a tribe when that tribe tries to exercise power beyond the realm of their sovereignty. *See id.* This principle actually narrowed the Court's position on the sovereign rights of tribal powers as it had previously decided in *Wheeler* that tribes retain their sovereign powers as long as Congress has not withdrawn them. *See* Andrew S. Montgomery, *Tribal Sovereignty and Congressional Dominion: Rights-Of-Way for Gas Pipelines on Indian Reservations*, 38 STAN. L. REV. 195, 212 (1985); *see also Wheeler*, 435 U.S. at 323. The Court in *Montana v. United States* stated that tribal power only exists by an express act of Congress or when internal relations require it. *See* Montana v. U.S., 450 U.S. 544, 564 (1981); *see also* Montgomery, *supra* at n.113.

258. *See Montana*, 450 U.S. at 564 (finding that "Indian tribes retain their inherent power to determine tribal membership, to regulate domestic relations among members and to prescribe rules of inheritance for members").

259. *See* Jimenez & Song, *supra* note 48, at 1642-43.

260. *See id.* at 1642.

261. *See id.*

dren's membership in a tribe, the tribe's interest is stronger.[262] The tribe has the cultural and ethnic background of the child at heart and will make sure that the child is not cut off from his heritage.[263] There is less danger of the child being cut off from the United States government as it is the dominant culture.[264] Allowing the tribe jurisdiction is consistent with the policy of self-determinism that the government has been applying recently.[265]

"As a general principle of Indian law, Indian tribes are qualified to exercise powers of self-government by virtue of their original tribal sovereignty."[266] This principle was first set out by Chief Justice Marshall in the Cherokee cases[267] and the Supreme Court consistently followed this standard, holding that the inherent sovereignty of the tribes has continued into modern times in certain areas such as domestic relations.[268] These powers exist unless Congress explicitly divests them.[269] State courts, such as Alaska's, must not ignore this principle.[270]

Historically, the question of whether a tribal court has jurisdiction over a child custody dispute has been dependent upon an interpretation of ambiguous authorities.[271] In the area of child custody and domestic affairs, the state and federal courts have been grappling with the effects of Public Law 280, ANCSA and ICWA since they were each enacted.[272] The courts reached differing results: the state courts in Alaska have continuously found that Native Alaskan Villages did not have sovereign powers while the federal courts have found the opposite.[273] In order to decide whether a

---

262. *See* Atwood, *supra* note 111, at 1082.

263. *See id.* at 1100. Atwood stated that tribal law should apply to interparental custody disputes except when there are virtually no contacts on the reservation and state law should apply unless the parties have virtually no contact outside the reservation. *See id.* at 1082. However, those disputes that fall within the gray area are those that must be examined closely. *See id.*

264. *See id.* at 1101.

265. *See* Atwood, *supra* note 111, at 1057. Under *Fisher*, a preference towards a tribe is justified under self-determinism grounds. *See id.* at 1105.

266. Di Pietro, *supra* note 72, at 338. For this reason, it was important that Native Alaskan tribes were recognized as Indian tribes. *See id.*

267. *See supra* Part I.A. (discussing the Cherokee cases and the principle of inherent sovereignty that was recognized in these early cases).

268. *See Wheeler*, 435 U.S. at 323.

269. *See Baker*, 982 P.2d at 751.

270. *See id.* at 751.

271. *See* Atwood, *supra* note 111, at 1053-54.

272. *See id.* at 1065.

273. *See* Thompson, *supra* note 71, at 380. The court in the early cases set a precedent for finding that Native Alaskan villages did not have sovereignty by concluding that there was no Indian country in Alaska. *See id.* These cases were

cause of action could be advanced against a village in Alaska, the state courts needed to decide whether such a village is a "tribe," so that the village could receive the protections accorded to tribes under federal law, such as sovereign immunity.[274]

A review of authorities reveals that state courts in Alaska have historically found that Native Villages in Alaska did not enjoy sovereignty because they did not qualify as "Indian country."[275] First, the Native villagers did not live on reservations like the Native Americans in other states.[276] Second, the same civil and criminal law applies to Natives as to non-Natives, while federal law governed American Indians.[277] In *Native Village of Stevens v. Alaska Management & Planning*,[278] the Supreme Court of Alaska decided that the village in question did not have sovereign immunity largely based on the fact that it is "like most native groups in Alaska, . . . not self-governing or in any meaningful sense sovereign."[279]

The Supreme Court of Alaska decided more recently that state law is exclusive in the area of domestic affairs since the "reassumption" requirement in Public Law 280 required that the village petition the Secretary of the Interior before they can be granted the jurisdiction to adjudicate the domestic dispute.[280] The Ninth Circuit in *Venetie I.R.A.*[281] held that Public Law 280 did not extinguish tribal sovereignty in child custody disputes

---

heard in local federal district courts before Alaska had their own system. *See id.* After the state court system was created, the state courts found that Alaskan Native Villages were not tribes. *See id.*

274. *See* Thompson, *supra* note 71, at 380.

275. *See id.* at 381.

276. *See id.*

277. *See id.* In 1884, Congress created a civil government for Alaska by enacting the Alaska Organic Act, which adopted the laws of Oregon. *See* Alaska Organic Act, ch. 53, 23 Stat. 24 (1884); *see also* Native Village of Stevens v. Alaska Mgmt. & Planning, 757 P.2d 32, 37 (Alaska 1988). Section 7 of the Act states: "That the general laws of the State of Oregon now in force are hereby declared to be the law in said district, so far as the same may be applicable and not in conflict with the provisions of this act or the laws of the United States." Ch. 53 § 7, 23 Stat. 24, 25 (1884). Federal laws through the creation of federal enclaves, on the other hand, govern native Indians. *See* Matal, *supra* note 27, at 301.

278. 757 P.2d at 32.

279. *Id.* The majority opinion was written by Justice Matthews, the author of the dissent in *John v. Baker*. *See* 982 P.2d at 738. Matthews' majority opinion in Stevens is consistent with his dissent in *Baker*, finding that the federal government had never recognized Native Alaskan groups as sovereign tribes. *See Stevens*, 757 P.2d at 34; *Baker*, 982 P.2d at 769 (Matthews, J., dissenting). However, Matthews did not account for activities between these two cases in his analysis in *Baker*. *See* 982 P.2d at 749-50.

280. *See Nenana*, 722 P.2d at 221.

281. 944 F.2d at 548.

because Public Law 280 granted only concurrent state jurisdiction.[282] Despite this federal ruling, in *In re F.P.*,[283] the Supreme Court of Alaska reiterated its conclusion that Public Law 280 does grant the state exclusive jurisdiction in child custody matters as it was combined with ICWA.[284] With the decision in John v. Baker, the Supreme Court of Alaska followed recent federal legislation which has strengthened the authority of Native American tribes and followed federal case law which suggested that tribes retain their authority to adjudicate child custody disputes.[285]

4. Congressional Intent Supports Tribal Jurisdiction Over Child Custody Disputes

The majority of the court in *Baker* interpreted ANCSA[286] and post-ANCSA federal statutes as supporting tribal jurisdiction over child custody disputes.[287] The court noted that ANCSA did not expressly divest nor did it state as a reason for ANCSA the intent to divest tribes of sovereign

---

282. *See id.* at 562.

283. 843 P.2d at 1214.

284. *See id.* at 1215.

285. *See Baker*, 982 P.2d at 754.

286. 43 U.S.C. §§ 1601-28 (1994).

287. *See Baker*, 982 P.2d at 753-54. Matthews, in comparison, discussed the effects of ANCSA in his dissent and came to the opposite conclusion. *See id.* at 770-72 (Matthews, J., dissenting). Pointing to his historical analysis, Matthews stated that Congress and the Department of the Interior assumed that ANCSA would grant the states exclusive jurisdiction. *See id.* at 770. However, following the enactment of ANCSA, the Department of the Interior reversed their position and said that Public Law 280 granted only concurrent jurisdiction and also recognized the Native Alaskan villages as tribes. *See id.* The Supreme Court followed with its decision in *Alaska ex. rel. Yukon Flats School Dist. v. Native Village of Venetie Tribal Gov't.*, holding that ANCSA does not terminate Indian country and that the Venetie Village occupied Indian country in Alaska. *See* 101 F.3d at 1300, 1302. The United States Supreme Court's decision two years later in *Venetie II* overturned the decision in *Yukon Flats*. *See Venetie II*, 522 U.S. at 520. The Supreme Court found that, after applying a two-part test, ANCSA lands were not "Indian country" as defined by federal statute 18 U.S.C. § 1151. *See id.* at 530-32. Matthews believed that the majority opinion in *John v. Baker* went beyond the holding in *Yukon Flats*, because under *Yukon Flats* tribal jurisdiction would have been limited to the boundaries of the "Indian country." *See Baker*, 982 P.2d at 771 (Matthews, J., dissenting). Matthews then offered an analysis of cases similar to the majority, using both *Fisher v. Dist. Court* and *DeCoteau v. Dist. County Court* to arrive at the opposite conclusion: previous courts have held that state law governs unless Congress explicitly removes jurisdiction and since Congress has not provided that tribal courts should have jurisdiction over child custody matters, state law governs. *See id.* at 772 (Matthews, J., dissenting).

authority.[288] After ANCSA, federal statutes such as the Tribe List Act,[289] ICWA[290] and the Indian Tribal Justice Act[291] demonstrated Congress' intent to allow the tribes to retain their authority.[292] Finally, the court stated that federal case law supported tribal court jurisdiction.[293] The Supreme Court of Alaska recognized that no federal case law could answer the question directly because the federal courts never had reason to decide whether tribes retained their authority even in the absence of land.[294] However, the court found that cases did state that the source of tribal authority is both tribal membership and the existence of tribal land.[295] According to the Supreme Court of Alaska, these cases treated tribal membership as a dispositive factor in determining jurisdiction while the tribal land factor

---

288. *See id.* at 753 (Matthews, J., dissenting). However, the dissent maintained that ANCSA does not support tribal jurisdiction outside Indian country since it states that it is not intended to extend any new rights or obligations. *See id.* at 791-92; 43 U.S.C. § 1601(b) (1988).

289. 25 U.S.C. §§ 470(a)-479(a-1) (1994).

290. 25 U.S.C. § 1901-03 (1994). The dissent claimed that, since Congress specifically took the child custody as a result of divorce issue out of ICWA, ICWA should not be used as a basis for deciding an unrelated issue as its intent applies only to issues presented here. *See Baker*, 982 P.2d at 794 (Matthews, J., dissenting). Even if ICWA could be used here, the dissent asserted that the intent of Congress was to build in protection for parents who did not want the tribal court to adjudicate the dispute outside of "Indian country" by allowing them notice and giving the parents veto power over tribal court jurisdiction. *See id.* at 795. However, the stated congressional policy of ICWA was:

> to protect the best interests of Indian children and to promote the stability and security of Indian tribes and families by the establishment of minimum Federal standards for the removal of Indian children from their families and the placement of such children in foster or adoptive homes which will reflect the unique values of Indian culture, and by providing for assistance to Indian tribes in the operation of child and family service programs.

Indian Child Welfare Act, 25 U.S.C. § 1902 (1994). Thus Congress intended this statute for the benefit of Indian families and their culture and not for the purpose of allowing parents to take their children out of tribal court jurisdiction. *See id.*

291. 25 U.S.C. §§ 3601-31 (1994).

292. *See id.*

293. *See Baker*, 982 P.2d at 755-58 (citing as examples: *Wheeler*, 435 U.S. at 313; *Montana*, 450 U.S. at 544; Duro v. Reina, 495 U.S. 676 (1990); *Merrion*, 455 U.S. at 130; Fisher v. Dist. Court, 424 U.S. 382 (1976); Kiowa Tribe of Okla. v. Mfg. Technologies, Inc., 523 U.S. 751 (1998)). The Supreme Court of Alaska in *Baker* asserted that these decisions "support the conclusion that Native American nations may possess the authority to govern themselves even when they do not occupy Indian country." *Baker*, 982 P.2d at 759.

294. *See id.* at 754.

295. *See id.*

was not dispositive.[296] The Supreme Court of Alaska embraced a membership-based sovereignty whereby the tribe can exercise its inherent sovereign powers over its members or potential members, as is the case here. The court extended membership-based sovereignty from the criminal arena into the civil arena thereby allowing the tribal court to practice its sovereignty over its members. Thus the question of whether the court will extend the membership-based jurisdictional view into civil jurisdiction over a dispute between members (or potential members) of the tribe when the dispute occurs off reservation land remains unanswered.[297] The court allowed this jurisdiction based on the membership of the parties themselves, rather than the geographic aspect of the dispute.[298] The majority concluded that tribal courts did have the authority, based on their inherent sovereign powers, to adjudicate child custody disputes in Alaska.[299]

B. The Existence of Indian Country Has Not Always Been a Critical Factor in Granting Jurisdiction to Tribes

The dissent in *John v. Baker* asserted that the United States Supreme Court has always considered the existence of Indian country as a threshold issue in determining whether to grant tribes jurisdiction.[300] Without Indian country, tribes cannot practice sovereign power over members of the tribe or the land.[301] The dissent next disputed the majority's reliance on case law by maintaining that the "United States Supreme Court has never held, as the majority does today, that a tribe's inherent sovereignty, in and of itself, independent of Indian country, can be the basis for tribal adjudicatory authority."[302] Matthews offered *United States v. Wheeler*[303] to prove that the Supreme Court previously stated that both the inherent powers prescribing and enforcing criminal sanctions, as well as regulating domestic disputes among tribal members will not follow tribal members outside the

---

296. *See id.* at 755; *see, e.g.*, *Wheeler*, 435 U.S. at 323.

297. *See* Atwood, *supra* note 111, at 1080.

298. *See id.*

299. *See Baker*, 982 P.2d at 765. Although the United States Supreme Court in the Cherokee cases recognized the sovereignty of Indian tribes based on both their control over certain land and their power over their members, later Supreme Court decisions "have demonstrated that, although the Court still states that tribal sovereignty includes aspects of control over both members and territory, the Court is increasingly relying on tribal membership as the sole basis for tribal authority." Dussias, *supra* note 45, at 96.

300. *See Baker*, 982 P.2d at 777 (Matthews, J., dissenting).

301. *See id.* at 778.

302. *Id.* at 780.

303. 435 U.S. at 313.

confines of Indian country.[304]

Indian country is not, as asserted by the dissent, a prerequisite to finding that Native Alaskan Villages can assert their inherent sovereign powers. In order to assert sovereignty, a Native Village must prove that it is a "tribe" as federally defined and that it has jurisdiction over the dispute.[305] Sovereign tribes can assert personal jurisdiction over their members but whether they can assert jurisdiction over non-members depends on whether the court is viewing the tribe's authority from a membership-based or a geographically-based point of view.[306]

The Supreme Court traditionally viewed tribal sovereignty from two different perspectives: membership-based or geographically-based, beginning with the Cherokee cases.[307] While the Court mainly had utilized geographically-based sovereignty in the past, the Court focused on membership-based sovereignty especially in the criminal context in recent years.[308] In *United States v. Wheeler*,[309] a criminal case discussed by the dissent, the Court viewed the tribe's sovereignty as over both its members and its territory.[310] The Court did not discuss, as asserted by the dissent, "the precise extent of the sovereignty that Indian tribes have over their territory, aside from sovereignty over their own members within that territory, when tribal court criminal jurisdiction is at stake."[311]

The Court in *Venetie II* held that the land in question did not constitute "Indian country." The Court did not decide that Native Alaskan Villages

---

304. *See Baker*, 982 P.2d at 780-88 (Matthews, J., dissenting).

305. *See* Thompson, *supra* note 5, at 429; *see also supra* note 101 (stating that the federal government has recognized most Native Alaskan villages, which includes the two at issue here, as tribes).

306. *See* Thompson, *supra* note 5, at 430-31. Personal jurisdiction is the "power to compel obedience to the court's orders" as opposed to subject matter jurisdiction which is the power of the court to hear the case. *See* Atwood, *supra* note 111, at 1065.

307. *See* Dussias, *supra* note 45, at 17. A geographically-based view recognized the tribe's authority over any person who is conducting an activity on the tribes' land. *See id.* at 4. A membership-based view recognized the tribe's authority over members of the tribe without regard to where the activity was conducted. *See id.*

308. *See id.* at 17.

309. 435 U.S. at 313 (1978). The Supreme Court in *Wheeler* discussed the tribes' authority over its members as well as over its land. *See id.* at 324.

310. *See* Dussias, *supra* note 45, at 24-25. The Supreme Court did not speak of geographically-based sovereignty in those terms; rather, the Court spoke of relations. *See id.* at 25. The relations between the tribe and non-members were the external relations and the relationship between the tribe and its members were the internal relations. *See id.*

311. *Id.*

did not have inherent sovereign powers; rather, the Court said that the Village did not occupy "Indian country" and thus could not practice this sovereignty over non-Indians in their borders. Thus, they rejected the geographically-based concept of sovereignty.[312] Accordingly, the Courts' holding mandated a membership-based view of sovereignty whereby the Native Villages in Alaska only have power over tribal members.[313] The Supreme Court of Alaska's decision in *John v. Baker* is consistent with *Venetie II* as the court in *Baker* held that the Native Village did have the inherent sovereign power to adjudicate the internal relations of its own members.

The State of Alaska is in the unique position of having within its borders Native Alaskan Villages which are federally recognized as tribes but which do not occupy "Indian country" under federal statutes.[314] The Supreme Court never stated that tribes could assert tribal sovereignty in the absence of Indian country because they have not yet had the opportunity to decide this issue.[315] However, the Court's previous decisions do not preclude such a finding and indeed support the idea that tribes have the inherent sovereignty to adjudicate internal affairs, such as child custody.[316]

C. The Native Alaskan Tribal Court has Sovereign Power Over a Non-Tribal Member

The dissent in *John v. Baker* argued that even if the tribal court did have inherent sovereign powers to decide cases outside the confines of Indian country, the majority wrongly decided this case because the court exercised their power over a non-tribal member.[317] The court granted jurisdiction not over the membership of Ms. John and Mr. Baker but over the membership, or potential membership, of the children.[318] The tribal court

---

312. *See* Thompson, *supra* note 5, at 449.

313. *See id.*

314. *See* John v. Baker, 982 P.2d 738, 754 (Alaska 1999) *cert. denied*, 528 U.S. 1182 (2000).

315. *See id.*

316. *See id.* The Alaskan Supreme Court cited *Kiowa Tribe of Oklahoma v. Manufacturing Technologies, Inc.* as the most recent example of the Supreme Court's willingness to allow Indian tribes jurisdiction over a dispute independent of whether the activity in question occurred in "Indian country." *Kiowa Tribe*, 523 U.S. at 751; *see also Baker*, 982 P.2d at 758.

317. *See Baker*, 982 P.2d at 796. Ms. John is not a member of Northway Village. *See id.* Hence, Mr. Baker argued, the tribe's authority to resolve the domestic disputes of members does not extend to Ms. John and she should not be bound by the decision of the tribal court. *See id.*

318. *See id.* at 759.

"should have jurisdiction over the child and the child's parents or other custodians."[319] If Congress decided to support the exercise of jurisdiction over non-members, this would be constitutionally suspect because it would deny access to the state courts.[320] Matthews merely mentioned that this would create a potentially suspect racial classification that would have to be struck down as discrimination.[321]

The Supreme Court of Alaska was correct in allowing the jurisdiction of the tribal court over the children in *John v. Baker* as they are eligible for membership with Northway Village and because both parents appear to have consented to the jurisdiction. The majority focuses on the potential membership of the children as a basis for allowing the tribe to maintain their inherent sovereignty.[322] Federal statutes such as ICWA follow this method which requires that the child be a member of the tribe or eligible for membership.[323] Ms. John was not a member of Northway Village, the village where Mr. Baker first filed a petition to gain custody of his children.[324] However, her children were eligible for membership there and thus the tribe could assert jurisdiction over them.[325]

An alternate foundation of jurisdiction is the consent of both parents.[326] Mr. Baker filed the petition with this tribal court for custody and Ms. John appears to have consented to Northway Villages' jurisdiction since she argued that the Village has the inherent sovereignty to decide the issue of her children's custody.[327] Thus, there is no issue in this case about the exercise of jurisdiction over a non-member, namely Ms. John.

In addition, if the exercise of jurisdiction over a non-member Indian were at issue here, a suspect racial classification would not result. The Supreme Court in *Fisher v. District Court*[328] rejected an argument similar to this by stating that:

---

319. *Id.* at 796.

320. *See id.* at 798.

321. *See id.* at 798. Although this racial discrimination is not at issue in *Baker*, as Mr. Baker consented to the tribe's jurisdiction, Matthews mentions that a denial of state court access violates equal protection rights. *See* John v. Baker, 982 P.2d 738, 798 (Alaska 1999). However, the majority in *Baker* did not deny access to state court; the court merely granted concurrent jurisdiction. *See id.* at 761.

322. *See id.* at 759 (stating that a tribe's ability to regulate internal disputes is dependent on the membership or potential membership of the children in question).

323. *See id.*

324. *See id.* at 743, 759.

325. *See* John v. Baker, 982 P.2d 738, 759 (1999).

326. *See id.* at 759 n.141.

327. *See id.*

328. 424 U.S. 382 (1976).

> [t]he exclusive jurisdiction of the Tribal Court does not derive from the race of the plaintiff but rather from the quasi-sovereign status of the Northern Cheyenne Tribe under federal law. Moreover, even if a jurisdictional holding occasionally results in denying an Indian plaintiff a forum to which a non-Indian has access, such disparate treatment of the Indian is justified because it is intended to benefit . . . Indian self-government.[329]

In *Fisher*,[330] the Court recognized exclusive jurisdiction of the tribal court over the state court in an adoption proceeding by Indian parents of an Indian child.[331] In this case, the jurisdiction recognized by the Supreme Court of Alaska is not exclusive but rather concurrent with the state.[332] Alaska Natives who wish to have their dispute settled in an Alaskan state court could do so and thus this classification will not keep Native Alaskans from seeking justice in an Alaskan state court.[333]

D. Tribal Court Rules Bind Non-Members

Dissenting Justice Matthews was worried about the effects of discriminatory rules enforced by tribal courts that would be intolerable in a state court.[334] He stated that tribal courts are not bound by the United States Constitution and the equal protection components of the Constitution are not applied to tribal courts as they are in state courts.[335] However, as Matthews admitted, the Indian Civil Rights Act imposes many of the protections in the Bill of Rights on tribal governments.[336] Also, citizens of foreign states who are unwilling to be subject to United States courts are not asked whether these courts guarantee them the fundamental rights of their home country.[337] If subject to United States laws, these citizens are treated

---

329. *Id.* at 390-91, (citing Morton v. Mancari, 417 U.S. 535, 551-55 (1974)).

330. *Id.* at 382.

331. *See id.* at 389.

332. *See Baker*, 982 P.2d at 761.

333. *See id.*

334. *See id.* at 804.

335. *See id.*

336. *See id.* at 804 n.298. The Indian Civil Rights Act was enacted in order to protect individual Native Americans from arbitrary decisions by the tribal courts. *See* Christina D. Ferguson, Martinez v. Santa Clara Pueblo: *A Modern Day Lesson on Tribal Sovereignty*, 46 ARK. L. REV. 275, 287 (1993). Most of the rights guaranteed in the Bill of Rights were included in the Indian Civil Rights Act except where Congress refrained from interfering with the tribes' culture and autonomy. *See id.* at 287 n.87.

337. *See* C.E. Willoughby, *Native American Sovereignty Takes a Back Seat to the "Pig in the Parlor": The Redefining of Tribal Sovereignty in Traditional Property Law Terms*, 19 S. ILL. U. L.J. 593, 620 (1995).

as U.S. citizens in the court system.[338] Similarly, if tribal court jurisdiction is administered over a non-member who is subject to the tribe's inherent sovereignty through the membership of their children, they must submit to the tribal court system.[339]

> E. Although Not Bound by Executive and Legislative Authority on Tribal Jurisdiction, Advocacy by Executive and Legislative Branches Support the Villages' Right to Jurisdiction in Child Custody Cases.

Tribal status is a non-justiciable political question and courts must adhere to the decisions of the executive branch and Congress.[340] Thus, if either recognizes a village as a sovereign tribe, the courts must do the same.[341] Although not bound by the two branches' determinations on the extent of power that tribes should be allowed to practice, their advocacy did support the Supreme Court of Alaska's decision.[342] The State of Alaska and the executive branch of the United States government both advocated this extension of jurisdiction.[343] Alaska Governor Tony Knowles and Alaska Attorney General Bruce Botelho issued a press release in May 1998 stating that they were in support of tribal sovereignty over child custody issues.[344] Attorney General Botelho filed an Amicus Curiae with the Supreme Court of Alaska in this case, arguing that state courts have concurrent, rather than exclusive jurisdiction, over child custody cases.[345]

The executive branch, through the Bureau of Indian Affairs, issued a list of federally recognized tribes which included 200 Native Alaskan Villages

---

338. *See id.*

339. *See id.*

340. *See Baker*, 982 P.2d at 753; *see also* Christopher A. Ford, *Executive Prerogatives in Federal Indian Jurisprudence: The Constitutional Law of Tribal Recognition*, 73 DENV. U. L. REV. 141, 163 (1995).

341. *See Baker*, 982 P.2d at 753.

342. *See id.*

343. *See id.* at 745. All amici advocated the Supreme Court of Alaska recognizing tribal jurisdiction. *See id.*

344. *See* Press Release, Governor Tony Knowles, May 4, 1998, *State Backs Expanded Tribal Role in Domestic Matters: Envisions Broader Tribal Powers in Adoption, Custody Cases between Members*, *available at* http://www.gov.state.ak.us/press/pr050498a.html (last visited Feb. 25, 2001). Botelho stated that the state "seek[s] to empower tribal governments and tribal courts to deal with issues that affect the lives of their members. Gov. Tony Knowles and I agree that self government and self determination are an essential part of addressing many of the social issues that exist today . . . ." *Id.*

345. *See id.*

in 1993.[346] Congress followed in 1994 with the Federally Recognized Tribe List Act of 1994.[347] Recently, the executive branch has shown its support for Indian self-determination. President Clinton, meeting with tribal leaders at the White House in 1994, said: "Today I reaffirm our commitment to self-determination for tribal governments."[348] In addition, the Department of Justice offered arguments in favor of an extension of tribal court jurisdiction.[349]

V. POLICY CONSIDERATIONS FAVOR CONCURRENT STATE AND FEDERAL JURISDICTION

Although the court did not base its decision on public policy, it cannot ignore the important role of policy in its decisions, as they affect so many citizens of the State of Alaska. The majority stated that although their opinion is based on precedent and the controlling authority of Congress, policy considerations support their decision.[350] The Court asserted that the best interests of Native American children are served by concurrent jurisdiction.[351] By acknowledging the power of Native American tribes to regulate internal affairs such as domestic relations and allowing parties to take their claims to state courts as well, the majority maintained that justice could be achieved.[352]

However, according to the dissent, the policy reasons stated by the majority as a basis for their holding were an inappropriate foundation for extending jurisdiction.[353] The majority's rationale for its decision was that allowing tribal governments to decide domestic disputes among themselves would help to promote justice for all.[354] They recognized that some members of the tribes lived far from the state courts or had cultural and

---

346. *See* Summit, *supra* note 3, at 627.

347. *See id.* at 628.

348. *Id.* at 626-27.

349. *See Baker*, 982 P.2d at 791. The dissent argued that the intent of the executive branch and the intent of Congress did not grant tribes the authority of sovereign power. *See id.* at 790. The dissent admitted that the executive branch (specifically, the Department of Justice's arguments in regard to this case) advocated an extension of tribal court jurisdiction, but asserts that the Supreme Court has reached different results from the executive branch on these issues. *See id.* Although the court is bound to follow the executive branch's recognition of certain groups as tribes, the dissent asserted that the court is not bound by the executive branch's advocacy for an extension of power. *See id.*

350. *See id.* at 760.

351. *See* John v. Baker, 982 P.2d 738, 760 (1999).

352. *See id.*

353. *See id.* at 799 (Matthews, J., dissenting).

354. *See id.* at 760.

language needs that state courts could not address.[355] This policy statement was suspect, according to the dissent, as it "raise[d] fundamental issues about judicial restraint, separation of powers, and the role of the judiciary."[356]

Although Congress is the appropriate branch to make policy that affects Native Alaskans, the Supreme Court of Alaska has the authority to interpret statutes and previous decisions of state and federal courts.[357] The majority of the court did not base its decision on policy. However, the justices believed that policy considerations supported their decision.[358] Policy issues dictate that tribes should have sovereign powers to adjudicate domestic disputes even in the absence of "Indian country."[359]

First, Native Alaskans face many obstacles in gaining access to the state court system in Alaska.[360] The members of the villages are often confronted with problems of language, geographic accessibility to the courts and a lack of resources.[361] According to a report by the Alaska Supreme Court Advisory Committee on Fairness and Access in 1997, the committee learned of instances of "unintentional bias, cultural misunderstandings, inadequate services and lack of accessibility."[362] Twenty-nine percent of Alaskans suffer from geographic isolation and do not have easy access to the court system.[363] In addition, approximately eleven percent of the Alas-

---

355. *See id.* Most Alaska Natives live in small villages reachable only by plane or boat and have been comparatively inaccessible for many centuries. *See* Summit, *supra* note 3, at 621. Native Alaskans can be divided into three categories: Aleuts, Eskimos and Indians (who can be further divided into Tlingit, Haida, and Athabascan) who are all lumped into the category of "Indian" for inclusion in federal programs. *See id.* at 608. These Native Alaskan groups are "genetically and geographically distinct." *Id.*

356. *Baker*, 982 P.2d at 799 (Matthews, J., dissenting). The Supreme Court of Alaska makes its decision based on policy, reason and precedent. *See id.* at 744.

357. *See id.* at 744.

358. *See id.* at 760.

359. *See id.*

360. *See id.*

361. *See Alaska Supreme Court Report: Fairness and Access Problems and Recommendations*, ALASKA JUST. F. (Univ. of Alaska Anchorage, Anchorage, AK), Fall 1997, at 1 [hereinafter *Alaska Supreme Court Report*].

362. *Id.*

363. *See id.* at 4. Compare with Justice Matthews' dissent, arguing that since more Native Alaskans live within driving distance of cities with state courts than do not and live far from their Native Villages, the majority of the Supreme Court of Alaska creates the problem they seek to resolve. *See* John v. Baker, 982 P.2d 738, 804 (Alaska 1999). For example, one-fifth of resident Native Alaskans live in Anchorage. *See id.* However, this argument does not bear up to scrutiny since the court is finding concurrent jurisdiction, not exclusive jurisdiction to the tribal court. *See id.* at 759-60.

kan population do not speak English and studies conducted "reveal problems with the provision of due process and equal protection when parties in a court procedure are not fluent in English."[364] Concurrent jurisdiction increases the chances that Native Alaskans will be able to receive justice by allowing a choice between state and tribal courts.[365] Parents will be able to bring suit in a state court if they have easy access or to a tribal court where their cultural and geographical needs can be met.[366]

Children, in particular, will be advantaged by the use of concurrent jurisdiction. Congress recognized that the issues of child custody from adoption or foster care placement should be litigated in tribal courts as the survival of tribes rests on their ability to raise their own children.[367] The Supreme Court of Alaska previously decided that the state of Alaska had exclusive jurisdiction over child custody cases.[368] The confusion resulting from different interpretations of child custody determinations left Indian communities "disrupted and unsettled, delayed placement of children, and diverted limited resources from needed services."[369] Indian children have been ripped apart from their culture and background through the enactment of assimilation acts, and through state and federal encroachments into Indian territory.[370] The child's best interest is served by allowing their tribe to deal with the cultural problems that are created by state courts over-

---

364. *Alaska Supreme Court*, *supra* note 361, at 4.

365. *See Baker*, 982 P.2d at 760.

366. *See id.* at 761.

367. *See* Michael E. Connelly, *Tribal Jurisdiction Under Section 1911(B) of the Indian Child Welfare Act of 1878: Are The States Respecting Indian Sovereignty?*, 23 N.M.L. REV. 479, 479 (1993). However, Congress explicitly did not include custody disputes between parents. *See* 25 U.S.C. § 1903(1) (1994). The jurisdiction granted by Congress when they enacted ICWA even extended to children living beyond the reservation. *See* Connelly, *supra* at 479. Although intended to grant exclusive jurisdiction to tribal courts on these matters, states have applied the good cause exception to ICWA to retain jurisdiction, thus creating de facto concurrent jurisdiction. *See id.* at 483. Congress specifically mentioned as a justification for ICWA that states have not historically recognized the "essential tribal relations of Indian people and the cultural and social standards prevailing in Indian communities and families." 25 U.S.C. § 1901(5) (1994). As a result, state courts have acted contrary to the intent of Congress by invading the cultural arena of a tribe and taking jurisdiction.

368. *See* In Re F.P., W.M. and A.M., 843 P.2d 1214, 1215-16 (1992).

369. Press Release, Governor Tony Knowles (May 4, 1998) *State Backs Expanded Tribal Role in Domestic Matters: Envisions Broader Tribal Powers in Adoption, Custody Cases between Members*, *at* http://www.gov.state.ak.us/press/pr050498a.html (last visited Feb. 25, 2001).

370. *See* B.J. Jones, *The Indian Child Welfare Act: In Search of a Federal Forum to Vindicate the Rights of Indian Tribes and Children Against the Vagaries of State Courts*, 73 N.D. L. REV. 395, 456 (1997).

reaching into areas of exclusive Indian control.[371] Although the problems that have been created by the federal and state governments should in reality be reversed through the use of exclusive tribal jurisdiction, the use of concurrent jurisdiction is a compromise that should quell fears about the imbalance between tribal and state courts.

VI. REVIEW OF SUPREME COURT OF ALASKA'S DECISION AND SUBSEQUENT RULINGS ON THIS CASE

In sum, the majority of the Supreme Court of Alaska held that the tribal court had the inherent sovereignty to decide this dispute and that the tribal court's jurisdiction is concurrent with that of the state of Alaska.[372] The majority reversed and remanded for the superior court to decide if the tribal courts' decision should be recognized under the principle of comity, which allows the court of one state to give effect to the decision of another court.[373] If comity was recognized, the tribal courts' exercise of jurisdiction would be upheld and the couple would share custody of their children. On remand, the superior court found that the comity doctrine does not apply since Mr. Baker was denied his Due Process rights to a fair tribunal.[374] Therefore, the state court retained jurisdiction and Mr. Baker continued to enjoy sole custody.

VII. CONCLUSION

The Supreme Court of Alaska allowed tribal courts to assert their jurisdictional powers over child custody disputes of their own members in *John v. Baker*.[375] This was not an extension of power for the tribes; rather, it is a product of their inherent sovereign adjudicatory authority that existed before the federal government even came into existence.[376] As Native Alaskan Villages are federally recognized tribes and as they did retain their powers over domestic relations, this decision was correct.[377] The Alaskan Supreme Court's recognition of concurrent jurisdiction between the tribal court and state court will help to alleviate some of the problems encoun-

---

371. *See id.*

372. *See Baker*, 982 P.2d at 765.

373. *See id.*

374. *See* Baker v. John, No. 4FA-95-3103 (order denying comity), *available at* http://www.alaska.net/~akctlib/4fa95310.txt. (Alaska Sup. Ct. Sept. 2000) (last visited Nov. 17, 2000)

375. *See* 982 P.2d at 749.

376. *See* Wolfley & Johnson, *supra* note 204, at 1.

377. *See supra* Part IV.

tered by Native American people who have been living in a jurisdictional labyrinth while maintaining a balance of power between state and tribal governments.

Alison DeCoste Bancroft

Ensuring Justice Prevails in the Wake of *Theresa Canavan's Case*: A Proposal for Reform

"While judges must determine and articulate a resolution, it is lawyers who frame the issues and marshal the evidence."[1]

I. INTRODUCTION

A dramatic change in the allocation of power between judge and jury marked the close of the 20th century as a result of a series of decisions rendered by the United States Supreme Court, and subsequently adopted by the Supreme Judicial Court of Massachusetts ("SJC"), on the issue of expert testimony. The Supreme Court's entry into the "battle of the experts"[2] began with *Daubert v. Merrell Dow Pharmaceuticals, Inc.*,[3] which was the first in a trilogy of cases that have entirely transformed the judge's

---

1. Chief Justice Margaret H. Marshall of the Supreme Judicial Court, Keynote Address at the Annual Meeting of the Massachusetts Bar Association (Jan. 22, 2000) *in* ANNUAL REPORT ON THE STATE OF THE MASSACHUSETTS COURT SYSTEM, Fiscal Year 1999, at 10.

2. Expert witnesses are an essential, arguably controversial, part of the justice system, and their involvement is not expected to decrease in the foreseeable future. *See* Eric G. Jensen, *When "Hired Guns" Backfire: The Witness Immunity Doctrine and the Negligent Expert Witness*, 62 UMKC L. REV. 185, 185 (1993). Many cases boil down to a battle of the experts where the winning party is the one with the more persuasive expert or "hired gun." *See* Randall K. Hanson, *Witness Immunity Under Attack: Disarming "Hired Guns,"* 31 WAKE FOREST L. REV. 497, 497 (1996).

> Witnesses have played a part in the common law courts for centuries. Initially witnesses were friends of the court who came forward to assist in the attempt to achieve justice and who were granted immunity from liability so as to encourage their open participation and assistance. Much has changed. There are now experts for hire on every imaginable topic, and even ordinary lawsuits seem to involve experts who testify in support of the party who hires them.

Id.

3. 509 U.S. 579 (1993).

role in determining the admissibility of expert testimony. As a result, "courts may effectively take any case which depends upon expert testimony away from the jury."[4] Historically, it was only in those rare cases, where expert testimony was based on truly novel theories, that judges excluded expert testimony under the *Frye* standard.[5] The *Frye* standard, a fairly vigorous test, required that an expert prove that his or her conclusions were based on principles generally accepted in the relevant scientific community.[6] This standard was used for most of the 20th century.[7] It was not until 1993, in *Daubert,* that the United States Supreme Court announced that *Frye* had been superseded by the Federal Rules of Evidence, which were enacted some twenty years earlier.[8] The *Daubert* decision was startling and it sent shock waves through the legal profession.[9] In place of the *Frye* test, the Court substituted a two-part standard requiring trial judges to act as "gatekeepers" to ensure that "any and all scientific testimony or evidence admitted is not only relevant, but reliable."[10] By implementing this new gatekeeping requirement, *Daubert* began the trend of diminishing the role of the jury. The result has been greater judicial involvement in the screening of expert testimony, thereby empowering trial judges.[11]

Subsequently, the Supreme Court decided the final two cases in the trilogy which continued the trend: *General Electric v. Joiner*[12] and *Kumho Tire v. Carmichael.*[13] In *Joiner*, the Court determined that abuse of discretion was the proper standard for appellate review of judicial rulings on the admissibility of expert testimony.[14] Applying the abuse of discretion standard increases judicial discretion over jury discretion by giving height-

---

4. Richard Collin Mangrum, *Kumho Tire Company: The Expansion of the Court's Role in Screening Every Aspect of Every Expert's Testimony at Every Stage of the Proceedings*, 33 CREIGHTON L. REV. 525, 525 (2000).

5. *See infra* notes 33-43 and accompanying text for a discussion of the *Frye* case.

6. *See infra* notes 33-43 and accompanying text.

7. The *Frye* standard was used for approximately seventy years. *See infra* notes 33-43 and accompanying text.

8. The Federal Rules of Evidence were enacted in 1974. *See infra* notes 44-47 and accompanying text.

9. *See* Daniel J. Capra, *The Daubert Puzzle*, 32 GA. L. REV. 699, 758 (1998) (suggesting that "[t]he shock waves from *Daubert* have reached the halls of Congress").

10. *Daubert*, 509 U.S. at 589.

11. *See infra* notes 44-59 and accompanying text.

12. 522 U.S. 136 (1997).

13. 526 U.S. 137 (1999).

14. *See infra* notes 60-68 and accompanying text for a discussion of the *Joiner* case.

ened deference to the decisions made by trial court judges to admit or exclude expert testimony.[15] In *Kumho Tire*, the Court expanded the judge's role as gatekeeper beyond scientific testimony to *all* expert testimony.[16]

This trilogy of cases, *Daubert*, *Joiner*, and *Kumho Tire*, has severely shifted the allocation of power in favor of the judge over the jury despite the Court's original intent to "liberalize" the admissibility of expert testimony.[17] The continuation of this trend threatens the very existence of the jury system. Concurrent with the shift in allocation of power between judge and jury at the close of the 20th century has been widespread criticism of the jury system.[18] Without caution, this shift of power combined with dissatisfaction with the jury system may be the first step toward abolishing the jury system, or at the very least altering it beyond recognition.

Massachusetts has systematically followed the Supreme Court with regard to admissibility of expert testimony. In fact, the SJC has adopted the principles set forth in the Supreme Court's trilogy almost verbatim.[19] In *Commonwealth v. Lanigan*,[20] the Supreme Judicial Court adopted, in part, the holding in *Daubert*.[21] And just recently, in *Theresa Canavan's Case*,[22] the SJC, while weighing in on the hotly debated issue of multiple chemical sensitivity ("MCS"), essentially adopted the principles set out in *Joiner* and *Kumho Tire*.[23] In doing so, the SJC clarified the law with respect to the admissibility of expert testimony in Massachusetts courts by

---

15. *See infra* notes 60-68 and accompanying text.

16. *See infra* notes 69-79 and accompanying text for a discussion of the *Kumho Tire* case.

17. The Court's intent to liberalize the admissibility of expert testimony is evidenced by its decision to abandon exclusive reliance on *Frye* in favor of *Daubert's* four-prong test. Although some deny the Court's intent was to liberalize admissibility of evidence, in *Joiner* the Court itself specifically recognized *Daubert's* effect. "Thus, while the Federal Rules of Evidence allow district courts to admit a somewhat broader range of scientific testimony than would have been admissible under *Frye*, they leave in place the 'gatekeeper' role of the trial judge in screening such evidence." *Joiner*, 522 U.S. at 142.

18. *See, e.g.*, Mark Curriden, *The Power of 12*, 87 A.B.A. J. 36 (Aug. 2001), which states that "[j]urors increasingly are sending loud messages of censure with megabuck verdicts . . . [b]ut critics charge that a jury is the least qualified to decide public policy." *Id.* at 37.

19. *See infra* notes 80-162 and accompanying text for a discussion of the admissibility of expert testimony in Massachusetts courts.

20. 641 N.E.2d 1342 (Mass. 1994).

21. The SJC refused to completely abandon the general acceptance test. *See infra* notes 86-96 and accompanying text.

22. 733 N.E.2d 1042 (Mass. 2000).

23. *See infra* notes 134-62 and accompanying text.

announcing two new rules.[24] First, the SJC adopted an abuse of discretion standard as the standard of review for judicial rulings on the admissibility of expert testimony, consistent with the Supreme Court's opinion in *Joiner*.[25] Furthermore, as the Supreme Court had done in *Kumho Tire*, the SJC in *Canavan* expanded the judge's role as gatekeeper so that all expert testimony is subject to judicial scrutiny.[26]

This comment attempts to answer two questions. How did we get here? Where do we go from here? Part II discusses the origins of the debate over the admissibility of expert testimony and discusses in detail the cases involved in the Supreme Court's trilogy.[27] Additionally, Part II discusses the admissibility of expert testimony in Massachusetts, highlighting those cases in which the SJC has adopted the principles set out by the Supreme Court.[28] Part III presents *Canavan's Case*, the SJC's most recent notable decision on the issue, including facts, procedural background, and the court's analysis.[29] Part IV addresses the future of admissibility of expert testimony in Massachusetts. It attempts to answer the question: Where do we go from here? Part IV begins with a discussion of issues left unresolved by *Canavan*, and concludes by offering suggestions on how to handle unresolved issues.[30] Part V examines the role of the judiciary, the Massachusetts bar, and contemporary society in shaping the future of the jury system.[31] An advocation for jury reform to preserve the jury system as our founding fathers envisioned it is also included in Part V. The suggested reform stresses the need to increase the efficiency of juries and to halt the continuing trend toward increased judicial power.[32]

II. BACKGROUND

A. The Origins of the Debate Over the Admissibility of Expert Testimony

For most of the 20th century, the standard for admissibility of scientific expert testimony was rigorous – expert testimony was admissible only if the proponent could show that the proffered evidence was a generally

---

24. *See infra* notes 142-44 and accompanying text.

25. *See Canavan's Case*, 733 N.E.2d at 1048-49. *See also infra* notes 144-53 and accompanying text.

26. *See infra* notes 154-62 and accompanying text.

27. *See infra* notes 33-79 and accompanying text.

28. *See infra* notes 80-96 and accompanying text.

29. *See infra* notes 97-162 and accompanying text.

30. *See infra* notes 163-278 and accompanying text.

31. *See infra* notes 206-78 and accompanying text.

32. *See infra* notes 206-78 and accompanying text.

accepted scientific principle.[33] This requirement became known as the *Frye* test, named after the case from which it was first enunciated.[34] The goal of the *Frye* test was to keep unreliable or "junk" evidence from jurors.[35] The most often quoted language from the *Frye* decision summarizes the *Frye* test succinctly:

> [j]ust when a scientific principle or discovery crosses the line between the

---

33. *See Daubert*, 509 U.S. at 585-86. This standard was first enunciated in 1923 and remained "good law" until 1993.

34. *See* Frye v. United States, 293 F. 1013 (D.C. Cir. 1923). The defendant, Frye, was convicted of murder in the second degree. *See id.* At trial, his attorney attempted to introduce expert testimony regarding the results of a systolic blood pressure deception test. *See id.* The test was similar to a present-day lie detector test. *See id; see also* J. Ken Thompson, Comment, *A Review of the Admissibility of Novel Scientific Evidence*, 17 AM. J. TRIAL ADVOC. 741, 744 (1994) (maintaining that the systolic blood pressure deception test in *Frye* was the "forerunner to the modern polygraph test"). In *Frye*, the court refused to admit expert testimony regarding the results of the blood pressure deception test because the test had "not yet gained such standing and scientific recognition among physiological and psychological authorities as would justify" its admission. *Frye*, 293 F. at 1014. Interestingly, *Frye* was only a two page decision containing no citations. *See id.* at 1013-14; *see also Daubert*, 509 U.S. at 585.

35. *See* Fred Galves, *Where the Not-so-Wild Things Are: Computers in the Courtroom, The Federal Rules of Evidence, and the Need for Institutional Reform and More Judicial Acceptance*, 13 HARV. J.L. & TECH. 161, 254 (2000).

> *Frye's* 'generally accepted' requirement was criticized, however, because although it was designed to keep out non-established, or 'junk' science, it also would keep out new, innovative scientific techniques, sometimes for many years, until they became established (or 'generally accepted'), necessarily placing courts one step behind society and denying them the use of helpful, cutting-edge scientific evidence.

Id. (citations omitted) The term "junk science" became a household word following release of Peter Huber's book entitled Galileo's Revenge: Junk Science in the Courtroom. According to Huber, "[j]unk science is the mirror image of real science, with much of the same form but none of the same substance." PETER HUBER, GALILEO'S REVENGE: JUNK SCIENCE IN THE COURTROOM 2 (1991). Huber categorically classifies some professions as junk science and suggests the presence of junk science in almost every aspect of our lives.

> Junk science cuts across chemistry and pharmacology, medicine and engineering. It is a hodgepodge of biased data, spurious inference, and logical legerdemain, patched together by researchers whose enthusiasm for discovery and diagnosis far outstrips their skill. It is a catalog of every conceivable kind of error: data dredging, wishful thinking, truculent dogmatism, and, now and again, outright fraud.

Id. at 3.

> experimental and demonstrable stages is difficult to define. Somewhere in this twilight zone the evidential force of the principle must be recognized, and while courts will go a long way in admitting expert testimony deduced from a well-recognized scientific principle or discovery, the thing from which the deduction is made must be sufficiently established to have gained general acceptance in the particular field in which it belongs.[36]

The *Frye* test was the dominant standard for determining the admissibility of scientific evidence for approximately seventy years.[37] The standard had been a source of controversy since its inception.[38] The major criticisms were that *Frye* was too restrictive,[39] it produced inconsistent results,[40] and it usurped the role of the jury.[41] It was not until the adoption of the Federal Rules of Evidence that the validity of the *Frye* test was eventually challenged.[42] The United States Supreme Court addressed the continuing validity of the *Frye* test in light of the adoption of the Federal Rules of Evidence in 1993, in its seminal decision, *Daubert v. Merrell*

---

36. *Frye*, 293 F. at 1014.

37. *See Daubert*, 509 U.S. at 585. Not all states utilized the *Frye* test. By the late 1970s, approximately forty-five states followed *Frye*. *See* Thompson, *supra* note 34, at 747. But, by 1994, only twenty-three states still followed *Frye*. *See id.* Massachusetts was one of the twenty-three states, the others were: Alabama, Alaska, Arizona, California, Colorado, Connecticut, Florida, Illinois, Indiana, Kansas, Kentucky, Louisiana, Maryland, Michigan, Minnesota, Mississippi, Missouri, New Hampshire, New Jersey, Oklahoma, Pennsylvania and South Dakota. *See id.*

38. *See* Thompson, *supra* note 34, at 747.

39. *See* Paul C. Giannelli, *The Admissibility of Novel Scientific Evidence: Frye v. United States, a Half-Century Later*, 80 COLUM. L. REV. 1197, 1223-24 (1980). "For example, the general acceptance standard has been criticized for excluding reliable evidence. Paradoxically, the standard also may permit the admission of unreliable evidence." *Id.* at 1223. Massachusetts has also recognized the restrictive nature of *Frye*. *See, e.g.*, Commonwealth v. Lanigan, 641 N.E.2d 1342, 1348 (Mass. 1994) (recognizing "the risk that reliable evidence might be kept from the factfinder by strict adherence to the *Frye* test"); Commonwealth v. Curnin, 565 N.E.2d 440, 443 n.8 (Mass. 1991) (suggesting that "strict adherence to the *Frye* rule may keep reliable probative evidence from the fact finder simply because the relevant scientific community has not yet digested and approved of its foundation"); *see also* Commonwealth v. Mendes, 547 N.E.2d 35, 41 (Mass. 1989) (Liacos, C.J., dissenting) (recognizing that in some situations, due to the changing nature of the profession, the relevant scientific community may not have had the opportunity to adequately analyze a scientific foundation).

40. *See* Giannelli, *supra* note 39, at 1219-21.

41. *See* Thompson, *supra* note 34, at 746-47. Critics of the *Frye* test argue that juries are capable of evaluating scientific evidence without judicial involvement. *See id.* and examples cited.

42. *See* Thompson, *supra* note 34, at 748-49. *See also Daubert*, 509 U.S. at 587.

Dow Pharmaceuticals, Inc.[43]

B. The Supreme Court's Trilogy on Admissibility of Expert Testimony

1. *Daubert v. Merrell Dow Pharmaceuticals*: The United States Supreme Court Provides the Foundation for Admissibility of Scientific Expert Testimony in Light of the Federal Rules of Evidence

Although the Federal Rules of Evidence were codified in 1974,[44] it was not until 1993, in the seminal *Daubert* case, that the continuing validity of the *Frye* test was challenged under the Federal Rules of Evidence.[45] In *Daubert*, the Supreme Court concluded that the Federal Rules of Evidence supersede the *Frye* test.[46] The Court also defined the appropriate analysis for scrutinizing the admissibility of scientific expert testimony in

---

43. 509 U.S. 579 (1993). *See infra* notes 44-59 and accompanying text.

44. *See* Clifford A. Knaggs, *The Admissibility of Evidence and Expert Testimony Based on Science, Technology or Other Specialized Knowledge—Is the* 'Frye' *Standard Consistent with the Federal Rules of Evidence?*, 4 T.M. COOLEY L. REV. 641 (1987). The Federal Rules of Evidence are codified at 28 U.S.C. § 702 (2000).

45. *See Daubert*, 509 U.S. at 585-87. *Daubert* involved a suit by two couples, on behalf of their sons, against a manufacturer to recover for birth defects sustained as a result of ingestion of the drug, Bendectin, during pregnancy. *See id.* at 582. The case was essentially a "battle of the experts" whose opinions differed on the side effects of Bendectin. The plaintiffs offered expert testimony by eight experts that Bendectin caused birth defects. *See id.* at 583. The district court granted Merrell Dow's motion for summary judgement using the *Frye* test, as did the Ninth Circuit Court of Appeals because the plaintiff's expert's testimony was based on a theory known as reanalysis, which had not been subject to peer review. *See id.* at 583-84. *See also* Daubert v. Merrell Dow Pharmaceutical, Inc., 727 F. Supp. 570, 572 (S.D. Cal. 1989), *aff'd* 951 F.2d 1128 (9th Cir. 1991). The Supreme Court granted the plaintiffs' petition for certiorari. *See id.* at 585. *See also* Daubert v. Merrell Dow Pharmaceuticals, Inc., 506 U.S. 914 (1992).

46. *See Daubert*, 509 U.S. at 587-89. "In principle, under the Federal Rules no common law of evidence remains." *Id.* at 588. However, in reality, common law fills in the gaps where the Federal Rules are lacking. *See id.* Because there is a specific Federal Rule which addresses the issue of the admissibility of expert testimony, the Federal Rule governs. *See id.* "Nothing in the test of [Rule 702] establishes 'general acceptance' as an absolute prerequisite to admissibility. Nor does respondent present any clear indication that Rule 702 or the Rules as a whole were intended to incorporate a 'general acceptance' standard." *Id.* As a result, the Court held that the general acceptance test set out in *Frye* is "at odds with the 'liberal thrust' of the Federal Rules and their 'general approach of relaxing the traditional barriers to opinion testimony,'" therefore *Frye* is incompatible with the Federal Rules of Evidence. *Id.* at 588 (citing Beech Aircraft Corp. v. Rainey, 488 U.S. 153, 169 (1988)).

light of the Federal Rules of Evidence ("FRE"), particularly FRE 702.[47]

Apparently, in an attempt to provide lower courts with guidance, the Court set out a framework for analyzing scientific expert testimony.[48] Under this framework, consistent with the Federal Rules of Evidence, the judge is to act as a "gatekeeper," ensuring that all expert scientific testimony is both relevant and reliable.[49] This gatekeeper function requires that a judge's analysis focus solely on principles and methodology rather than the conclusions that the expert generates.[50] The Court set out four non-exclusive factors to be used in determining admissibility: testability,[51] peer review and publication,[52] rate of error,[53] and general acceptance in the scientific community.[54] The Court cautioned that these factors were not exclusive, rather, they were suggested standards by which a trial judge

---

47. FED. R. EVID. 702. According to Rule 702, "[i]f scientific, technical, or other specialized knowledge will assist the trier of fact to understand the evidence or to determine a fact in issue, a witness qualified as an expert by knowledge, skill, experience, training, or education, may testify thereto in the form of an opinion or otherwise." *Id.*

48. *See Daubert*, 509 U.S. at 593-95.

49. *See id.* at 594-95. The focus on relevance is derived from FRE 402 which provides that "[a]ll relevant evidence is admissible, except as otherwise provided by the Constitution of the United States, by Act of Congress, by these rules, or by other rules prescribed by the Supreme Court pursuant to statutory authority. Evidence which is not relevant is not admissible." FED. R. EVID. 402. Relevant evidence is defined as evidence that has "any tendency to make the existence of any fact that is of consequence to the determination of the action more probable or less probable than it would be without the evidence." FED. R. EVID. 401. The Supreme Court interpreted this definition of relevance liberally. *See Daubert*, 509 U.S. at 587. Similarly, the focus on reliability is derived from FRE 702. *See id.* at 588. Recognizing that there are uncertainties in science, the Court concluded that "[p]roposed testimony must be supported by appropriate validation—i.e., 'good grounds,' based on what is known. In short, the requirement that an expert's testimony pertain to 'scientific knowledge' establishes a standard of evidentiary reliability." *Id.* at 590.

50. *See Daubert*, 509 U.S. at 595.

51. *See id.* at 593. "Scientific methodology today is based on generating hypotheses and testing them to see if they can be falsified; indeed, this methodology is what distinguishes science from other fields of human inquiry." *Id.* (citations omitted).

52. *See id.* Although not fool-proof "submission to the scrutiny of the scientific community is a component of 'good science,' in part because it increases the likelihood that substantive flaws in methodology will be detected." *Id.* (citation omitted).

53. *See id.* at 594. This includes the known, or potential, rate of error. *See id.*

54. *See id.* General acceptance should be a factor to consider, but the Court cautions that a "reliability assessment does not require, although it does permit, explicit identification of a relevant scientific community and an express determination of a particular degree of acceptance within that community." *Id.*

should evaluate scientific testimony in light of the Federal Rules of Evidence.[55] The Court emphasized that "[t]he inquiry envisioned by Rule 702 is . . . a flexible one."[56]

It did not take long before the *Daubert* framework, like the *Frye* test, became the subject of intense debate. In fact, much of the criticism aimed at the new *Daubert* framework was the very same as that previously aimed at the *Frye* test.[57] One such criticism, a topic of intense debate, is that "*Daubert* was a wrong turn because it presumed that judges could know better than scientists whether the scientists' opinions are reliable."[58] Further debate focused on issues unanswered by *Daubert*, for example: the standard of appellate review of a trial judge's decision to admit or deny testimony and whether *Daubert* applied only to "scientific" evidence. Despite the continued controversy, most agree that *Daubert* was intended to expand, rather than limit, admissibility of expert testimony.[59]

---

55. *See Daubert*, 509 U.S. at 592-93. "Many factors will bear on the inquiry, and we do not presume to set out a definitive checklist or test." *Id.* at 593.

56. *Id.* at 594.

57. *See* Thompson, *supra* note 34, at 751. *See also supra* notes 39-41 and accompanying text for a discussion of the criticisms of *Frye*.

58. Michael H. Gottesman, *From* Barefoot *to* Daubert *to* Joiner: *Triple Play or Double Error?*, 40 ARIZ. L. REV. 753, 775 (1998).

> The Court [in *Daubert*] did not suggest that it is the judge's function to decide whether the expert honestly believes in the opinion he or she is proffering. Rather, *Daubert* assumed that an expert might *really believe* the proffered testimony but authorized judges nonetheless to exclude the testimony as unreliable. This approach necessarily assumes that a judge is better equipped than an honestly-testifying expert to know whether the expert's opinion is reliable. That is an unlikely premise. To be sure, the judge will have heard the views of experts who disagree, but the testifying expert is just as aware of those competing views and, with far greater expertise than the judge, has concluded that the contrary opinions are unconvincing.

Id. at 758-59. Gottesman theorizes that *Daubert* gives judges the freedom to exclude expert testimony where the judge questions the credibility of the expert rather than the scientific nature of the proffered testimony; "[*Daubert*] facilitates the exclusion of opinions the judge believes are not honestly held by the expert proffering them." *Id.* at 759.

59. *See* David L. Faigman, *The Law's Scientific Revolution: Reflections and Ruminations on the Law's Use of Experts in Year Seven of the Revolution*, 57 WASH. & LEE L. REV. 661, 663-64 (2000). In support of this theory, most commentators point to the language in *Daubert* itself in which the Court observed that "[the] rigid 'general acceptance' requirement would be at odds with the 'liberal thrust' of the Federal Rules." *Id.* at 664 (citing *Daubert*, 509 U.S. at 588).

2. *General Electric Co. v. Joiner*:[60] Defining the Standard of Appellate Review of a Trial Judge's Decision to Admit or Deny Expert Testimony. The Supreme Court Attempts to Address the Issues Left Unanswered by *Daubert*.

Since the *Daubert* opinion in 1993, "*Daubert* already has been cited in some 2,000 published opinions, with nearly half the states adopting its gatekeeper doctrine."[61] Not surprisingly, there have been innumerable appeals of decisions involving admissibility of expert testimony under *Daubert*. Consequently, courts have struggled with the standard of appellate review to be applied in these cases. In *Joiner*, the second case in the trilogy, the Supreme Court granted certiorari to "determine what standard an appellate court should apply in reviewing a trial court's decision to admit or exclude expert testimony under [*Daubert*]."[62] All of the Justices joined Chief Justice Rehnquist in holding that abuse of discretion is the proper standard for appellate review of a lower court's evidentiary ruling.[63] Furthermore, abuse of discretion is the proper standard regardless of whether a lower court's ruling allowed or excluded expert testimony.[64]

---

60. 522 U.S. 136 (1997). The facts of *Joiner* are as follows: As part of his employment, the plaintiff had to immerse his hands and arms into the coolant of electrical transformers. *See id.* at 139. Occasionally, the coolant would splash into his eyes and mouth. *See id.* The coolant contained polychlorinated biphenyls (PCBs), which are hazardous to human health. *See id.* The plaintiff developed small-cell lung cancer and sued the manufacturer of the coolant along with the manufacturer of the PCBs. *See id.* The district court granted summary judgment in favor of the defendants on the grounds that the testimony of the plaintiff's experts "did not rise above 'subjective belief or unsupported speculation,'" thus it was inadmissible. *See id.* at 140. The Court of Appeals for the Eleventh Circuit reversed, applying a "particularly stringent standard of [appellate] review to the trial judge's exclusion of expert testimony" in light of the preference for admissibility under the Federal Rules of Evidence. *Id.* The Supreme Court granted certiorari. *See id.* at 141. *See also* General Electric Co. v. Joiner, 520 U.S. 1114 (1997).

61. William C. Smith, *No Escape from Science*, 86 A.B.A. J. 60, 62 (Aug. 2000).

62. Joiner, 522 U.S. at 138-39.

63. *See id* at 141. The Court was unequivocal on this issue. "We have held that abuse of discretion is the proper standard of review of a district court's evidentiary rulings." *Id.* Apparently, the Court of Appeals suggested that *Daubert* altered the standard of review of scientific evidence, however the Supreme Court rejected this idea. *See id.* at 142.

64. *See id.* at 141-43. In *Joiner*, the Appeals Court apparently applied a more stringent abuse of discretion standard on the theory that its decision was outcome determinative. The Supreme Court responded that "[a] court of appeals applying 'abuse of discretion' review . . . may not categorically distinguish between rulings allowing expert testimony and rulings disallowing it." *Id.* at 142. The Court also rejected the "respondent's argument that because the granting of summary judgment in

The Court then did what it had not done in *Daubert*, it examined the record and ruled on the admissibility of the expert testimony.[65]

Joiner seems to contemplate that courts should apply the abuse of discretion standard to review the admissibility of expert testimony in the same manner the standard is applied to other evidentiary issues.[66] Consequently, appellate courts may give considerable deference to trial court rulings on the admissibility of expert testimony.[67] The result is that appellate courts will have a limited role in deciding whether to admit or exclude expert testimony and the bulk of the power remains in the hands of the trial judge.[68]

3. *Kumho Tire v. Carmichael*:[69] Does *Daubert* Apply to All Evidence or Only Scientific Evidence?

In the post-*Daubert* era, lower courts struggled with interpreting *Daubert*. In particular, whether the *Daubert* factors applied to non-scientific expert testimony.[70] Since *Daubert* did not address this issue, lower court judges were left to make their own conclusions as to the appli-

---

this case was 'outcome determinative,' it should have been subjected to a more searching standard of review." *Id.* at 142-43.

65. *See id.* at 146-47. Instead of remanding the case, the Court concluded that it was "within the District Court's discretion to conclude that the studies upon which the experts relied were not sufficient, whether individually or in combination, to support their conclusions that Joiner's exposure to PCBs contributed to his cancer," therefore the District Court did not abuse its discretion. *Id.*

66. *See id* at 141-42. *See also* Randolph N. Jonakait, *The Standard of Appellate Review for Scientific Evidence: Beyond* Joiner *and* Scheffer, 32 U.C. DAVIS L. REV. 289, 294 (1999).

67. *See* Jonakait, *supra* note 66, at 294-95. The result will be that "appellate courts will seldom overrule trial rulings on the admissibility of scientific evidence." *Id.* at 295.

68. *See* Jonakait, *supra* note 66, at 294-95. Appellate courts will be limited to determining whether trial judges acted correctly as gatekeepers without having the ability to set rules that will apply to future cases dealing with the same subject matter. *See id.* at 295.

69. 526 U.S. 137 (1999).

70. Examples of non-scientific testimony include the testimony of engineers or accountants. The struggle by lower courts is evidenced by their split on the issue of whether testimony by an engineer is subject to the gatekeeping requirements of *Daubert*. *See, e.g.*, Vadala v. Teledyne Indus., Inc., 44 F.3d 36, 39 (1st Cir. 1995) (applying *Daubert* to testimony offered by a mechanical engineer); *compare with* Freeman v. Case Corp., 118 F.3d 1011, 1016 n.6 (4th Cir. 1997) (holding that *Daubert* does not apply to expert testimony of a mechanical engineer because the expert relied on his own training and experience as the basis of his testimony, rather than any particular methodology).

cability of *Daubert* to non-scientific evidence.[71] Less than two years after *Joiner*, the Supreme Court decided the third case in the trilogy when it granted certiorari to determine whether a judge's gatekeeper obligation under *Daubert* applied only to scientific evidence or whether it extended to 'technical' and 'other specialized' knowledge.[72] Specifically, the Supreme Court was asked to address whether *Daubert* applies to testimony of engineers and those who are not scientists.[73]

The Court held that the *Daubert* "gatekeeping" obligation applies to all expert testimony, not only scientific testimony.[74] Furthermore, the Court held that trial courts may consider *Daubert* factors to the extent relevant, recognizing that "the test of reliability is 'flexible,' and *Daubert*'s list of specific factors neither necessarily nor exclusively applies to all

---

71. *See* Derek L. Mogck, *Are We There Yet?: Refining the Test for Expert Testimony Through* Daubert, Kumho Tire *and Proposed Federal Rule of Evidence 702*, 33 CONN. L. REV. 303, 308 (2000). "Indeed, among the eleven circuit courts of appeals that addressed the question, five courts concluded that *Daubert*'s factors applied to non-scientific testimony while six courts concluded that they did not." *Id.* Those that did apply *Daubert* to non-scientific evidence include: Watkins v. Telsmith, Inc., 121 F.3d 984, 991 (5th Cir. 1997); Surace v. Caterpillar, Inc., 111 F.3d 1039, 1056 (3d Cir. 1997); Smelser v. Norfolk S. Ry. Co., 105 F.3d 299, 303 (6th Cir. 1997); Peitzmeier v. Hennessy Indus., Inc., 97 F.3d 293, 296-97 (8th Cir. 1996); and Cummins v. Lyle Indus., 93 F.3d 362, 367 n.2 (7th Cir. 1996). Those that chose not to apply *Daubert* to non-scientific evidence include: Desrosiers v. Flight Int'l, Inc., 156 F.3d 952, 960 (9th Cir. 1998), Carmichael v. Samyang Tire, Inc., 131 F.3d 1433, 1436 (11th Cir. 1997); Freeman v. Case Corp., 118 F.3d 1011, 1016 n.6 (4th Cir. 1997); Stagl v. Delta Airlines, Inc., 117 F.3d 76, 81 (2d Cir. 1997); Bogosian v. Mercedes-Benz, Inc., 104 F.3d 472, 479 (1st Cir. 1997); and Compton v. Subaru, Inc., 82 F.3d 1513, 1518 (10th Cir. 1996).

72. *See Kumho Tire*, 526 U.S. at 141. "This case requires us to decide how *Daubert* applies to the testimony of engineers and other experts who are not scientists." *Id. Kumho Tire* was a products liability case against a tire manufacturer and distributor alleging manufacturing and design defects following an accident in which a tire on the plaintiffs' minivan blew out. *See id.* at 142. The plaintiffs offered the testimony of a tire failure analyst whose opinion was based upon a visual and tactile inspection of the tire. *See id.* at 142-47. The defendants sought to exclude the testimony proffered by the plaintiffs' expert arguing it was insufficient under FRE 702. *See id.* at 145. The District Court, applying *Daubert*, excluded the testimony of the plaintiffs' expert and granted the defendant's motion for summary judgment. *See id.* at 146. The Eleventh Circuit Court of Appeals reversed the district court's decision on the grounds that *Daubert* did not apply to non-scientific expert testimony such as the skill or experience-based testimony proffered by the plaintiff's expert. *See id.* at 146. *See also* Carmichael, 131 F.3d at 1435-37, *cert. granted sub nom.* Kumho Tire Co. v. Carmichael, 526 U.S. 137 (1999).

73. *See Kumho Tire*, 526 U.S. at 141.

74. *See id.*

experts or in every case."[75] The Court uniformly rejected the Eleventh Circuit's distinction between *Daubert*'s application to expert testimony based on scientific principles and expert testimony based on personal observation or clinical experience.[76] As a result of the Court's decision in *Kumho Tire*, trial judges must employ their "gatekeeper" function in assessing the admissibility of all expert testimony. However, this does not mean that a trial judge must always employ the specific factors set out in *Daubert*, i.e., testability, peer review and publication, rate of error and general acceptance.[77] The *Daubert* factors are not a definitive checklist, rather the *Daubert* inquiry is a flexible one and the specific factors may be considered if applicable.[78] The Court acknowledged the trial judges' role as a gatekeeper in evaluating all expert testimony, but it did not endorse the strict application of the *Daubert* factors because, as Justice Scalia notes in his concurring opinion, the *Daubert* factors are not a "holy writ."[79]

In terms of the delicate balance of power between judge and jury, the *Kumho Tire* opinion, the last case in the Supreme Court's trilogy, significantly tipped the scales in favor of judicial discretion over jury discretion in evaluating expert testimony by awarding almost unfettered discretion to trial judges to determine the credibility and admissibility of expert testimony. Massachusetts followed suit.

---

75. *See id.* The Court also held that the trial court did not abuse its discretion in applying the *Daubert* factors to exclude the testimony of the plaintiff's expert. *See id.* at 158. As a result of the Supreme Court's decision, in federal court, all expert testimony is subject to a *Daubert* analysis to ensure that such testimony is both relevant and reliable. *See id.*

76. *See id.* at 148.

> [I]t would prove difficult, if not impossible, for judges to administer evidentiary rules under which a gatekeeping obligation depended upon a distinction between "scientific" knowledge and "technical" or "other specialized" knowledge. There is no clear line that divides the one from the others. Disciplines such as engineering rest upon scientific knowledge. Pure scientific theory itself may depend for its development upon observation and properly engineered machinery. And conceptual efforts to distinguish the two are unlikely to produce clear legal lines capable of application in particular cases.

Id.

77. *See id.* at 145. *See also supra* notes 51-55 and accompanying text for a discussion of the *Daubert* factors.

78. The Court stressed the flexibility of *Daubert*, recognizing that in some cases all of the *Daubert* factors may be applicable while in others only some of the factors may apply. *See Kumho Tire*, 526 U.S. at 141.

79. *See id.* at 158-59 (Scalia, J., concurring).

C. The Admissibility of Expert Testimony in Massachusetts

Since the law of evidence in Massachusetts is based on common law and no comprehensive code of evidence has been adopted,[80] Massachusetts courts remain free to develop common law principles relating to the admissibility of evidence.[81] During the last ten years, Massachusetts' courts have embraced this opportunity, particularly with respect to the admissibility of expert testimony, resulting in a myriad of new rules.[82] Most of these rules have been patterned after those prescribed by the United States Supreme Court.[83] For example, just as *Daubert* dramatically changed the environment in federal courts with respect to admissibility of expert testimony, *Lanigan* dramatically changed the environment in Massachusetts state courts.[84] Likewise, just as *Daubert* signaled the end of exclusive reliance on the *Frye* test in determining admissibility of expert testimony in federal courts, *Lanigan* signaled the end of exclusive reliance on the *Frye* test in state courts.[85]

---

80. *See* 19 HON. WILLIAM G. YOUNG ET AL., MASSACHUSETTS PRACTICE § 101.1 (2d ed. 1998). In Massachusetts, the common law has been "overlaid by an extraordinary variety of statutory provisions enacted over many years, most designed to accomplish a particular public good without regard to a comprehensive evidentiary framework." *Id.* Although the law of evidence has not been codified to date, Massachusetts gave great consideration to the issue in the late 1970's and early 1980's. "On November 22, 1976, the Supreme Judicial Court . . . appointed an Advisory Committee to consider whether the Massachusetts rules of evidence should be codified or promulgated. In July of 1980 the Advisory Committee transmitted proposed Rules of Evidence to the Court." 20 HON. WILLIAM G. YOUNG ET AL., MASSACHUSETTS PRACTICE app. D (2d ed. 1998). The court received briefs, and oral arguments were heard. *See id.* A majority of the justices decided not to adopt the proposed rules. *See id.* The justices concluded that promulgation of the proposed rules would restrict the development of the common law, that the proposed rules would not necessarily result in uniformity of Federal and State practice, and the Federal Rules of Evidence have not resulted in uniform results. *See id.* Despite some distinctions, the proposed rules track the Federal Rules of Evidence. Although the court declined to adopt the proposed rules, the court announced that "[p]arties are invited to cite the Proposed Rules, wherever appropriate, in briefs and memoranda submitted." *Id. See also* Grant v. Lewis/Boyle, Inc., 557 N.E.2d 1136, 1139 n.3 (Mass. 1990).

81. *See* YOUNG ET AL., *supra* note 80, at app. D.

82. *See infra* notes 86-96 for a discussion of the *Lanigan* case and notes 97-162 and accompanying text for a discussion of the *Canavan* case.

83. *Compare infra* notes 86-96 and accompanying text *with supra* notes 44-79 and accompanying text.

84. *See infra* notes 86-96 and accompanying text.

85. *See infra* notes 86-96 and accompanying text.

1. *Commonwealth v. Lanigan*: The SJC Adopts, in Part, the *Daubert* Holding

In *Commonwealth v. Lanigan*, the SJC was asked to evaluate the admissibility of expert testimony regarding DNA evidence.[86] In so doing, the SJC took the opportunity to analyze the applicability of *Daubert* under Massachusetts evidentiary law.[87] The SJC chose to adopt, in part, the holding of *Daubert*.[88] Although Massachusetts historically relied on the *Frye* test to determine the admissibility of scientific expert testimony, the court concluded that the ultimate test was the reliability of the expert's theory or process.[89] Recognizing "the risk that reliable evidence might be kept from the factfinder by strict adherence to the *Frye* test" the SJC, at the urging of the Commonwealth, adopted the *Daubert* reasoning.[90] The SJC adopted "the basic reasoning of the *Daubert* opinion because it [was] consistent with [Massachusetts'] test of demonstrated reliability."[91] The SJC accepted the principle that "a proponent of scientific opinion evidence may demonstrate the reliability or validity of the underlying theory or process by other means, that is, without establishing general acceptance [in the scientific community]."[92] Accordingly, a trial judge must determine "whether the reasoning or methodology underlying the testimony is scientifically valid" and "whether that reasoning or methodology properly can be applied to the facts in issue."[93] This test is virtually identical to the *Daubert* test.[94]

However, the SJC did not unquestionably adopt the *Daubert* holding. Rather, the SJC cautioned that "[w]e suspect that general acceptance in the relevant scientific community will continue to be the significant, and often

---

86. In *Lanigan*, the defendant appealed his conviction for rape of a child and indecent assault and battery on three minors on the grounds that expert testimony on behalf of the Commonwealth regarding DNA evidence was improperly admitted because the expert's testimony, which was based on the ceiling principle to prove the probability of a DNA match, had been the subject of criticism in the scientific community. *See Lanigan*, 641 N.E.2d at 1347. After abrogating strict adherence to *Frye* in favor of the more liberal standard in *Daubert*, the SJC held that expert testimony regarding the ceiling principle is admissible under *Daubert* and "[u]nanimity of opinion among the relevant scientists is not essential even under the general acceptance test." *Id.* at 1349.

87. *See Lanigan*, 641 N.E.2d at 1348. The court chose to adopt *Daubert* at the Commonwealth's urging. *See id.*

88. *See supra* notes 44-59 for a discussion of the *Daubert* holding.

89. *See Lanigan*, 641 N.E.2d at 1348.

90. *See id.*

91. *Id.* at 1349.

92. *Id.*

93. *Id.* (citation omitted)

94. *See supra* notes 44-59 for a discussion of *Daubert*.

the only, issue."[95] Essentially, the standard for admissibility of scientific expert testimony set out in *Lanigan*, although fundamentally analogous to *Daubert*, may be distinguished from *Daubert* primarily by Massachusetts' continued reliance on the general acceptance test.[96]

III. THERESA CANAVAN'S CASE: THE SJC ADOPTS THE PRINCIPLES SET FORTH IN *JOINER* AND *KUMHO TIRE*

A. The Employee-Employer Relationship

Theresa Canavan, a nurse, had been employed by Brigham and Women's Hospital since 1983.[97] In January of 1990, she began working in the Hospital's operating room.[98] "[W]hile in the operating room, she was subjected to various chemicals including ethylene oxide, formaldehyde, and diesel fuel."[99] On August 6, 1993, at the end of a 10-hour shift, Theresa Canavan felt ill, she complained of "a severe headache, nasal congestion, and dizziness. She returned to work on August 9, 1993, but still suffered symptoms, including a fever, headache, and swelling of her nose and right cheek."[100] Theresa was diagnosed and treated for chronic sinusitis.[101] It was determined that she was disabled; Brigham and Women's Hospital accepted her disability and paid her workers' compensation benefits.[102] Theresa was treated with antibiotic therapy for her chronic sinusitis

---

95. *Id.*

96. *See Lanigan*, 641 N.E.2d at 1350. Although the SJC adopted the principles set forth in *Daubert*, it determined that general acceptance may still be the most significant, and often the only, issue for the trial judge to decide. *See id.*

97. *See* Canavan's Case, 733 N.E.2d 1042, 1045 (Mass. 2000). *See also* http://www.brighamandwomens.org (last visited on August 27, 2001) for information regarding the Hospital. Brigham and Women's Hospital is located in Boston, Massachusetts, is an affiliate of Harvard Medical School and is recognized internationally. *See id.*

98. *See Canavan's Case*, 733 N.E.2d at 1045. "As an operating room nurse, she was responsible for the safe care of patients during surgery, including preparing the room for surgery, caring for the surgical instruments, and assisting the surgeons." *Id.* From 1983 until January of 1990, Theresa worked in the Hospital's recovery room. *See id.*

99. *Id.*

100. *Id.*

101. *See id.* She was treated with antibiotics by Arthur Laurentano, M.D. *See id.*

102. *See id.* The Hospital accepted Theresa Canavan's medical condition. *See id.* This acceptance became one of the issues on appeal. *See id.* at 1046-47. Theresa Canavan claimed, among other things, that Brigham and Women's Hospital was barred by the doctrine of judicial estoppel from denying the existence of her multiple chemical sensitivity (MCS) as a compensable disease under the Workers' compensation Act. *See id.* In *Niles-Robinson v. Brigham and Women's Hospital, Inc.*, 711

which proved ineffective and, as a result, she was referred to a second physician, N. Thomas LaCava, M.D., who eventually became her treating physician.[103] Dr. LaCava is certified in pediatrics by the American Board of Medical Specialties and is certified in Environmental Medicine by the American Board of Environmental Medicine.[104] As part of her diagnosis and treatment, Dr. LaCava took an extensive history from Theresa, performed a physical exam and conducted diagnostic tests.[105] "Dr. LaCava concluded that Theresa suffered from arthritis, paresthesias, organic brain syndrome, chemical induced headaches, immunodeficiency, and multiple chemical sensitivities (MCS) secondary to chemical poisoning, which [Dr. LaCava] believed was caused by exposure during her employment at the hospital."[106] He prescribed intravenous vitamins, vitamin C, nutritional

---

N.E.2d 940 (Mass.App.Ct. 1999), Theresa Canavan was one of approximately fifteen plaintiffs who filed suit against Brigham and Women's Hospital for tort damages for multiple chemical sensitivity. *See id.* at 941-42. Brigham and Women's Hospital, in its motion to dismiss pursuant to Mass.R.Civ.P. 12(b)(6), argued that Theresa Canavan, and the others, did not have a cause of action due to the exclusivity provisions of the Workers' compensation Act. *See id.* at 942. *See also Canavan*, 733 N.E.2d at 1046-47. Essentially, Theresa Canavan argued, in her appeal to the SJC, that the Hospital's 12(b)(6) motion in the *Niles-Robinson* case was an admission that she, and the other employees, suffered from multiple chemical sensitivity. *See Canavan*, 733 N.E.2d at 1045. Therefore, they argued, the Hospital should now be barred from asserting that Theresa did not suffer from multiple chemical sensitivity. *See id.* The SJC rejected this argument. The court determined that Brigham and Women's Hospital was required to assume that the employees had acquired MCS while at the workplace when it filed its 12(b)(6) motion, and that "[b]y filing a motion to dismiss, the hospital was not admitting that the employee acquired MCS at work; it was only stating that, even if she had acquired MCS at work, she could not maintain an action for common-law tort." *Id.* Essentially, the Hospital's 12(b)(6) motion did not amount to an admission of liability or causation. *See id.*

103. *See Canavan's Case*, 733 N.E.2d at 1045. Although I hesitate to include the names of specific doctors involved, this case was essentially a battle of the experts and it would be extremely difficult to follow the discussion without using specific names.

> Dr. LaCava is a private practitioner, an instructor in pediatrics at the University of Massachusetts Medical School, a staff pediatrician at Holden District Hospital, St. Vincent Hospital, Worcester Hahnemann Hospital, and The Memorial Hospital, and is on the medical staff at University of Massachusetts Hospital.

Id.

104. *See id.* Environmental medicine is not recognized by the American Board of Medical Specialties.

105. *See id.* at 1045. Dr. LaCava formed his opinion based on the results of the medical history, physical exam and diagnostic tests. *See id.*

106. *Id.*

supplements, antibiotics and heat therapy.[107]

B. The Workers' Compensation Hearing

Brigham and Women's Hospital challenged Dr. LaCava's diagnosis and treatment of Theresa Canavan. On April 13, 1995, Theresa Canavan's case was heard before a single judge at the Department of Industrial Accidents (DIA).[108] The issues before the DIA included: whether Theresa was unable to work, whether her disease was causally related to her employment, and whether her treatment was reasonable and necessary.[109]

Dr. LaCava's deposition testimony was offered on Theresa Canavan's behalf at her hearing.[110] Dr. LaCava testified that Theresa's condition was causally related to her work environment at Brigham and Women's Hospital, and that her treatment was reasonable and necessary.[111] Dr. LaCava defined MCS as "a systemic reaction of the body with multiple symptoms to multiple kinds of chemicals, which may be chemically unrelated, which are commonly present in the every day working and living environment."[112]

Brigham and Women's Hospital offered the deposition testimony of Donald D. Accetta, M.D.[113]

> Dr. Accetta is certified by the American Board of Allergy and Immunology. He is a private practitioner, a consultant in allergy at New England Medical Center, the secretary of the New England Society of Allergy, and has served on the board of directors for both the New England Society of Allergy and the Massachusetts Allergy Society.[114]

The testimony proffered by Dr. Accetta directly contradicted that of

---

107. *See id.*

108. The case was heard by the Honorable John C. Bradford. *See* Canavan v. Brigham and Women's Hospital and Brigham and Women's Hospital as Self-Insurer, Case No. 03332093, 22 Feb. 1996, MA Industrial Accident Board.

109. *See id.* at 4. The three issues in controversy were:

> A. Is the employee's diagnosis of multiple chemical sensitivity disease causally related?
> B. Is the treatment of this diagnosed disease reasonable and necessary?
> C. Assuming a causal relationship, is the employee incapacitated and to what extent?

Id.

110. *See id.* at 7-9; *see also Canavan's Case*, 733 N.E.2d at 1045.

111. *See Canavan's Case*, 733 N.E.2d at 1045.

112. *See id.*

113. *See id.* at 1046.

114. *Id.*

Dr. LaCava.[115] Dr. Accetta examined Theresa Canavan twice and testified that her condition was not a result of chemical exposure in the workplace.[116] Dr. Accetta believed Theresa suffered from "chronic nonallergic rhinitis caused by nonspecific stimuli that exist in the every day environment . . . [and] that [Theresa's] symptoms have a psychogenic component."[117] He also testified that MCS "is not accepted as a diagnostic disease by mainstream allergists/immunologists and occupational medicine physicians."[118]

The workers' compensation judge adopted the opinion of Dr. LaCava over that of Dr. Accetta, and concluded that Theresa's condition arose out of her employment and that intravenous vitamin C treatment and antibiotic therapy were reasonable and necessary.[119] The Industrial Accident Reviewing Board reviewed and affirmed the decision.[120] Brigham and Women's Hospital appealed the Board's decision directly to the Appeals Court pursuant to G.L. c. 152, § 12(2).[121]

C. The Appeals Courts' Decision

The principal issue presented to the Appeals Court was "the admission in evidence of the opinions of [Dr. LaCava] on diagnosis, disability, and causation."[122] Brigham and Women's Hospital argued that Dr. LaCava's testimony should not have been admitted under the principles set forth in *Daubert* and *Lanigan*.[123]

At the outset the court recognized its ability to hear the case, noting that the rules of evidence which are applicable in Massachusetts Courts are also applicable in cases before the Department of Industrial Accidents.[124] Furthermore, decisions rendered by the Reviewing Board of the Depart-

---

115. *See id.*

116. *See id.*

117. *Canavan's Case*, 733 N.E.2d at 1046. *See also* Canavan v. Brigham and Women's Hospital and Brigham and Women's Hospital as Self-Insurer, Case No. 03332093, 22 Feb. 1996, MA Industrial Accident Board, at 9-10. In layman's terms, she had a chronic runny nose and psychological problems.

118. *Canavan's Case*, 733 N.E.2d at 1046.

119. *See* Canavan v. Brigham and Women's Hospital and Brigham and Women's Hospital as Self-Insurer, Case No. 03332093, 22 Feb. 1996, MA Industrial Accident Board at 13.

120. *See* Canavan's Case, 720 N.E.2d 43, 44 (Mass. App. Ct. 1999).

121. *See Canavan's Case*, 733 N.E.2d at 1044. *See also* MASS. GEN. LAWS ch.152, § 12 (2) (2000) which states: "[a]ny appeal from a decision by a reviewing board shall . . . be filed with the appeals court of the commonwealth." *Id.*

122. *Canavan's Case*, 720 N.E.2d 43, 44 (Mass. App. Ct. 1999).

123. *See id. See also supra* notes 44-59 and 86-96 for a discussion of the *Daubert* and *Lanigan* cases respectively.

124. *See Canavan's Case*, 720 N.E.2d at 45.

ment of Industrial Accidents brought pursuant to G.L. c. 152, § 12(2) are subject to the provisions set out in G.L. c. 30A, § 14(7)(a)-(d), (f), and (g),[125] which define the scope of appellate review.

After establishing its authority to hear the case, the Appeals Court held that it was within the Administrative Law Judges' authority to admit Dr. LaCava's deposition testimony.[126] As a foundation for its holding, the court concluded that "certain expert testimony based on personal observations, clinical experience, or generally accepted scientific techniques need not be subject to the *Lanigan* analysis."[127] In the alternative, the court held that it is well established that a treating physician can testify to a patient's condition,[128] and that an adequate foundation was laid for admission of Dr. LaCava's testimony regarding diagnosis.[129] Therefore, it was not error for the workers' compensation judge to admit Dr. LaCava's causation testimony. [130] The testimony, which was based on knowledge that other pa-

---

125. *See* Case of Scheffler, 419 Mass. 251, 257-58 (1994). MASS. GEN. LAWS ch. 30A, § 14(7)(a)-(d), (f), and (g) provide:

> The court may affirm the decision of the agency, or remand the matter for further proceedings before the agency; or the court may set aside or modify the decision, or compel any action unlawfully withheld or unreasonably delayed, if it determines that the substantial rights of any party may have been prejudiced because the agency decision is—
>
> (a) In violation of constitutional provisions; or
>
> (b) In excess of the statutory authority or jurisdiction of the agency; or
>
> (c) Based upon an error of law; or
>
> (d) Made upon unlawful procedure; or ...
>
> Unwarranted by facts found by the court on the record as submitted or as amplified under paragraph (6) of this section, in those instances where the court is constitutionally required to make independent findings of fact; or
>
> Arbitrary or capricious, an abuse of discretion, or otherwise not in accordance with law.
>
> The court shall make the foregoing determinations upon consideration of the entire record, or such portions of the record as may be cited by the parties. The court shall give due weight to the experience, technical competence, and specialized knowledge of the agency, as well as to the discretionary authority conferred upon it.

Id.

126. *See Canavan's Case*, 720 N.E.2d at 46.

127. *See id.* This was contrary to the Supreme Court's holding in Kumho Tire v. Carmichael, 526 U.S. 137 (1998).

128. *See Canavan's Case,* 720 N.E.2d at 46. "It is well established that a treating physician may testify to a patient's 'ailments, bodily condition, and extent to which a person was affected [by them].'" *Id.* (quoting Kramer v. John Hancock Mut. Life Ins. Co., 336 Mass. 465, 467 (1957)).

129. *See Canavan's Case*, 720 N.E.2d at 47.

130. *See id.* at 47-48. "[T]he judge did not err in admitting Dr. LaCava's opinion on causation and adopting it over that of the conflicting testimony of Dr. Acetta." *Id.*

tients of his who worked in the same area of the hospital also suffered from similar symptoms, met the low threshold requirement for admissibility in a workers' compensation setting.[131] Therefore, the Appeals Court affirmed the decision of the Board,[132] and the Hospital applied to the Massachusetts Supreme Judicial Court for further appellate review.[133]

D. The Supreme Judicial Court's Decision

The primary issue before the Massachusetts Supreme Judicial Court was the admissibility of Dr. LaCava's diagnosis and causation testimony.[134] The importance of the outcome of this issue is reflected in the quality and quantity of amicus briefs submitted. Theresa Canavan's Case engendered comment by four powerful "friends of the court": The Chemical Manufacturers Association together with the Massachusetts Chemical Technology Alliance,[135] the Product Liability Advisory

---

131. *See id.* at 47. In workers' compensation cases the employee is "not required to exclude all other possible sources of [her] injury." *Id.* (citing Rodrigues's Case, 296 Mass. 192, 195, 5 N.E.2d 42,44 (1936)).

132. *See generally Canavan's Case*, 720 N.E.2d 43 (Mass. App. Ct. 1999).

133. *See* Theresa Canavan's Case, 733 N.E.2d 1042, 1045 (Mass. 2000).

134. *See id.* at 1046. Before discussing the admissibility of Dr. LaCava's expert testimony, the court addressed two preliminary issues. First, whether Brigham and Women's Hospital was barred, by judicial estoppel, from denying the existence of multiple chemical sensitivity. *See id. See also infra* note 102. The second preliminary issue was the preservation of the Hospital's objections to Dr. LaCava's testimony. *See Canavan's Case*, 733 N.E.2d at 1047. Theresa Canavan argued that Brigham and Women's Hospital did not sufficiently preserve its objections to Dr. LaCava's testimony, therefore, the Hospital's objections were not a proper issue for review on appeal. *See id.* Dr. LaCava's testimony was taken by deposition. *See id.* The court determined that the steps taken by the attorney representing Brigham and Women's were sufficient to preserve the issue for review on appeal. Those steps included:

> At each point during the deposition prior to Dr. LaCava's offering a conclusion regarding the nature of the employee's condition, . . . the hospital's attorney objected to his conclusion as lacking a foundation. The hospital's attorney reiterated these objections in his written closing submission to the judge in which he explicitly referenced the reliability standards established in Commonwealth v. Lanigan.

Id.

135. *See* Brief of Amici Curiae Chemical Manufacturers Ass'n. and the Massachusetts Chemical Tech. Alliance, Inc., Canavan's Case, 733 N.E.2d 1042 (Mass. 2000) (No. SJC 08226).

> The Chemical Manufacturers Association, also known as the CMA, is a non-profit trade association whose member companies produce, market,

Council,[136] the Massachusetts Defense Lawyers Association,[137] and Marcia

---

and use industrial chemicals . . . [and] [t]hough neither CMA nor any of its members has a direct financial interest in this case, many of its members conduct operations or do business in Massachusetts and are potentially subject to suit in Massachusetts courts.

Id. at 1. The Massachusetts Chemical Technology Alliance, also known as the MCTA, is also a non-profit organization "whose member companies manufacture, distribute, or use chemical technologies." *Id.* The CMA and MCTA disagreed with the Appeals Court's decision that expert testimony based on personal observation or clinical experience is not subject to a *Lanigan* analysis. *See id.* at 1-2. Specifically, in their brief, the CMA and MCTA stated as their single issue "whether expert testimony about an alleged causal relationship between exposure to chemicals and an exposed person's 'multiple chemical sensitivity' is admissible if based only on the expert's training, experience, and clinical observations, with no rational explanation and no empirical support." *Id.* at 2. The CMA and MCTA dispute the validity of MCS as a disease entity. *See id.* at 6-16. They argued that any valid scientific methodology can be proven by testing and generating hypotheses. *See id.* at 6. Furthermore, they argued, as there is no empirical evidence of the existence of MCS, there can be no rational explanation as to the cause of MCS. *See id.* at 18. Thus, the CMA and the MCTA urged the court to reverse the decision of the Appeals Court and render judgement in favor of Brigham and Women's Hospital. *See id.* at 19.

136. *See* Brief of Amicus Curiae The Product Liability Advisory Council, Inc., Canavan's Case, 733 N.E.2d 1042 (Mass. 2000) (No. SJC 08226). The Product Liability Advisory Council, Inc., (PLAC), "is a non-profit corporation whose membership is composed of 127 corporations from a broad cross-section of American industry, and a number of the product liability lawyers who represent these corporations." *Id.* at 1. The primary purpose of PLAC is to lobby on issues that affect product liability law. *See id.* "PLAC has submitted over 400 *amicus* briefs to state and federal courts, including [the SJC]." *Id.* PLAC's brief addressed the sole issue of whether the Appeals Court erred in holding that expert testimony based on personal observations or clinical experience is exempt from a *Lanigan* analysis. *See id.* at 2. PLAC argued that *Lanigan* applied to *all* types of expert opinion evidence and that there is no exception, and there should be no exception to this rule. *See id.* at 22-29.

137. *See* Brief of Amicus Curiae The Massachusetts Defense Lawyers Association, Canavan's Case, 733 N.E.2d 1042 (Mass. 2000) (No. SJC 08226). The Massachusetts Defense Lawyers Association (MDLA) is a voluntary, non-profit professional association of attorneys. The purpose of the MDLA is "to uphold and defend the constitutions of the United States and the Commonwealth of Massachusetts, to promote the administration of justice, to uphold the honor of the legal profession, and to apply the knowledge and experience of its members so as to promote the public good." *Id.* at 1. The brief of the MDLA addressed four issues. First, they argued that expert testimony based on personal observation and clinical experience should be subject to a *Lanigan* analysis. *See id.* at 7-16. Second, Massachusetts does not recognize any exception under which Theresa Canavan's treating physician could offer testimony. *See id.* at 25-27. Third, a judicial determination as to the admissibility of expert testimony should be subject to a de novo standard of review. *See id.* at 27-30. And fourth, the assertions made by Theresa Canavan's treating physician are not sufficiently reliable when applying *Lanigan*. *See id.* at 30-41.

Angell and Others.[138]

In its' opinion, the court first addressed the standard of review on appeal, and then proceeded to evaluate the admissibility of Dr. LaCava's expert testimony.[139] Ultimately, the SJC reversed the decisions of the Industrial Accident Board and the Appeals Court, on the grounds that it was an abuse of discretion for the workers' compensation judge to allow Dr. LaCava's diagnosis testimony[140] and the Appeals Court incorrectly determined that expert testimony based on "personal observations and clinical experience" is not subject to a *Lanigan* analysis.[141] In doing so, the court set out two new rules which are expected to have a great impact on the practice of law in Massachusetts.[142] First, the court adopted an abuse of discretion standard for appellate review of a decision to admit or exclude expert testimony.[143] Second, expert testimony based on personal observa-

---

138. *See* Brief of Amici Curiae Marcia Angell, Alvan R. Feinstein, Ronald E. Gots, Michael Gough, Leonard D. Hamilton, Dudley Herschbach, Steven H. Lamm, Robert J. McCunney, Robert L. Park, Sally L. Satel, Daniel I. Sessler, Arthur C. Upton, James D. Watson, James D. Wilson, and Richard Wilson, Canavan's Case, 733 N.E.2d 1042 (Mass. 2000) (No. SJC 08226). Marcia Angell and Others are a group of 15 scientists "who have studied the role that scientific issues play in public affairs and in particular the way in which they can illuminate disputes between different persons or elements of society in the courts of law." *Id.* at 1. The group is composed of medical doctors from a variety of specialties, professors, and scientists from the environmental field. *See id.* They addressed the following issue:

> [W]hether expert testimony about an alleged causal relationship between exposure to chemicals and an exposed individual's 'multiple chemical sensitivity' is admissible if it is based only on the expert's personal observations, clinical experience, and his own methodology, with no rational explanation, no empirical support, no peer review, no publication of data, hypotheses or methodology, and without adherence to accepted scientific methodology.

Id. at 2-3. Their brief takes the position that it was error for the Appeals Court to admit evidence without ensuring that the underlying principles satisfied epidemiologic criteria with respect to causation. *See id.* at 31-32.

139. *See* Canavan's Case, 733 N.E.2d 1042, 1047-52 (Mass. 2000).

140. *See id.* at 1051. "We cannot conclude that the expert's mere assertion that a methodology is reliable is sufficient to pass the *Lanigan* test absent any other evidence showing its reliability." *Id.*

141. *See id.* at 1049. "Because Dr. LaCava based his opinion on his clinical experience and personal observations of the employee, the [Appeals] court concluded that the expert evidence was exempt from a *Lanigan* analysis. We disagree with this conclusion." *Id.*

142. *See* Meghan S. Laska, *Decision On Experts Ties Hands Of Plaintiff's Bar*, MASS. LAWYER'S WEEKLY, Sept. 4, 2000, at 1. "Plaintiffs' attorneys will face more expert-testimony challenges, bear increased expense in proving scientific methods and may have to turn certain cases away in the wake of [*Canavan*]." *Id.*

143. *See Canavan's Case*, 733 N.E.2d at 1049.

tion or clinical experience is subject to a *Lanigan* analysis.[144]

1. The Standard of Review

In *Commonwealth v. Vao Sok*,[145] decided in August of 1997, the Massachusetts Supreme Judicial Court adopted a de novo standard of review to determine whether a judge's decision to admit or exclude evidence was proper.[146] This was a departure from the position taken by federal appellate courts, and the Supreme Court had not yet addressed the issue.[147]

In December of 1997, the United States Supreme Court finally had occasion to address the issue in *General Electric Co. v. Joiner*.[148] In *Joiner*, the Supreme Court determined that an abuse of discretion standard should be used to review admissibility of evidence under *Daubert*.[149] As a result, when the Supreme Judicial Court of Massachusetts was again presented with the issue, it chose to abrogate its prior decision in *Vao Sok* in favor of an abuse of discretion standard. Keeping in line with the Supreme Court, the SJC concluded that an abuse of discretion standard is proper as "a judge's determination on the reliability of scientific testimony is no different from other evidentiary decisions by a trial judge . . . ."[150] Although the court recognized the advantages and disadvantages of both a de

---

144. *See id.*

145. 683 N.E.2d 671 (Mass. 1997). *Vao Sok* involved the introduction of DNA evidence at three murder proceedings.

146. *See id.* at 677. At the time of the *Vao Sok* decision, applying a de novo standard was contrary to some federal appellate courts, and the United States Supreme Court had not yet addressed the issue. The SJC's rationale in adopting a de novo standard of review was that, in cases involving scientific validity, "a trial judge's conclusion will have applicability beyond the facts of the case before him." *Id.*

> The question of general acceptance of a scientific technique, while referring to only one of the criteria for admissibility of expert testimony, in another sense transcends that particular inquiry, for, in attempting to establish such general acceptance for purposes of the case at hand, the proponent will also be asking the court to establish the law of the jurisdiction for future cases.

Id. (citing Jones v. United States, 548 A.2d 35, 40 (D.C. App. 1988)). Thus, the court reasoned that similar claims could have inconsistent results if a de novo standard of review was not applied. *Id.* Furthermore, the court concluded that "[t]he question of the validity of a particular scientific methodology is thus entitled to the same standard of review as a conclusion of law." *Id.*

147. *See Canavan's Case*, 733 N.E.2d at 1047.

148. 522 U.S. 136 (1997). *See supra* notes 60-68 and accompanying text for a discussion of the *Joiner* case.

149. *See Joiner*, 522 U.S. at 141.

150. *Canavan's Case*, 733 N.E.2d at 1048.

novo and an abuse of discretion standard,[151] it ultimately determined that an abuse of discretion standard would "allow trial judges the needed discretion to conduct the inherently fact-intensive and flexible *Lanigan* analysis, while preserving a sufficient degree of appellate review to assure that *Lanigan* determinations are consistent with the law and supported by a sufficient factual basis in the particular case."[152] The court's decision to abrogate *Vao Sok* in favor of an abuse of discretion standard is the first significant new rule set out in *Canavan*, and it was a conscious effort by the court to follow "the general proposition set forth in *Daubert*."[153]

2. Expert Testimony Based on Personal Observation and Clinical Experience

Whether expert testimony based on personal observation or clinical experience is subject to a *Lanigan* analysis was first presented to the Supreme Judicial Court in *Vassallo v. Baxter Healthcare Corp.*[154] However, the court chose not to address the issue, merely noting the trend toward applying *Daubert* to *all* testimony.[155] The Supreme Judicial Court was again presented with the issue in *Canavan* and the court took the opportunity to adopt the positions set forth in *Joiner* and *Kumho Tire*. Relying on the Supreme Court's decision in *Kumho Tire*,[156] the court ruled that "[t]here is no logical reason why conclusions based on personal observa-

---

151. *See id.* The advantage of a de novo standard of review is that it allows an appellate court to conduct a thorough review of applicable literature in an attempt to reach a well-reasoned decision. *See id.* The disadvantage of a de novo standard of review is that without it, the danger exists of "freezing" perceptions as to the validity of a scientific theory "without accounting for the evolving state of scientific knowledge." *Id.*

152. *Id.* at 1049.

153. *Id.* at 1048.

154. *See* Vassallo v. Baxter Healthcare Corp., 696 N.E.2d 909, 919 (Mass. 1998).

> In view of the narrow issue before us, we need not decide whether, as argued by the plaintiffs, the opinions of their experts fall within those cases which accept expert testimony based on personal observations, clinical experience or generally accepted scientific techniques without application of the possibly more rigorous analysis set out in *Lanigan*, . . . or whether, as argued by the defendants, the opinions require detailed consideration under *Lanigan* for determination of their inherent scientific reliability. We note, however, that the Federal courts (and other State courts) have taken the latter view and have conducted inquiries under the *Daubert* decisions (or the particular State counterpart).

Id.

155. *See id.*

156. *See supra* notes 69-79 and accompanying text.

tions or clinical experience should not be subject to the *Lanigan* analysis . . . . [O]bservation informed by experience is but one scientific technique that is no less susceptible to *Lanigan* analysis than other types of scientific methodology."[157]

Applying *Lanigan* to Dr. LaCava's testimony, the court concluded that Dr. LaCava failed to prove that he "used a reliable methodology to transform his general finding of chemical exposure to his more specific diagnosis of MCS."[158] Also, "Dr. LaCava did not identify any specific studies that show the existence of MCS based on specific symptoms and did not identify tests that can be performed to prove that a patient suffers from MCS."[159] As there is dispute among the medical community as to the existence of MCS as a disease entity, Dr. LaCava's assertion that Theresa Canavan suffered from MCS is supported merely by the *ipse dixit* of the expert, and therefore, is insufficient.[160]

Although the court's conclusion that Dr. LaCava's diagnosis testimony was improperly admitted basically disposed of the case, the court chose to address Dr. LaCava's causation testimony "because of the importance of the question."[161] Despite the fact that tests confirmed that Theresa Canavan had been exposed to chemicals and that other patients of his who had worked at Brigham and Women's Hospital had suffered from similar symptoms, Dr. LaCava's methodology had not been subjected to a *Lanigan* analysis, therefore, his causation testimony was inadmissible.[162]

IV. WHERE DO WE GO FROM HERE?: EXPERT TESTIMONY IN THE 21ST CENTURY

A. Issues Unresolved by *Canavan*

Although the SJC set a bright line rule with respect to admissibility of expert testimony based on personal observation or clinical experience,[163]

---

157. *Canavan's Case*, 733 N.E.2d at 1050.

158. *Id.* at 1051. Dr. LaCava testified that the tests he utilized were generally accepted to prove chemical exposure. *Id.* at 1050. The court took issue with Dr. LaCava's interpretation of these results, in particular his transformation of chemical exposure into a diagnosis of MCS. *See id.* The court took the opportunity to apply *Lanigan* to the expert's testimony instead of remanding the case, similar to the Supreme Court's approach in *Joiner*. *See supra* note 65 and accompanying text.

159. *Canavan's Case*, 733 N.E.2d at 1051.

160. *See id.* at 1051. "[N]othing in *Daubert* . . . requires a . . . court to admit opinion evidence that is connected to existing data only by the *ipse dixit* of the expert." *Id.* (citing General Elec. Co. v. Joiner, 522 U.S. 136, 146 (1997)).

161. *Id.*

162. *See id.* at 1051.

163. Such testimony is subject to judicial "gatekeeping." *See supra* notes 139-44 and accompanying text.

there are a number of issues that remain unresolved by the *Canavan* decision. First, it has been well established from the outset that the *Daubert* factors are not exhaustive;[164] yet neither the Supreme Court in *Kumho Tire* nor the SJC in *Canavan* provide any guidance as to what other factors may be necessary to evaluate expert testimony based on personal observation or clinical experience, testimony previously free from the requirements of *Daubert* and *Lanigan*.[165] In *Kumho Tire*, the Supreme Court recognized the difficulty in applying the *Daubert* factors to non-scientific evidence,[166] a position which the Supreme Judicial Court also took in *Canavan's Case*, however, neither Court offered any suggestions.[167] Certainly, neither court can set out factors which would apply in every case involving non-scientific testimony as the nature and array of such testimony is vast.[168] However, whether judges are better suited to determine these factors in the first place is suspect, bearing in mind that expert testimony based on personal observation or clinical experience may not rely on scientific principles but rather on the experience of the expert. Juries are just as capable of evaluating experi-

---

164. *See Daubert*, 509 U.S. at 593. "[W]e do not presume to set out a definitive checklist or test." *Id. See also Kumho Tire*, 526 U.S. at 141. "[A]s the Court stated in *Daubert*, the test of reliability is 'flexible,' and *Daubert's* list of specific factors neither necessarily nor exclusively applies to all experts or in every case." *Id.*

165. *See Canavan's Case*, 730 N.E.2d at 1050. Although the Court recognized that the reliability of testimony based on personal observation or clinical experience may involve examining factors beyond those enunciated in *Daubert*, and adopted in *Lanigan*, the Court failed to offer any suggestions, or provide any guidance, as to what other factors should be considered and/or how to formulate new factors.

166. *See Kumho Tire*, 526 U.S. at 150.

> We agree with the Solicitor General that "[t]he factors identified in *Daubert* may or may not be pertinent in assessing reliability, depending on the nature of the issue, the expert's particular expertise, and the subject of his testimony." The conclusion, in our view, is that we can neither rule out, nor rule in, for all cases and for all time the applicability of the factors mentioned in *Daubert*, nor can we now do so for subsets of cases categorized by category of expert or by kind of evidence. Too much depends upon the particular circumstances of the particular case at issue.

Id. (citations omitted).

167. *See Canavan's Case*, 733 N.E.2d at 1050 n.5.

> Differing types of methodology may require judges to apply differing evaluative criteria to determine whether scientific methodology is reliable. In the *Lanigan* case, we established various guideposts for determining admissibility including general acceptance, peer review, and testing. Establishing the reliability of personal observations may in some circumstances require examining other criteria.

Id. (citations omitted).

168. *See supra* note 166.

ence-based testimony as judges; extending the judicial "gatekeeping" requirement of *Daubert* to non-scientific testimony merely allows judges to evaluate the credibility of an expert and to weigh the evidence, a function traditionally reserved for juries.[169]

Second, one of the more pressing issues is that, neither *Lanigan* and its progeny, nor *Canavan*, address the procedural issues involved in challenging expert testimony.[170] When and how a party should challenge expert testimony remains an open procedural issue, as does the requisite burden of proof. The problem with leaving the procedural aspects unanswered is that challenging expert testimony may be the basis for a summary judgment motion.[171] When expert testimony proffered to prove an essential element of a party's case is excluded, the result may be summary judgment.[172] Thus, success at challenging expert testimony may signal the end of a case.

B. The Impact of *Canavan's* Unresolved Issues

Canavan's unresolved issues have the potential to permanently damage the delicate balance of power of the Massachusetts judicial system. For example, expanding judicial "gatekeeping" to non-scientific evidence improperly expands the judge's role over that of the jury.[173]

Furthermore, the danger in combining a challenge of expert testimony with a motion for summary judgment is that it may unfairly end a case. The

---

169. *See* Jeffrey M. Schumm, *Precious Little Guidance to the "Gatekeepers" Regarding Admissibility of Nonscientific Evidence: An Analysis of* Kumho Tire Co. v. Carmichael, 27 FLA. ST. U. L. REV. 865, 890 (2000). "Judges are laypersons as well, and there is no inherent reason to believe that the layperson on the bench is more competent at dealing with scientific issues than the laypersons in the jury box." *Id.*

170. *See* Marianne C. LeBlanc, *Theresa Canavan's Case and Its Progeny: Where is the Daubert/Lanigan Line Drawn?, in* CANAVAN'S CASE : WHERE IS THE DAUBERT/LANIGAN LINE DRAWN? 43, 47 (Mass. B. Inst. Nov. 2000) (suggesting that *Canavan* raises "pressing questions with respect to the scope of the scientific analyses to be made by the courts, and the procedural mechanisms by which the principles articulated in *Daubert/Lanigan/Canavan* should be addressed by counsel and the judiciary"). Note, that for purposes of this section, reference will be made to *Lanigan* challenges, but challenges to the reliability of expert testimony may be labeled *Daubert*, *Kumho* or *Canavan* challenges; often these titles are used interchangeably.

171. *See* Douglas K. Sheff, *Debunking the Daubert Delusion*, MATA LEGAL NOTES (Massachusetts Academy of Trial Attorneys, Boston, MA), Oct. 2000, at 2. Attorney Sheff, President of the Massachusetts Academy of Trial Attorneys suggests that this is unfair: "[p]re-trial Motions, predicated upon a *Daubert* challenge, may become mere disguises for Motions for Summary Judgment, except that the quantum of proof to counter a *Daubert* motion will be greater than the legal standards pursuant to [Massachusetts Rule of Civil Procedure] 56." *See id.*

172. *See* FEDERAL JUDICIAL CENTER, REFERENCE MANUAL ON SCIENTIFIC EVIDENCE 54 (2d ed. 2000).

173. *See supra* note 169 and accompanying text.

purpose of a motion for summary judgment, authorized under Rule 56 of the Massachusetts Rules of Civil Procedure,[174] is to facilitate the disposition of a case when only issues of law are involved and there is no dispute over material facts.[175] In a motion for summary judgment, "[t]he moving party bears the burden of establishing the absence of a triable issue . . . [o]nce this is satisfied, the burden shifts to the party opposing summary judgment to allege specific facts establishing the existence of a genuine issue or issues of material fact."[176] Thus, the initial burden is on the moving party. More importantly, "[i]n assessing whether each party has met its burden, the court is not permitted to weigh the evidence, to determine the credibility of any witnesses or make any findings of fact."[177] The court should not assess the credibility of the evidence, rather the evidence should be considered in the light most favorable to the non-moving party.[178] Conversely, under *Lanigan* and *Canavan*, the moving party must merely challenge the admissibility of proffered expert testimony, the burden is on the non-moving party to prove the reliability of its expert's testimony.[179]

A challenge to expert testimony and a motion for summary judgment address two different issues which should not be confused. A challenge to

---

174. *See* MASS. R. CIV. P. 56(c) which states: "[t]he judgment sought shall be rendered forthwith if the pleadings, depositions, answers to interrogatories, and admissions on file, together with the affidavits, if any, show that there is no genuine issue as to any material fact and that the moving party is entitled to a judgment as a matter of law." *Id.*

175. *See* Cassesso v. Commissioner of Correction, 456 N.E.2d 1123, 1125 (Mass. 1983) (citing Community Nat'l Bank v. Dawes, 340 N.E.2d 879 (Mass. 1976) for the principle that the purpose of a motion for summary judgment is "to make possible the prompt disposition of controversies on the merits without a trial, if in essence there is no real dispute as to the salient facts or if only a question of law is involved").

176. Goodwin v. Youth Sports Ass'n, 2001 WL 128442, at *2 (Mass. Super. Feb. 14, 2001) (citing Pederson v. Time, Inc., 532 N.E.2d 1211, 1213 (Mass. 1989)).

177. *Id.*

178. *See* Anthony's Pier Four, Inc. v. Crandall Dry Dock Engineers, Inc., 489 N.E.2d 172, 175 (Mass. 1986) (recognizing that in a motion for summary judgment, the evidence should be "considered with an indulgence in the [opposing party's] favor").

179. *See* Margaret A. Berger, *Procedural Paradigms for Applying the Daubert Test*, 78 MINN. L. REV. 1345, 1365-66 (1994). This is another criticism of *Daubert*. In order to ensure efficiency and fairness, the burden of proof should be on the opponent of expert testimony to show that the proffered testimony is not relevant or unreliable under *Daubert* or *Lanigan* before a court addresses the issue. Otherwise, "if the defendant can force the plaintiffs to put in their scientific evidence without the defendant having to supply affirmative proof, the defendant will have no incentive to conduct discovery. The defendant will instead seek to bring a summary judgment motion as soon as possible." *Id.* at 1376. This has the potential to induce harassment and it deprives the judge of a sufficient record upon which to base his or her decision. *See id.* at 1377.

expert testimony requires an assessment of the relevance and reliability of evidence in an effort to ascertain admissibility, while a motion for summary judgment addresses the sufficiency of evidence. When the two are combined, the judge must first address the expert testimony issue and then address the summary judgment motion, being careful not to confuse the two.[180] "Admissibility entails [the limited] *threshold* inquiry over whether a certain piece of evidence ought to be admitted at trial . . . . A sufficiency inquiry, asks whether the collective weight of a litigant's evidence is adequate to present a jury question."[181] Furthermore, in light of the fact that success at challenging an expert can dramatically affect the outcome of a case, courts must be cautious not to sacrifice fairness and efficiency when executing their "gatekeeper" function.[182]

Many scholars believe that challenging expert testimony will have the greatest impact on pre-trial proceedings, thus procedural rules addressing the application of *Daubert* and *Lanigan* during discovery and motion practice are needed.[183] The next section discusses the timing of *Lanigan* motions, an issue left open by *Canavan*, and offers suggestions as to how to ensure that *Lanigan* motions are not used improperly to end a case.

C. Timing of Lanigan Motions

Challenges to the admissibility of evidence may be brought in a pre-trial motion[184] or during trial.[185] Since success on a challenge to the admis-

---

180. As each is intended to evaluate a different perspective.

181. Judge Harvey Brown, *Procedural Issues Under Daubert*, 36 HOUS. L. REV. 1133, 1153 (1999) (citation omitted).

182. *See id. See also* Berger, *supra* note 179, at 1386 (suggesting that *Daubert* will have the greatest effect on pre-trial proceedings, therefore attention must be given to how *Daubert* "intersects with discovery and motion practice").

183. *See* Berger, *supra* note 179, at 1386.

184. Pre-trial motions usually take the form of a motion in limine. 43 MASSACHUSETTS PRACTICE § 16.1 (West Supp. 2001). "The motion in limine prevents inquiry that poses a substantial danger of eliciting evidence that is clearly prejudicial and inadmissible, or evidence whose probative value is far surpassed by its potential prejudicial effect." *Id.* (citation omitted). The authority to bring a motion in limine does not come from the Rules of Civil or Criminal Procedure, rather "[t]he courts' authority is found in their inherent power to supervise and control pretrial matters and the trial." *Id; see also* Sperberg v. Goodyear Tire & Rubber Co., 519 F.2d 708, 712 (6th Cir. 1975), *cert. denied* 423 U.S. 987 (1975) (prohibiting reference to other pending cases to which the plaintiff was a party properly kept extraneous issues from the jury where allowing such references "would have lengthened the trial, confused the issues and possibly misled the jury").

185. Objections to the admissibility of evidence may be made during trial pursuant to MASS. R. CIV. P. 46 and MASS. R. CRIM. P. 22. Both rules include almost identical language, an attorney who makes an objection must make "known to the court the action which he desires the court to take or his objection to the action of the court."

sibility of expert testimony can signal the end of a case,[186] this Comment suggests that these challenges should be raised pre-trial.

The most important advantage of requiring *Lanigan* motions prior to trial is that, should the proffered experts' testimony be inadmissible, it allows the party the opportunity to find a new expert,[187] assuming of course, that the inadmissibility of the expert testimony does not entirely dispose of the case.[188] Another advantage is that "[r]equiring the motion before trial avoids jurors waiting out in the hallway or judges rendering quick decisions without adequate reflection, research, or both."[189] Challenging expert testimony often involves complex issues, such as the evaluation of complex scientific or medical evidence. Resolving these issues by applying the *Daubert* factors[190] can involve intense research, a time consuming task which can be compromised by a hasty decision. Moreover, many federal courts have chosen innovative approaches in dealing with these complex issues such as selecting neutral experts.[191] "Such innovation is not possible if the trial court is not given advance warning,"[192] and the presence of a jury may put time restraints on such a decision. Furthermore, a successful challenge of expert testimony during trial, in the presence of jurors, may prejudice the party proffering the expert by undermining the credibility of the attorney whose expert is struck "because presentation of the expert may have been promised or referred to during voir dire or opening statement."[193]

Although there are benefits to allowing *Lanigan* motions during trial, these benefits are unconvincing and are outweighed by the advantages of pre-trial motions. One benefit of allowing *Lanigan* motions to be raised during trial is that "all of the discovery will have been conducted and the

---

MASS. R. CIV. P. 46; MASS. R. CRIM. P. 22.

186. *See Daubert*, 509 U.S. at 596 (recognizing that a court may grant summary judgment or a directed verdict where the evidence is insufficient to reach a jury).

187. *See* Brown, *supra* note 181, at 1143 (citing Summers v. Missouri Pac. R.R. Sys., 132 F.3d 599, 605 (10th Cir. 1997) as an example of a case in which the plaintiffs experts were inadmissible under *Daubert* and the court extended the tracking/scheduling order so that the plaintiffs could find new experts).

188. *See supra* note 172 and accompanying text.

189. Brown, *supra* note 181, at 1143.

190. Or other appropriate factors. *See supra* notes 51-54 and accompanying text for a discussion of the *Daubert* factors.

191. *See* Maritime Overseas Corp. v. Ellis, 971 S.W.2d 402, 414-15 (Tex. Sup. Ct. 1998) (Gonzalez, J., concurring) (advocating the use of neutral experts such as specially trained scientists or a special master in cases involving complex issues). *See also Justice Breyer Calls for Experts to Aid Courts in Complex Cases*, N.Y. TIMES, Feb. 17, 1998, at A17. Admittedly, this procedure is more common in federal courts.

192. *Maritime Overseas Corp.*, 971 S.W.2d at 414 (Gonzalez, J., concurring).

193. Brown, *supra* note 181, at 1144.

court will have all of the necessary information to make a ruling," [194] thus, the court may have a better understanding of all the issues involved. However, if a challenge is made pre-trial, the parties may request an evidentiary hearing wherein they would have the opportunity to adequately present "all of the necessary information to make a ruling."[195]

Another benefit of waiting until the time of trial to challenge expert testimony is that it can avoid unnecessary hearings because a majority of cases settle prior to trial.[196] However, this argument is circuitous. If a party anticipates a challenge to the admissibility of expert testimony, the prudent attorney would wait until the court rules on the issue before considering settlement under the present system. For example, if a defense attorney anticipated bringing a *Lanigan* motion, that attorney would be wise to wait until after the court rules on the issue because success may dispose of the plaintiffs case.[197] A third benefit of waiting until trial to challenge expert testimony is that it may reduce costs by only requiring the expert's presence in court once, at the time of trial, "as opposed to twice—once for the hearing and a second time for trial."[198] Although this may be true if the only consideration is the cost of paying the expert whose testimony is challenged, however, when weighed against the cost of preparing for and trying a case before a jury, the savings are minimal. Therefore, although there are benefits of waiting until the time of trial to challenge expert testimony; fairness, efficiency, and justice require that such challenges be raised prior to trial.

V. RECOMMENDATIONS FOR THE FUTURE

A. Require Challenges of Expert Testimony to be Made Pre-Trial

The purpose of this Comment is to advocate for procedural rules to address the timing of challenging expert testimony, whether these challenges are labeled *Daubert*, *Lanigan*, or *Canavan* motions, particularly in civil cases.[199] Such rules are needed in order to ensure that judges adequately respond as "gatekeepers" and to facilitate the orderly flow of the litigation process by ensuring efficiency, fairness, and maximization of judicial resources.[200] This is particularly true in light of the fact that such decisions are reviewed under an abuse of discretion standard with consid-

---

194. *See id.* at 1143.

195. *Id.*

196. *See id.*

197. *See supra* note 172 and accompanying text.

198. Brown, *supra* note 181, at 1144.

199. There is also a need for such rules in criminal cases, but the focus of this section is on civil cases.

200. *See* Brown, *supra* note 181, at 1160 n. 176.

erable deference given to trial judges.[201]

As discussed in Part IV (C), the advantages of requiring expert challenges prior to trial clearly outweigh the advantages of allowing such challenges during trial.[202] Therefore, this Comment suggests that courts adopt a rule which requires that challenges of expert testimony be brought prior to trial, for example thirty days prior to trial or at the time of the final pre-trial conference.[203] In civil cases, compliance with the Massachusetts Rules of Civil Procedure requires the proponent of expert testimony to identify expert witnesses and the substance of their opinions in response to appropriate discovery.[204] This requirement gives the opposing party sufficient notice of proffered expert testimony. "The opponent of such testimony should bring its objections to the trial court's attention so that the trial court may resolve them without interfering with the eventual trial."[205] Such a rule could easily be drafted to ensure fairness to all parties by simply including language which permits challenging expert testimony during trial, in the event that such a challenge is not made prior to trial, and provided that there is a showing of good cause and no fault on the part of either party.

The purpose of advocating for adoption of a procedural rule to address when and how to challenge expert testimony is to aid in the resolution of issues left unresolved by *Canavan*. However, the adoption of such a rule does not resolve the entire problem because the *Canavan* decision also raises concerns over the effectiveness of the legal system as a whole. The next Section analyzes the need to re-evaluate the role that

---

201. *See supra* notes 62-68 and 148-53 for a discussion of the abuse of discretion standard.

202. *See supra* notes 184-98 and accompanying text for a discussion of the advantages and disadvantages of the timing of expert challenges.

203. *See* Brown, *supra* note 181, at 1144. This suggestion is based on Judge Gonzalez' concurring opinion in *Maritime Overseas Corp. v. Ellis*, 971 S.W.2d 402, 414-15 (Tex. Sup. Ct. 1998). Judge Gonzalez suggests that procedural rules are needed to address expert challenges. Specifically, Judge Gonzalez suggests that courts "require parties to notify opponents and the court sufficiently in advance of the trial of plans to either offer scientific evidence or challenge an opponent's evidence; [and] render expert testimony inadmissible or rule objections waived unless the parties fully comply with the notice requirements." *Id.*

204. *See* MASS. R. CIV. P. 26(b)(4)(A)(i) which provides:

> A party may through interrogatories require any other party to identify each person whom the other party expects to call as an expert witness at trial, to state the subject matter on which the expert is expected to testify, and to state the substance of the facts and opinions to which the expert is expected to testify and a summary of the grounds for each opinion.

Id.

205. *Maritime Overseas Corp.*, 971 S.W.2d at 414 (Gonzalez, J. concurring).

judges and jurors play in our current system in the wake of *Canavan*.

B. A Concerted Effort

It will take a concerted effort by the judiciary, the Bar, juries, expert witnesses and litigants to settle *Canavan's* unresolved issues. A constant struggle for trial judges in the post-*Daubert* era, and consequently the post-*Lanigan* era, has been how to sift through evidence in an effort to exclude that which is unreliable while remaining faithful to the purpose of *Daubert* (which was to liberalize the admissibility of expert testimony).[206] Recent cases merely instruct the judge that he or she must act as "gatekeeper" in an ever-increasing number of situations without offering any guidance as to how to carry out this function.[207] The trend in these recent cases has been to expand the court's role in screening every aspect of every expert's testimony at every stage of a case, thus expanding judicial discretion over jury discretion by creating a significant obstacle that every case must satisfy in order to be presented to a jury.[208] By extending the reach of *Daubert* and *Lanigan* to all expert testimony, a judge has the ability to take any case which relies on expert testimony away from the jury.[209] *Canavan* is the most recent in this line of cases which follows the trend in expanding judicial over jury discretion.[210] Furthermore, the court in *Canavan*, by adopting an abuse of discretion standard, further heightened judicial scrutiny over jury discretion by elevating the screening role of the judge as "gatekeeper" above the liberalization of admissibility of expert testimony, a result not intended by *Daubert*.[211] The result is that judges are free to choose between competing expert testimony without much guidance or recourse. "To say that a [judge] has discretion in a given area of law is to say that [he or she] is not bound to

---

206. *See* Daubert v. Merrell Dow Pharmaceuticals, 509 U.S. 579, 595-96 (1992). *See also* Richard Collin Mangrum, *Kumho Tire Company: The Expansion of the Court's Role in Screening Every Aspect of Every Expert's Testimony at Every Stage of the Proceedings*, 33 CREIGHTON L. REV. 525, 538 (2000). "Following *Daubert*, the courts were faced with reconciling the seemingly contradictory 'liberalization' impulse of *Daubert* with the screening qualifier." *Id.*

207. *See supra* notes 48-79 and accompanying text, which illustrate the expanding role of judges as gatekeepers.

208. *See* Mangrum, *supra* note 206, at 538-39. *Daubert* and *Kumho Tire* affect every aspect of a case, from pre-trial discovery to appeals. This continuous scrutiny creates new hurdles for attorneys. *See id.*

209. *See id.* at 538.

210. *Canavan* stands for the same principle as *Kumho Tire*, that expert testimony based on personal observation and/or clinical experience is subject to judicial "gatekeeping," therefore the criticisms of *Kumho Tire* are also applicable to *Canavan*.

211. *See* Mangrum, *supra* note 206, at 538. *See also supra* notes 149-53 and accompanying text for a discussion of the abuse of discretion standard.

decide the question one way rather than another."[212] Clearly, this power is extraordinary. Unfortunately, the problem with this trend toward judicial over jury discretion is that it usurps the role and duty of the jury.[213] This continuous trend toward judicial discretion over jury discretion threatens the very existence of our jury system, a system guaranteed by both the United States Constitution[214] and the Massachusetts Declaration of Rights.[215] "Throughout history, the right to a trial by jury has been viewed by our founding fathers, the framers of our Constitution, and all citizens of the United States since its inception, as essential to the freedoms that make our society great."[216]

Despite the deeply rooted history of the jury system, there has been considerable debate over the credibility of the jury system in recent

---

212. Maurice Rosenberg, *Judicial Discretion of the Trial Court, Viewed from Above*, 22 SYRACUSE L. REV. 635, 636-37 (1971). The result is that there is no "officially" wrong answer. *See id.*

213. *See* Mangrum, *supra* note 206, at 537-38.

214. The Seventh Amendment to the Constitution of the United States provides: "In Suits at common law, where the value in controversy shall exceed twenty dollars, the right of trial by jury shall be preserved, and no fact tried by a jury, shall be otherwise re-examined in any Court of the United States, than according to the rules of common law." U.S. CONST. amend. VII. With respect to criminal cases, the Sixth Amendment to the U.S. Constitution provides that in all criminal matters, trial shall be by a jury:

> In all criminal prosecutions, the accused shall enjoy the right to a speedy and public trial, by an impartial jury of the State and district wherein the crime shall have been committed, which district shall have been previously ascertained by law, and to be informed of the nature and cause of the accusation; to be confronted with the witnesses against him; to have compulsory process for obtaining witnesses in his favor; and to have the Assistance of Counsel for his defence.

U.S. CONST. amend. VI.

215. Likewise, Article XV of the Massachusetts Declaration of Rights provides:

> In all controversies concerning property, and in all suits between two or more persons, except in cases in which it has heretofore been otherways used and practiced, the parties have a right to a trial by jury; and this method of procedure shall be held sacred, unless, in causes arising on the high seas, and such as relate to mariners' wages, the legislature shall hereafter find it necessary to alter it.

MA. CONST. pt.1, art. XV. Furthermore, Article XII of the Massachusetts Constitution guarantees that "no subject shall be arrested, imprisoned, despoiled, or deprived of his property, immunities, or privileges, put out of protection of the law, exiled, or deprived of his life, liberty, or estate, but by the judgment of his peers, or the law of the land." MA. CONST. pt.1, art. XII.

216. Michael Sudman, *The Jury Trial: History, Jury Selection, and the Use of Demonstrative Evidence*, 1 J. LEGAL ADVOC. & PRAC. 172, 173 (1999).

years.[217] The most salient criticism of the jury system is that juries are ill-equipped to decide complex issues.[218] However, empirical research shows that, in most cases, juries are competent and they effectively perform their duties.[219] "When problems arise, it is often the quality of the presentation that is implicated, rather than inherent deficiencies in juror's abilities to process the information provided."[220] That jurors are able to comprehend and analyze complex evidence and that many difficulties that jurors encounter can be attributed to an attorney's presentation of such evidence should be a wake-up call to all attorneys. As difficult as it may be, attorneys must accept some responsibility when difficulties arise during presentation of complex issues.

Although this debate has been brought to the forefront in recent years, it is not new. There has been debate over the jury system for, at least, the past fifty years.[221] Despite the criticisms, the jury system persists. It is inherently American.[222] The jury system ensures the allocation of power in the American courtroom, "[i]n essence, according the jury a greater voice during trial reallocates the division of power, providing a symbolic and perhaps even an actual check on the court."[223] Although some of the criticisms of the jury system are valid, [224] they do not warrant total abandonment of the jury system, nor do they require that judges assume a greater role in the decision-making process. Rather, the solution lies somewhere in between. Instead of shifting the allocation of power between judge and jury, on the mistaken belief that juries are incapable of understanding complex evidence, focus should be placed on empowering the jury because

---

217. *See, e.g.*, FRANKLIN STRIER, RECONSTRUCTING JUSTICE 107 (1994) (suggesting that juries are often portrayed "as both a convenient scapegoat and a panacea for our major complaints about the adversary trial").

218. *See* Steven I. Friedland, *The Competency and Responsibility of Jurors in Deciding Cases*, 85 NW. U. L. REV. 190, 190 (1990).

219. *See* G. THOMAS MUNSTERMAN ET AL., JURY TRIAL INNOVATIONS 7 (1997). *See also*, Joe S. Cecil et al., *Citizen Comprehension of Difficult Issues: Lessons From Civil Jury Trial*, 40 AM. U. L. REV 727, 745 (1991) (analyzing studies concerning the ability of jurors to comprehend and apply complex issues).

220. MUNSTERMAN ET AL., *supra* note 219, at 7.

221. *See* Singer v. United States, 380 U.S. 24, 35 (1965) (recognizing that "trial by jury has its weaknesses and the potential for misuse").

222 *See, e.g.*, Duncan v. Louisiana, 391 U.S. 145, 149 (1968) (suggesting that "trial by jury . . . is fundamental to the American scheme of justice"); STRIER, *supra* note 217, at 110 (distinguishing the American system from the English, French and German systems).

223. Friedland, *supra* note 218, at 207-08 (1990). The presence of a jury, at least theoretically, ensures a fair trial by distributing the power between the judge, attorneys and the community. *See id.*

224. Studies have proven that jurors have more difficulty understanding and utilizing complex evidence. *See, e.g.*, MUNSTERMAN ET AL., *supra* note 219, at 7.

"[t]he capabilities of jurors—perhaps not as individuals but as a group—even appear to extend to cases of the greatest complexity."[225] In an effort to empower juries, experiments in jury reform are taking place across the country.[226] Most of these experiments involve expanding the role of the jury during trial in an effort to make jurors more active participants, this has been referred to as the "active juror model."[227] Proponents of jury reform argue that the passive role juries are forced to assume interferes with their ability to learn and comprehend, thereby decreasing their efficiency and interfering with their ability to act as community representatives in the decision making process.[228] The benefits of the active juror model have been well documented.[229] A more active juror is more likely to remember the evidence and/or the law, and is less likely to be confused.[230] Furthermore, when jurors take a more active role, it enhances public confidence in the decisionmaking process.[231] The American Bar Association has endorsed the active juror model:

> Jurors need not and should not be merely passive listeners in trials, but instead should be given the tools to become more active participants in the search for just results. To that end, trial procedures and evidentiary rules should take greater advantage of modern methods of communication and recognize modern understanding of how people learn and make future decisions.[232]

Despite the benefits of the active juror model, efforts to implement this system have been slow to gain support. In the last twenty years, Massachusetts has been on the cutting edge in terms of jury reform efforts.

---

225. Cecil, *supra* note 219, at 745. *See also* MUNSTERMAN ET AL., *supra* note 219, at 7 (suggesting that "improvements in jury performance can be achieved by improving the quality of communications with the jury").

226. Arizona has been the leader in jury reform. *See, e.g.*, Janessa E. Shtabsky, Comment, *A More Active Jury: Has Arizona Set the Standard for Reform With Its New Jury Rules?*, 28 ARIZ. ST. L.J. 1009 (1996); CHARTING A FUTURE FOR THE CIVIL JURY SYSTEM, BROOKINGS INSTITUTION 3 (1992).

227. *See* BROOKINGS INSTITUTION, *supra* note 226, at 3.

228. *See* Hon. B. Michael Dunn, *"Learning Lessons" and "Speaking Rights": Creating Educated and Democratic Juries*, 68 IND. L.J. 1229, 1229-30 (1993). "[E]nforced passivity interferes with learning and reduces opportunities for jurors, individually and collectively, to perform to their potential as community representatives and decision makers in trials of criminal and civil cases." *See id.* at 1236.

229. *See*, e.g., Friedland, *supra* note 218, at 207. *See also* Shtabsky, *supra* note 226, at 1012 -13.

230. *See* Dunn, *supra* note 228, at 1231.

231. *See* Friedland, *supra* note 218, at 207.

232. BROOKINGS INSTITUTION, *supra* note 226, at 3.

Although initial attempts at jury reform have been adopted with great success, more recent efforts have been slow to gain uniform acceptance.

Massachusetts was one of the first states in the nation to adopt the one day/one trial system.[233] This system was adopted in response to *Commonwealth v. Bastarache*.[234] In *Bastarache*, the defendant, convicted of manslaughter, appealed his conviction by challenging the composition of the grand jury that indicted him, and the trial jury that convicted him.[235] The defendant argued that both the grand jury, and the trial jury, were underrepresented by individuals between the ages of eighteen and thirty-four.[236] The Appeals Court agreed, reversed the defendant's conviction and dismissed the indictment.[237] The SJC, on further appellate review, determined that the procedures used to establish jury lists in Franklin County, where the case was heard, did not amount to a Constitutional violation.[238] However, the procedures did result in under-representation by younger individuals.[239] Therefore, the SJC asked the Attorney General of the Commonwealth "to prescribe procedures for the compilation of jury lists in those cities and towns that are not now using a substantially random selection process."[240] According to the SJC, the jury selection process "should increase confidence in the jury system, enhance the appearance of fairness, and distribute more evenly the civic responsibility to serve on juries."[241] At the time the *Bastarache* case was decided, the "jury statute"[242] was a pilot program in effect in Middlesex County only.[243] The

---

233. *See* Munsterman ET AL., *supra* note 219, at 29. In the "one day/one trial system," jurors fulfill their obligation by performing one day of jury duty service even if not chosen to sit on a jury. If a juror is chosen to sit on a jury, the juror's obligation is fulfilled when the case is completed. *See id.* This system was first implemented in Texas in 1972. *See id.* Some states have implemented this procedure statewide, while others use it selectively. It "is [used] statewide in Massachusetts, Connecticut, Florida, and Colorado, and is used in most courts in New York, Arizona, North Carolina and Texas." *See id.*

234. Commonwealth v. Bastarache 414 N.E.2d 984 (Mass. 1980). *See also* http://www.state.ma.us/courts/jury/introduc.htm (last visited on Aug. 27, 2001).

235. *See Bastarache*, 414 N.E.2d at 987. Initially the defendant challenged the jury composition in a pretrial motion. *See id.*

236. *See id.*

237. *See id.*

238. *See id.* at 993. The procedure for compiling jury lists in Franklin County authorized the board of selectmen to "prepare a list of such inhabitants of the . . . town, qualified, . . . of good moral character, of sound judgment and free from all legal exceptions, not exempt from jury service . . . as they think qualified to serve as jurors." *Id.* at 989.

239. *See Bastarache*, 414 N.E. 2d at 995.

240. *See id.* (recognizing that involvement by the Legislature may be needed).

241. *Id.*

242. *See* OFFICE OF JURY COMMISSIONER, THE MASSACHUSETTS JURY SYSTEM IN

"jury statute"[244] eventually expanded the one day/one trial system statewide.[245] The advantage of the one day/one trial system is that it promotes juror satisfaction and increases the number of potential jurors which, in turn, "increases the representativeness and inclusiveness of the jury pool."[246]

Although Massachusetts has been eager to enact programs which focus on getting prospective jurors in the doors of courthouses, in addition to the one day/one trial system Massachusetts has also enacted the delinquent juror prosecution program[247] and a public outreach program,[248] Massachusetts has only recently begun to experiment with methods of increasing the effectiveness of the jury system.

C. Jury Trial Innovations in Massachusetts[249]

The National Center for Citizen Participation in the Administration of Justice and the Flaschner Judicial Institute[250] with the assistance of G. Thomas Munsterman, Director of the Center for Jury Studies of the Na-

---

BRIEF 2 (1998) citing Chapter 415 of the Acts of 1977, later codified under chapter 234A, which was passed on January 2, 1979.

243. *See* http://www.state.ma.us/courts/jury/introduc.htm *supra* note 234. *See also* OFFICE OF JURY COMMISSIONER, *supra* note 242, at 2.

244. *See* OFFICE OF JURY COMMISSIONER, *supra* note 242, at 2, citing Chapter 234A amended by chapter 298 of the Acts of 1982.

245. *See* OFFICE OF JURY COMMISSIONER, *supra* note 242, at 2. The one day/one trial system was fully operational by March of 1988. "Massachusetts became the first in the nation to operate this then relatively new, highly regarded, and nationally renowned jury system on a statewide basis." *Id.*

246. *See* MUNSTERMAN ET AL., *supra* note 219, at 30. Limiting the length of juror service also reduces the hardship associated with jury duty and encourages courts to use juror time more efficiently. *See id.*

247. *See* OFFICE OF JURY COMMISSIONER, *supra* note 242, at 15. The delinquent juror prosecution program is authorized by MASS. GEN. LAWS ch. 234A, §§ 42-44 (1979). Under the delinquent juror prosecution program, "[u]pon a finding by the court that a juror will not appear to perform or complete juror service . . . the court may issue a warrant for the arrest of the juror or may take such other appropriate actions as are likely to compel the juror to appear before the court." *Id.*

248. *See* OFFICE OF JURY COMMISSIONER, *supra* note 242, at 25. The goal of the public outreach project, a division of the Office of Jury Commissioner, is to educate the public concerning the importance of jury duty. *See id.*

249. This section analyzes more recent jury trial reform efforts. Most of the information is taken from a book by the same name. HON. PETER M. LAURIAT ED., JURY TRIAL INNOVATIONS IN MASSACHUSETTS (2000).

250. *See* LAURIAT, *supra* note 249, at xi. The Flaschner Institute was established in 1978 as a memorial to the late Chief Justice Franklin N. Flaschner. *See id.* The Flaschner Institute has been characterized as a "self-help judges' organization." *See id.* Its goal is to improve professionalism by judicial education. *See id.*

tional Center for State Courts, and the Honorable B. Michael Dann of the Arizona Superior Court[251] founded the Massachusetts Project on Innovative Jury Practices (hereinafter "the Project").[252] "The Massachusetts Project on Innovative Jury Practices is a demonstration project intended to encourage the use of jury trial practices that enhance juror performance and satisfaction with jury service."[253] The Project involved a year-long study in which participating judges experimented with jury trial innovations in their courtrooms.[254] Participating judges, jurors and attorneys were asked to complete questionnaires assessing their views of the innovations.[255] The innovative jury trial practices employed were: allowing juror notetaking;[256] distributing juror notebooks;[257] preinstructing the jury;[258] permitting jurors to question witnesses;[259] using "plain English" at

---

251. *See* Hon. B. Michael Dann & George Logan III, *Jury Reform: The Arizona Experience*, 79 JUDICATURE 280 (1996). Arizona has been a pioneer in jury reform. Judge Dann sits in the Maricopa County Superior Court and is chair of the Arizona Supreme Court Committee on More Effective Use of Juries. *See id.*

252. *See* LAURIAT, *supra* note 249, at xiii (Hon. Peter M. Lauriat ed.) (2000). *See also* NATIONAL CENTER FOR STATE COURTS, DRAFT REPORT FOR THE MASSACHUSETTS PROJECT ON INNOVATIVE JURY TRIAL PRACTICES 1 (April 5, 2000) *available at* http://www.masslaw.com/matreas/dfrinnov.htm.

253. NATIONAL CENTER FOR STATE COURTS, *supra* note 252, at ii *available at* http://www.masslaw.com/matreas/dfrinnov.htm.

254. *See id.* at 1-2. The Project began with a conference in which participating judges discussed innovative jury practices. *See id.* The participating jurors also met periodically throughout the year to discuss the Project. *See id.* at 2.

255. *See id.* at 2.

256. *See id.* at 5. Juror notetaking has been an approved practice in Massachusetts since 1978. *See* MASS. SUP. CT. R. 8A. *See also* Bencosme v. Kokoras, 507 N.E.2d 748 (Mass. 1987). It was the most commonly used technique in the Project and it had unanimous approval from those who tried it. *See* NATIONAL CENTER FOR STATE COURTS, *supra* note 252, at 5.

257. *See* NATIONAL CENTER FOR STATE COURTS, *supra* note 252, at 18.

> This technique consists of providing notebooks for jurors to use, particularly in lengthy trials and trials of complex cases. The contents of the notebooks will vary with the needs of the case. In highly complex cases, a notebook might consist of paper for taking notes, preliminary jury instructions, a short statement of the parties' claims and defenses, a list of witnesses by name (including identifying information and phonetic spellings when helpful[,] copies of key exhibits, glossary of technical terms, and the final jury instructions. In simpler cases, however, a 'notebook' may not be necessary at all. Instead, it may be sufficient to provide only copies of important documents or exhibits.

Id.

258. *See id.* at 7. "Preinstructions for the jury consist fundamentally of an introduction to the parties and their claims, a presentation of matters not in dispute, and guidance on the contested issues and governing legal principles of the case." *Id.* This

trial;[260] using mini-opening statements together with interim commentary;[261] placing limits on each party's time at trial;[262] permitting jurors to discuss the evidence prior to deliberating;[263] using "plain English" jury instructions;[264] permitting jurors to ask questions about final instruc-

---

technique was favored by attorneys, but there was concern that it might be difficult in complex cases involving numerous issues. *See id.*

259. *See* NATIONAL CENTER FOR STATE COURTS, *supra* note 252, at 12-14. This is one of the most controversial innovations. *See id.*

> Permitting jurors to submit questions to witnesses provides a procedural mechanism for jurors to ask questions about the evidence or testimony presented at trial with minimal risk to the due process rights of the litigants. Jurors are instructed at the beginning of trial that they may submit questions that they would like to have asked of the witnesses in writing. If jurors submit questions, they are first marked for identification and then the judge reviews the questions with the attorneys out of the presence of the jury.

Id. at 12. This technique is specifically targeted to aid jurors in clarifying any areas of confusion. A comment by one of the judges on the benefits of permitting jurors to ask questions:

> [Permitting jurors to submit questions to witnesses] enhances the jurors' understanding of the important fact issues and clears up significant misunderstandings. It enhances the attorneys' ability to intelligently address the key issues by informing the attorneys of specific issues that the jurors think are important and giving the attorneys the opportunity to clarify and further address the areas raised in the jurors' questions. It enhances the overall quality of the trial process by treating the jurors as important, active, responsible participants. And it enhances the experience of the jurors as active, intelligent, responsible participants and citizens of a democracy.

Id. at 13.

260. *See* NATIONAL CENTER FOR STATE COURTS, *supra* note 252, at 8-9. The goal of this innovation is to enhance jurors' understanding of a case by eliminating "legalese" because studies have shown "[a] strong correlation exists between juror satisfaction with jury service and how well jurors understood the proceedings in which they participated and their role within those proceedings." *Id.* at 8.

261. *See id.* at 20-21. This innovation allows attorneys to give a "short opening statement at the beginning of trial and subsequent statements periodically throughout the trial." *Id.* at 20. This innovation was not frequently used in the Project. *See id.*

262. *See* NATIONAL CENTER FOR STATE COURTS, *supra* note 252, at 20. This technique was unanimously favored by the three judges who used it, but the attorneys "were less approving." *See id.*

263. *See* NATIONAL CENTER FOR STATE COURTS, *supra* note 252, at 21. This is another controversial technique. It allows jurors to discuss the case under three conditions: (1) all jurors must be present; (2) discussions must be in the jury room; and (3) jurors can only discuss the evidence and not the outcome of the case. *See id.* Jurors were especially receptive to this practice. *See id.*

264. *See* NATIONAL CENTER FOR STATE COURTS, *supra* note 252, at 10-11.

tions;[265] providing jurors with written or taped copies of final instructions;[266] giving jurors final instructions before closing arguments;[267] giving jurors suggestions on conducting deliberations;[268]conducting post-verdict meetings with the judge and jurors;[269] debriefing jurors;[270] and conducting post-verdict meetings with judges, attorneys and jurors.[271] The overall results were positive.[272] The most frequently cited obstacle to implementing these reform practices was time, as many of the reform practices "require a degree of preparation that cannot usually be accomplished on the morning of the day of trial."[273] Consequently, the draft report for

---

> This technique requires judges and trial attorneys to draft jury instructions in comprehensible language, with specific attention to the overall character and structure of the jury charge. Although it is lauded as one of the most sensible innovations in jury trial management, it is a time and labor-intensive process that requires significant involvement by the judge and trial attorneys either in pretrial meeting or during trial – something that many are unwilling to do without assurances that specially drafted jury instructions will survive appellate scrutiny.

Id. at 10.

265. *See id.* at 16. This technique had unanimous support from judges who used it. *See id.*

266. *See id.* at 16-17. While the goal is to aid in juror comprehension of jury instructions, providing written or taped instructions is limited by court resources. *See id.*

267. *See* NATIONAL CENTER FOR STATE COURTS, *supra* note 252, at 19. There was little feedback on this technique. *See id.*

268. *See id.* at 11. "Suggestions for conducting deliberations can help overcome the awkwardness that often accompanies inexperience, and can provide a practical framework for jurors to undertake their decision-making tasks." *Id.* Although most judges felt this was useful, jurors were less enthusiastic about this practice. *See id.* at 12.

269. *See* NATIONAL CENTER FOR STATE COURTS, *supra* note 252, at 9. This provides judges with the opportunity to explain post-verdict procedures to the jurors and to get feedback from the jurors. *See id.*

270. *See id.* at 14-15. This technique is generally used where the subject matter of a trial is emotionally upsetting. *See id.* at 15. It had unanimous acceptance by those judges who used it. *See id.*

271. *See* NATIONAL CENTER FOR STATE COURTS, *supra* note 252, at 19. This practice gives attorneys the opportunity to hear jurors' comments. *See id.*

272. *See* NATIONAL CENTER FOR STATE COURTS, *supra* note 252, at 4. "Overall, the participating judges were enthusiastic about the majority of reform practices that they introduced in their courtrooms." *Id.* Justice Patrick F. Brady commented "[a]ll of the practices used in the project seem to involve the jury more actively in the learning process necessary to a rational decision. Jurors in my experience have reacted very favorably to the practices. I have found virtually no drawbacks in or contraindications to using these new practices." *Id.* at 4-5.

273. *Id.* at 23.

the Project recommended restructuring the pretrial management system.[274] Surprisingly, despite the positive results overall, the participating judges encouraged the use of innovative jury reform practices but did not support mandatory implementation of reform practices via statute or court rules.[275]

Restructuring the pretrial management system, although a monumental task, seems a small price to pay for the potential rewards in terms of increasing the active role of jurors in the decision-making process. This is the only way to empower the jury in an effort to reallocate power in the courtroom. Without jury reform, jurors are forced to sit in courts across Massachusetts silent and confused. "Jurors are more attentive when given the opportunity to ask questions, even if they choose not to exercise it."[276] Furthermore, restructuring the pretrial management system would be invaluable in resolving the procedural problems raised by *Canavan*,[277] in light of the fact that one of the biggest obstacles to ordering challenges of expert testimony prior to trial is the ability of the pretrial system to handle these challenges.[278]

VI. CONCLUSION

In the last decade of the twentieth century, judges have attained immense power and our jury system has suffered as a result. The expansion of power began with the Supreme Court's entry into the battle over admissibility of expert testimony. The Supreme Court's trilogy of cases strategically produced a procedural framework wherein the trial judge screens every aspect of every expert's testimony at every stage of a case.[279] This increase in judicial power began with *Daubert* which required that judges act as gatekeepers in an effort to screen expert testimony.[280] *Joiner* enhanced judicial power by establishing abuse of discretion as the standard

---

274. *See id.* at 23. "If jury reform practices . . . are to have a reasonable chance of success in Massachusetts, the Superior and District Court procedures should be restructured to provide judges with sufficient time and resources to more effectively conduct pretrial management of the cases on their dockets." *Id.*

275. *See* NATIONAL CENTER FOR STATE COURTS, *supra* note 252, at 24. Although many "judges indicated that it would be useful to have court rules issued clarifying that use of these practices was permissible under Massachusetts law," they had reservations about implementing some of the innovations, such as juror notebooks and debriefing, in simple cases. *See id.*

276. Judge Ken Curry & M. Beth Krugler, *The Sound of Silence: Are Silent Juries the Best Juries?*, 62 TEX. B.J. 441, 447 n.32 (1999).

277. *See supra* notes 189-190 and accompanying text.

278. *See supra* notes 184-198 for a discussion of the timing of expert challenges.

279. *See* Mangrum, *supra* note 206, at 525. *See supra* notes 45-79 for a discussion of the trilogy of cases.

280. *See supra* notes 45-59 for a discussion of *Daubert*.

of review of judicial rulings on admissibility of expert testimony.[281] Finally, *Kumho Tire* expanded the judge's role as gatekeeper so that a judge must evaluate all testimony based on "scientific, technical, or other specialized knowledge."[282] *Kumho Tire* is "an expansion of the court's role in screening every aspect of every expert's testimony at every stage of the proceedings."[283] "The fact that the *Daubert*, *Joiner*, and *Kumho* decisions charge trial judges with the responsibility of evaluating the reliability of expert testimony presupposes that the traditional adversary process is insufficient to enable opposing counsel and their expert to ferret out inaccuracies and bias in expert testimony."[284] Massachusetts subsequently adopted the principles set forth in *Daubert*, *Joiner* and *Kumho Tire* in *Lanigan* and *Canavan*.[285] Unfortunately, these cases do not address the real problem. Instead of changing the jury system to keep up with changing times, the courts have chosen to screen all expert testimony before it is presented to a jury. The jury system in America "has increasingly become a source of social ridicule, criticism, and concern," and the recent cases do nothing to minimize this.[286] The reality is that "[o]verly restrictive gatekeeping . . . may be selling juror's abilities short and depriving juries of valuable information that could help to reach a correct verdict."[287] This is particularly true in the case of nonscientific testimony which is essentially experience-based; the reliability of such evidence is not validated through external testing, rather experience-based testimony is validated by examining the principles and methodologies used.[288] There is no reason to believe that a jury is incapable of evaluating the reliability of non-scientific evidence, i.e., evidence based on personal observations or clinical experience.[289] Furthermore, even in the case of purely scientific evidence, as Justice Breyer recognized in his concurring opinion in *Joiner*, "judges are not scientists and do not have the scientific training that can facilitate the making of such decisions." [290]

With all due respect, the purpose of this comment is to advocate restructuring of the pre-trial process and jury reform. Jury reform has the ability to empower juries so that they may regain their rightful place in the

---

281. *See supra* notes 62-68 for a discussion of *Joiner*.

282. *See supra* notes 69-79 for a discussion of *Kumho Tire*.

283. Mangrum, *supra* note 206, at 525.

284. *Id.* at 536.

285. *See supra* notes 86-162 and accompanying text.

286. Natasha K. Lakamp, *Deliberating Juror Predeliberation Discussions: Should California Follow the Arizona Model?*, 45 UCLA L. REV. 845, 847 (1998).

287. *Id.* at 890.

288. *See id.* at 890.

289. *See id.*

290. *Joiner*, 522 U.S. at 148.

American courtroom. Massachusetts courts are not prepared to deal with the issue of challenging expert testimony in the twenty first century without establishing procedural rules to govern these challenges and without implementing jury reform statewide. This must involve a concerted effort by the judiciary and the Massachusetts Bar.

Maryellen Ryan

IMS Technology, Inc. v. Haas Automation, Inc.: Does the Next Mutation[1] of Means-Plus-Function Claim Infringement Analysis Comport with the Purpose Envisioned by the Drafters of Section 112, Paragraph 6?

"[Paragraph six of section 112] (the scope clause in particular) has presented its share of interpretative and logical difficulties . . . . The resolutions to these difficulties have serious and often dispositive consequences."[2]

I. INTRODUCTION

Judicial methods for determining structural equivalents of means-plus-function claims for the purpose of finding literal infringement under 35 U.S.C. § 112, ¶ 6 are evolving.[3] All patent claims must conform to certain

---

1. Contrary to common thinking, a mutation does not always create an undesirable result. Mutations sometimes occur during meiosis (reductive cell division to produce reproductive cells) when the crossing over and exchange of a segment of alleles takes place between two homologous chromosomes. *See* ROBERT F. WEAVER & PHILIP F. HENDERICK, GENETICS 120-23 (3d. ed. 1997). The analogy between genetic mutation and the development of means-plus-function claim infringement case law seems especially fitting in the context of the *IMS Technology* case because once again a court borrows from the Doctrine of Equivalents, a somewhat homologous doctrine, to develop judicial tools for determining literal infringement of means-plus-function patent claims. *See infra* Part II.B. (discussing the *IMS Technology* case and other areas of means-plus-function law for which the courts have borrowed from the Doctrine of Equivalents). The analogy is also suitable here because it helps to cast a shadow of skepticism that should accompany all developments in case law. *See infra* note 349.

2. Rudolph P. Hofmann, Jr. & Edward P. Heller, III, *The Rosetta Stone for the Doctrines of Means-Plus-Function Patent Claims*, 23 RUTGERS COMPUTER & TECH. L.J. 227, 232 (1997).

3. *See infra* Part II.A.2.c. (discussing infringement analysis of means-plus-function claims); *see also infra* note 209 and accompanying text (discussing the habit

statutory standards.[4] General policy considerations support these statutory requirements.[5] Means-plus-function claims are a special category of patent claims that were intended by Congress to provide patentees with flexibility when designing organization schemes involving the interaction of structural components.[6] Section 112, ¶ 6 furthers the drafters' purpose because it does not require a patentee to recite structure within each means claim.[7] Instead, a patentee claims the innovative function she wishes to patent in a means claim that references a corresponding structure identified elsewhere in the patent.[8] A means-plus-function claim protects the patentee from devices that use structures identical to or structurally equivalent to the corresponding structures to perform the same function as the claimed structures.[9] Part II.A. discusses these background issues.[10]

Parts II.B.2.c. and III.C.2. of this Comment discuss the *IMS Technology* court's analysis of structural equivalents for infringement of a means-plus-function claim.[11] Parts III.A. and III.B. discuss the propriety, in terms of the drafters' goals and policy considerations, of borrowing from the judicial applications of the Doctrine of Equivalents to develop a contextual analysis for means-plus-function claim infringement.[12] Part III.C.3. addresses the ambiguities of the *IMS Technology* court's contextual analysis test and suggests that the proper scope for a contextual analysis to determine structural equivalents of means-plus-function claims involves evaluating the particular physical characteristics of the corresponding structure with respect to the overall function of the invention.[13] Part III.D. evaluates how this Comment's proposal affects the level of uncertainty inherent within means-plus-function claims.[14] Finally, Part III.E. analyzes the *Sunrise Medical* court's approach to structural equivalency of means-plus-function claims and illustrates not only the need for a clear test for contextual structural equivalency analyses of means-plus-function claims, but also how this Comment's proposed contextual analysis will work when

---

courts have of borrowing from the Doctrine of Equivalents to analyze literal infringement of means-plus-function claims).

4. *See infra* Part II.A.1. (discussing patent drafting).

5. *See infra* Part III.C.3. (discussing general policy considerations of patent law).

6. *See infra* notes 201-03 and accompanying text (discussing the basic purpose behind §112, ¶ 6).

7. *See infra* Part II.A.1. (discussing drafting of means-plus-function claims); *see also infra* Part II.A.2.b. (discussing judicial construction of means-plus-function claims).

8. *See infra* Part II.A.1. (discussing drafting of means-plus-function claims).

9. *See infra* Part II.A.2. (discussing infringement analysis of means-plus-function claims).

10. *See infra* Part II.A. (discussing pre-*IMS Technology* patent law).

11. *See infra* Parts II.B.2.c. & III.C.2.

12. *See infra* Parts III.A. & III.B.

13. *See infra* Part III.C.3.

14. *See infra* Part III.D.

applied.[15]

II. DEVELOPMENT OF MEANS-PLUS-FUNCTION PATENT CONSTRUCTION

A. Pre-*IMS Technology*[16] Patent Law

1. Fundamentals of Patent Drafting

The manner in which a patent application is drafted is important because the validity and infringement of patents are determined through the analysis of patent applications.[17] A valid patent must be drafted such that it provides adequate notice to all others as to what is being claimed to be protected from duplication.[18] Federal law mandates how a patent application must be drafted.[19] In order to be a valid patent, the application must provide a written description called a "specification."[20] The specification must be written so that "one having ordinary skill in the technology of the invention can make and use the invention."[21] A patent must also contain "claims," which are sentences that explicitly indicate "the subject matter which the applicant regards as his invention."[22] Claims contain "elements" or "limitations," which are the individual components of the claimed invention or process.[23] The United States Supreme Court has interpreted the requirements of a patent to include a specification that describes the invention with enough detail to satisfy § 112, ¶ 1 and at least one claim that points out and claims the subject matter of the invention.[24] A typical claim should define the scope of the invention by including descriptions of the process, machinery, material and the like, but not the function or result.[25]

Notwithstanding the typical claim format that recites the structure of the

---

15. *See* Sunrise Medical HHG, Inc. v. Airsep Corp., 95 F. Supp. 2d 348 (W.D. Pa. 2000). *See infra* Part III.E.

16. IMS Tech., Inc. v. Haas Automation, Inc., 206 F.3d 1422 (Fed. Cir. 2000).

17. *See* Yoncha Lynn Kundupoglu, *The Law of Means-Plus-Function Language*, 28 AIPLA Q.J. 39, 41 (2000).

18. *See* Brooks v. Fiske, 56 U.S. 212, 214-15 (1853); *see also* Markman v. Westview Instruments, Inc., 517 U.S. 370, 373 (1996) (quoting McClain v. Ortmayer, 141 U.S. 419, 424 (1891)).

19. *See, e.g.*, 35 U.S.C. § 112, ¶ 6 (1994).

20. *See* Kundupoglu, *supra* note 17, at 40.

21. *Id. See also* 35 U.S.C. § 112, ¶ 1 (1994).

22. 35 U.S.C. § 112, ¶ 2 (1994).

23. *See* Kundupoglu, *supra* note 17, at 41.

24. *See* 35 U.S.C. § 112, ¶ 1 (1994). An invention must be described "in such full, clear, concise, and exact terms as to enable any person skilled in the art . . . to make and use the same." *Id. See also* Markman v. Westview Instruments, Inc., 517 U.S. 370, 373 (1996).

25. *See Markman*, 517 U.S. at 373.

invention within itself, 35 U.S.C. § 112, ¶ 6 allows a patent claim to be drafted using "means-plus-function language" or " functional language."[26] A "means" claim, which is one half of the means-plus-function format, recites a particular means to be used to achieve a particular function.[27] A "means" claim does not contain the structure of the indicated means.[28] Instead of existing as an invalid claim for its failure to recite a structure, § 112, ¶ 6 allows a structural claim (one that recites the machinery, material, etc.) to correspond to the recited function of the means claim.[29]

The most usual way to draft a patent in means-plus-function language is to use the words "means for" or "step for" followed by a reference to a corresponding structural claim.[30] The federal circuit courts have addressed the issue of whether a specific claim limitation should be construed pursuant to ¶ 6 in a series of patent infringement cases.[31]

2. Patent Infringement Litigation: From Patent Application to Patent Protection

a. Adjudicatory Infringement Claims

Patent infringement occurs "when a competitor makes, uses, sells, offers

---

26. *See* 35 U.S.C. § 112, ¶ 6 (1994).

> An element in a claim for a combination may be expressed as a means or step for performing a specified function without the recital of structure, material, or acts in support thereof, and such claim shall be construed to cover the corresponding structure, material, or acts described in the specification and equivalents thereof.

Id.

27. *See* 35 U.S.C. § 112, ¶ 6 (1994).

28. *See id.*

29. *See* Ishida Co., Ltd. v. Taylor, 221 F.3d 1310, 1312, 1314 (Fed. Cir. 2000) (finding that the district court did not err in its claim construction because the claim related to another structural claim, not a functional claim); *see also* C.R. Bard, Inc. v. United States Surgical Corp., 102 F. Supp. 2d 199, 212 (D. Del. 2000). If a claim fails to "recite sufficient structural limitation," or if it uses "purely functional terms," the court will find that the claim is drafted in means-plus-function language and § 112, ¶ 6 is thereby implicated. *See id.* The *Bard* court found § 112, ¶ 6 to be invoked because the claim at issue was sufficiently lacking in structural elements. *See id.* at 216.

30. *See* Laitram Corp. v. Rexnord, Inc., 939 F.2d 1533, 1535, 1538 (Fed. Cir. 1991); *see also* Greenberg v. Ethicon Endo-Surgery, Inc., 91 F.3d 1580, 1583 (Fed. Cir. 1996) (stating that most means-plus-function claims are drafted in the traditional means-plus-function language using terms like "means for" and "step for").

31. *See, e.g.*, IMS Tech., Inc. v. Haas Automation Inc., 206 F.3d 1422, 1429-30 (Fed. Cir. 2000) (finding ¶ 6 applied because the limitations "recite[d] a specified function to be performed rather than the structure, materials, or acts for performing the function").

to sell, or imports a product containing all the limitations recited in a valid patent claim."[32] A judicial determination that a patent has been infringed requires a finding that "every limitation set forth in a claim [is] in an accused product or process exactly or by a substantial equivalent."[33] As a result of this statutory mandate, "failure to meet a single limitation is sufficient to negate infringement of the claim[.]"[34]

Patent infringement analysis involves two basic levels of adjudication.[35] First, because the claim construction has been established as a question of law properly left to the determination of the judge,[36] the court must determine the meaning of the claims.[37] After the court has determined the meaning of the terms, has construed the claim and patent accordingly, and has found there to be a genuine issue of material fact, the jury performs step two by comparing the claimed devices to the accused device to decide the factual issues of equivalents and ultimately, of infringement.[38] The

---

32. Kundupoglu, *supra* note 17, at 41. *See also* 35 U.S.C. § 271 (Supp. 1998).

33. Johnston v. IVAC Corp., 885 F.2d 1574, 1577 (Fed. Cir. 1989). The *Johnston* court identifies this maxim as the "All Elements Rule" or the "All Limitations Rule." *See id.* at 1577 n.3; *see also Laitram*, 939 F.2d at 1538.

34. *Laitram*, 939 F.2d at 1535.

35. *See* Cybor Corp. v. FAS Tech., Inc., 138 F.3d 1448, 1454 (Fed. Cir. 1998) (en banc) (discussing the two steps of patent infringement litigation); *see also* Symbol Tech., Inc. v. Opticon, Inc., 935 F.2d 1569, 1574 (Fed. Cir. 1991).

36. *See* P.M. Palumbo v. Don-Joy Co., 762 F.2d 969, 974 (Fed. Cir. 1988). The *Palumbo* court stated that, while construction of patent claims was ordinarily a question of law, determining the meaning of disputed terms required consideration of extrinsic evidence. *See id.* As a result, claim construction that included the explanation of disputed terms became a factual determination that could be properly decided by the judge or by the jury with an "appropriate instruction." *See id.* Since *Palumbo*, the United States Supreme Court has ruled that patent construction is always a decision of law properly delegated to the court. *See* Markman v. Westview Instruments, Inc., 517 U.S. 370, 372 (1996) (The "construction of a patent, including terms of art within its claim is exclusively within the province of the court."). The *Markman* court concluded that patent construction is not a jury decision because, based on a historical analysis, there existed no right to a jury at the time the Constitution was adopted to be preserved by the 7th Amendment. *See id.* at 372, 378-82 (finding that, historically, judges construed written documents and later, as patent law developed, specific terms of patents). As a result, appellate courts review district court construction determinations de novo. *See Cybor Corp.*, 138 F.3d at 1451, 1454; *see also Markman*, 517 U.S. 370; *see also* Ishida Co., Ltd. v. Taylor, 221 F.3d 1310, 1315 (Fed. Cir. 2000).

37. *See Cybor Corp.*, 138 F.3d at 1454; *see also Palumbo*, 762 F.2d at 974 (stating that to analyze infringement, the court must determine the scope of the claim and then the trier of fact determines if the claims "cover" the accused device).

38. *See Cybor Corp.*, 138 F.3d at 1453; *see also Symbol Tech.*, Inc., 935 F.2d at 1574; *see also Palumbo*, 762 F.2d at 974.

patentee must persuade the jury on the factual question of infringement[39] by a preponderance of the evidence.[40]

b. Principles of Patent Construction

In all three types of patent infringement litigation that this Comment will address, literal infringement of a means-plus-function claim,[41] literal infringement of a non-means-plus-function claim,[42] and infringement under the Doctrine of Equivalents,[43] the first task of the court is to determine how the patent at issue is to be construed.[44] Patent construction requires analysis of the terms used to describe the invention[45] and a resulting determination of the scope the claim covers.[46] The court must also determine if § 112, ¶ 6 is implicated.[47]

---

39. *See Symbol Tech.*, 935 F.2d at 1574.

40. *See id.*; *see also* Laitram Corp. v. Rexnord, Inc., 939 F.2d 1533, 1535 (Fed. Cir. 1991). *See* Lawrence M. Sung, *Echoes of Scientific Truth in the Halls of Justice: The Standards of Review Applied by the United States Court of Appeals for the Federal Circuit in Patent-Related Matters*, 48 AM. U. L. REV. 1233, 1278-79 (1999).

41. *See infra* notes 72-74 and accompanying text.

42. *See infra* Part II.A.2.c.i.

43. *See infra* Part II.A.2.c.ii.

44. *See* Cybor Corp. v. FAS Tech., Inc., 138 F.3d 1448, 1454 (Fed. Cir. 1998) (en banc) (discussing the two steps of patent infringement litigation); *see supra* notes 35-38 and accompanying text.

45. *See infra* notes 49-55 and accompanying text.

46. *See Cybor Corp.*, 138 F.3d at 1457 (describing how certain term definitions and structural equivalents can be excluded from the scope of a patent by prior judicial rulings and concessions); *see also* C.R. Bard, Inc. v. United States Surgical Corp., 102 F. Supp. 2d 199, 210 (D. Del. 2000) (discussing how prosecution history and prior art can supplement the plain meaning of a patent's terms). It is important to limit the scope of a patent by its prosecution history because "[t]he public has the right to rely on statements made by the patent applicant or his attorney during prosecution that define the scope of the claims." *Id.* Such reliance is permissible because patents are required to put others on notice as to what uses and developments are barred by the patent. *See* Markman v. Westview Instruments, Inc., 517 U.S. 370, 373 (1996) (stating that a patent must describe the scope of the invention exactly to protect the patentee and to give notice to the public of what inventions are covered by the patent); *see also* 35 U.S.C. § 112, ¶ 1 (1994). When determining the scope of a claim, the court must consider the claim language, the patent's specification, the prosecution history, the language of other claims within the patent, and expert testimony regarding the likely meaning of terms within the claim. *See* P.M. Palumbo v. Don-Joy Co., 762 F.2d 969, 975 (Fed. Cir. 1988). Expert testimony introduced and considered by the court on the meaning of patent terms is permissible under Federal Rule of Evidence 704. *See* Symbol Tech., Inc. v. Opticon, Inc., 935 F.2d 1569, 1575 (Fed. Cir. 1991).

47. *See* 35 U.S.C. § 112, ¶ 6 (1994); *see also Cybor Corp.*, 138 F.3d at 1454, 1457.

General rules of construction have been developed through patent validity and infringement litigation to define the scope of patents. Patents are to be construed "based on patent documents themselves."[48] When construing a patent, terms are given meaning with respect to the context of the entire patent document.[49] Expert witness testimony can be used to help the court define the terms of the patent without the need for a jury to assess credibility.[50] The court can also appropriately consider dictionary definitions of ordinary terms.[51]

When the terms of a patent are susceptible to more than one possible interpretation, the patent should be construed so that the patent will be valid.[52] Unless the specification or prosecution history indicate otherwise, words will usually be given their ordinary meaning.[53] The terms of a claim may only be limited by reference to a specification "when the patentee has set forth an explicit definition of a claim term in the specification, or when the claim terms chosen by the patentee are so ambiguous as to deprive the claim of clarity."[54] Claims can be construed beyond the scope of the specification but the meaning of terms should be limited to meanings that are compatible with the context of the specification.[55]

---

48. *Cybor Corp.*, 138 F.3d at 1454 (quoting Markman v. Westview Instruments, Inc., 52 F.3d 967, 981 (Fed. Cir. 1995)); *see also* Hofmann & Heller, *supra* note 2, at 252.

49. *See Markman*, 517 U.S. at 389. "[Terms] can be defined only in a way that comports with the instrument as a whole." *Id.*; *see also* Scripps Clinic & Research Foundation v. Genetech, Inc., 927 F.2d 1565, 1580 (Fed. Cir. 1991).

50. *See Cybor Corp.*, 138 F.3d at 1455.

51. *See id.* at 1459 (discussing how the lower court's use of the Webster Dictionary definition of "to" was appropriate in interpreting the term).

52. *See* IPPV Enters., LLC v. Echostar Communications Corp., 106 F. Supp. 2d 595, 603 (D. Del. 2000).

53. *See id.* at 603-04 (finding the applicant's view that the encryption technique was independent of the invention and type of signal used important in the construction of the terms of the claim). Evaluation of prosecution history involves examination of the record before the United States Patent and Trademark Office (USPTO) when the patentee was attempting to patent his invention. *See* Sung, *supra* note 40, at 1281. A patentee will not be allowed to characterize his invention differently during litigation than he did during prosecution of the patent. *See* James W. Dabney, *Fundamentals of Patent Litigation*, *in* c909 ALI-ABA COURSE OF STUDY, FUNDAMENTALS OF PATENT LAW AND PRACTICE 277, 293-94 (1994).

54. *IPPV Enters.*, 106 F. Supp. 2d at 602.

55. *See* Maytag Corp. v. Whirlpool Corp., 95 F. Supp. 2d 888, 891 (N.D. Ill. 2000). In *C.R. Bard*, the plaintiff argued that allowing additional limitations from the specification to define the clear terms of the claims would impermissibly limit the scope of the disputed claims to the exactly described devices. *See* C.R. Bard, Inc. v. United States Surgical Corp., 102 F. Supp. 2d 199, 213 (D. Del. 2000) (finding that meaning of terms should be taken in the context of the specification because the disputed claims lacked sufficient structural recitation and had to be construed as means-

Usually a means-plus-function claim includes the term "means for" or "means" within the language of the claim along with a reference to a corresponding structure in another specification.[56] However, the presence or absence of "means" language is not dispositive on the issue of whether a patent should be construed as a means-plus-function claim.[57] If the key language is not used, there arises a presumption that § 112, ¶ 6 does not apply.[58]

The recitation of some structure within a claim, with or without the "means" language, does not automatically bar § 112, ¶ 6 from applying.[59] Conversely, the presence of the word "means"[60] or a device that takes its name from the function it performs does not automatically mandate the application of § 112, ¶ 6.[61] However, absent specific "means" language, a court will find that § 112, ¶ 6 is implicated if the claim at issue does not identify a particular structure within itself but instead describes a means for performing a specific function with reference to another claim describing the corresponding structure.[62]

---

plus-function claims).

56. *See C.R. Bard*, 102 F. Supp. 2d at 212; *see also* Mas-Hamilton Group v. LaGard, Inc., 156 F.3d 1206, 1211 (Fed. Cir. 1998).

57. *See* Laitram Corp. v. Rexnord, 939 F.2d 1533, 1536 (Fed. Cir. 1991); *see also* Greenberg v. Ethicon Endo-Surgery, Inc., 91 F.3d 1580, 1583 (Fed. Cir. 1996).

58. *See C.R. Bard*, 102 F. Supp. 2d at 212; *see also* Personalized Media Communications v. International Trade Comm'n., 161 F.3d 696, 703-04 (Fed. Cir. 1998).

59. *See Laitram*, 939 F.2d at 1536 (stating that the structure described in the disputed limitation served to explain what the means clause did, not to explain the structure of the device, so the means claim was lacking a sufficient structural recitation to preclude application of § 112, ¶ 6).

60. *See Greenberg*, 91 F.3d at 1584. The term "detent means" was used in the patent's specification as a short hand reference to the structure described in the claim not as an expression of how a function was to be performed. *See id.* at 1583-84.

61. *See id.* at 1583 (concluding that a finding by the trial court that "detent mechanism" takes its name from the "detent" function it performs is an insufficient basis for invoking § 112, ¶ 6). The court should consider the dictionary definition of the device and the definition commonly understood in the art to aid it in deciding whether a device that takes its name from its function is a structure or a function. *See id.*

62. *See* Ishida Co., Ltd. v. Taylor, 221 F.3d 1310, 1312 (Fed. Cir. 2000) (finding that the district court did not err in its claim construction because the claim related to another structural claim, not a functional claim); *see also* C.R. Bard, Inc. v. United States Surgical Corp., 102 F. Supp. 2d 199, 212 (D. Del. 2000). If a claim fails to recite "sufficient structural limitations" or uses a "purely functional term," the court will find that the claim is drafted in means-plus-function language and § 112, ¶ 6 is thereby implicated. *See id.* The *Bard* court found § 112, ¶ 6 to be invoked because the claim at issue was sufficiently lacking in structural elements. *See id.* at 216. *See also* Maytag Corp. v. Whirlpool Corp., 95 F. Supp. 2d 888, 900 (N.D. Ill. 2000) (stating

When two limitations that are structurally different from one another are contained in a single claim corresponding to a functional claim, the claim should be construed narrowly so as not to include both devices unless the court finds the two devices are alternates.[63] However, courts try not to limit the scope of a claim to a "preferred embodiment" or any specific example contained within a corresponding specification.[64] Similarly, the court can not read limitations from other claims into an independent claim.[65] The Federal Circuit has stated that "claim differentiation," a "judicially developed guide" for claim construction, cannot trump the statutory requirements of § 112, ¶ 6.[66] Means claims will be construed to include and be limited by their corresponding structural claims even when such construction would create similarities between or overlap of two claims within the same patent.[67]

Once the court determines § 112, ¶ 6 applies, the court must construe a means-plus-function limitation by identifying the claimed function and the corresponding structure or act in the specification.[68] The court will usually examine the language of the means claim and consider statements made by the patentee during litigation and the application process to determine the proper function and corresponding structure.[69] Section 112, ¶ 6 may be

---

that if no structure is recited in a claim, the claim is governed by 35 U.S.C. § 112, ¶ 6). *See also* Symbol Tech., Inc. v. Opticon, Inc., 935 F.2d 1569, 1574 (Fed. Cir. 1991) (finding § 112, ¶ 6 applied because the claim was described as a means of performing a particular function).

63. *See Ishida*, 221 F.3d at 1316 (discussing how the notice function of patents would be defeated by a broad construction including all devices within a claim and stating the court's duty under § 112, ¶ 6 is to merely identify the structures that perform the specific function). *See also* Serrano v. Telular Corp., 111 F.3d 1578, 1583 (Fed. Cir. 1997) (holding that alternative structures within a corresponding structure claim must be included in the scope of the patent); *see also* Micro Chem., Inc. v. Great Plains Chem. Co., 194 F.3d 1250, 1258-59 (Fed. Cir. 1999). The *Ishida* court interprets *Serrano* and *Micro Chem.* as requiring only the identification of alternative structures within a claim without requiring the claim to be construed to include both identified alternative structures. *See Ishida*, 221 F.3d at 1316.

64. *See C.R. Bard*, 102 F. Supp. 2d at 210.

65. *See* P.M. Palumbo v. Don-Joy Co., 762 F.2d 969, 977 (Fed. Cir. 1988). An independent claim is one that recites structure within it and is not linked by reference to a means claim. *See* Markman v. Westview Instruments, Inc., 517 U.S. 370, 375 (1996).

66. *See* Laitram Corp. v. Rexnord, Inc., 939 F.2d 1533, 1538 (Fed. Cir. 1991). The doctrine of claim differentiation provides guidance for judges when construing patent claims by suggesting that claims within a patent should be interpreted, whenever possible, to describe different "inventions" from one another. *See id.*

67. *See id.* at 1538.

68. *See, e.g.*, IMS Tech., Inc. v. Haas Automation, Inc., 206 F.3d 1422, 1431 (Fed. Cir. 2000); *see also infra* Parts II.B.2.c.i & II.B.2.c.ii.

69. *See IMS Technology*, 206 F.3d at 1431.

found to apply to "[t]hat part of a claim contain[ing] means-plus-function language" without applying to the entire claim or entire patent.[70] If means-plus-function analysis is appropriate for a given claim, that claim will be limited to the specific means or function recited.[71]

c. Infringement

i. Literal Infringement

After the court has determined the proper construction of the patent document, the fact finder must compare the allegedly infringing invention or process with the construed claim to determine whether there is a "literal infringement" or infringement under the "Doctrine of Equivalents."[72] In general, a device has literally infringed a patented device when the accused device is identical to the patented device.[73] If an allegedly infringing invention or process performs an identical function to a valid claim and limitations, the court determines if the accused device has the same structure, materials or acts as described in the specification of the valid patent.[74] However, literal infringement can be found under § 112, ¶ 6 even when the structures of the two devices are not identical.[75] Literal infringement under

---

70. *See* Johnston v. IVAC Corp., 885 F.2d 1574, 1580 (Fed. Cir. 1989).

71. *See* Ishida Co., Ltd. v. Taylor, 221 F.3d 1310, 1314 (Fed. Cir. 2000).

72. *See, e.g.*, Chiuminatta Concrete Concepts, Inc. v. Cardinal Indus., Inc., 145 F.3d 1301, 1309-10 (Fed. Cir. 1998); Dawn Equip. Co. v. Ky. Farms, Inc., 140 F.3d 1009, 1021-22 (Fed. Cir. 1998).

73. *See Cybor Corp.*, 138 F.3d at 1458 (discussing that the district court did not err in instructing the jury that the second pumping device means limitation "'refers to a structure identical to the structure disclosed in the specification of the patent, or the equivalent to that structure, which performs the [same] function'"). The district court's jury instruction allowed the jury to find "literal infringement" if one of two circumstances existed: the accused device was identical to the patented device or the accused device had the same function and equivalent structure as the patented device. *See id. See also* P.M. Palumbo v. Don-Joy Co., 762 F.2d 969, 974 (Fed. Cir. 1988). Literal infringement may be found if the "accused device falls within the scope of the asserted claims as properly interpreted." *Id.*

74. *See generally* IMS Tech., Inc. v. Haas Automation, Inc., 206 F.3d 1422 (Fed. Cir. 2000).

75. *See* Micro Chem., Inc. v. Great Plains Chem., 103 F.3d 1538, 1548 (Fed. Cir. 1997); *see also Cybor Corp.*, 138 F.3d at 1458. *But see Johnston*, 885 F.2d at 1580 ("Where a claim does not read on an accused device exactly, there can be no literal infringement."). The *Johnston* court later considered literal infringement under a means-plus-function analysis. *See id.* A means-plus-function infringement analysis necessarily requires consideration of equivalents, or, a "non-exact" reading of a claim. *See supra* Parts II.A.1. & II.A.2.b. By acknowledging that literal infringement can be found even for patents that are drafted using the means-plus-function language permitted in § 112, ¶ 6, the *Johnston* court contradicts its strict literal infringement inter-

these circumstances occurs when the accused device performs a function identical to the function of the patented device and has a structure equivalent to that described in the patent.[76]

Literal infringement analysis under § 112, ¶ 6 has been described as a "reverse doctrine of equivalents" because it works to limit the literal meaning of a claim's language to the corresponding means specification of the patent application.[77] Therefore, a "means clause does not cover *every means* for performing the specified function."[78] However, while a means-plus-function construction limits the means to that described in the corresponding structure, the means-plus-function claim protects the patentee from any equivalent structures of those devices referenced in, and therefore corresponding to, the means claim.[79] It is important to note that, for the purposes of this Comment, a patent's scope is limited by the structures or means described in a claim and is not limited to the *product* that is described in the claim to be prepared by the process or device.[80] This means that the *function*, though sometimes referred to by courts as the *result*, is not described in terms of the product but instead in terms of the way in which the product is processed or produced.[81]

---

pretation. *See Johnston*, 885 F.2d at 1579; *see also* Symbol Tech., Inc. v. Opticon, Inc., 935 F.2d 1569, 1575. *See* 35 U.S.C. § 112, ¶ 6 (1998).

76. *See Micro Chem.*, 103 F.3d at 1548; *see also Ishida*, 221 F.3d at 1316-17 (stating that literal infringement of means-plus-function language requires a finding of identical function achieved by devices with structures that are "insubstantially different" from the corresponding structure). In *Ishida*, the court defined the "insubstantial difference" analysis as one in which the court identifies the way the function of the accused device is performed (the means) and the result accomplished by the performance (the function) and compares it to determine if the accused device's performance or result are "substantially different" from those of the patented device. *See id.* If the accused device performs in a "substantially different" way or achieves a "substantially different" result, there is no infringement. *See id. See also Laitram*, 939 F.2d at 1536 (stating that the patentee must show that the means described in the claim is the structural equivalent of the accused device if the structure of the accused device is different from the patent's corresponding structural claim).

77. *See Johnston*, 885 F.2d at 1580; *see also Ishida*, 221 F.3d at 1314 (discussing that if means-plus-function construction is not implicated, a claim is not limited to any means or function); *see Symbol Tech.*, 935 F.2d at 1575 (stating that the scope of a means-plus-function claim is limited by its coordinating structure claim and equivalents); *Laitram*, 939 F.2d at 1536; Greenberg v. Ethicon Endo-Surgery, Inc., 91 F.3d 1580, 1582 (Fed. Cir. 1996).

78. *Laitram*, 939 F.2d at 1536 (emphasis in original).

79. *See* P.M. Palumbo v. Don-Joy Co., 762 F.2d 969, 974 (Fed. Cir. 1988).

80. *See* Scripps Clinic & Research Found. v. Genetech, Inc., 927 F.2d 1565, 1583 (Fed. Cir. 1991).

81. *See id.* The function, which is the way in which a product or result is achieved, must be distinguished from the means, which is the way in which the function is achieved. *See id.*; *cf.* Markman v. Westview Instruments, Inc., 517 U.S. 370,

Even though the analysis of literal infringement of a means-plus-function claim allows some "structural equivalents" and "substantially same" structures to be infringing, the literal infringement cannot extend a claim to include developments in technology made after the date the patent is issued.[82] For this reason, an "equivalent structure" under the Doctrine of Equivalents will not necessarily satisfy an "equivalent structure" under a literal infringement analysis of a means-plus-function claim.[83] However, equivalents under literal infringement analysis of a means-plus-function claim and equivalents under the Doctrine of Equivalents do share some similarities.[84] For example, basic identification of equivalents in both literal infringement under § 112, ¶ 6 and under the Doctrine of Equivalents involves an evaluation of the reasonable "interchangeability" of the structures of the patented device and the accused device.[85]

In summary, while means-plus-function language allows a finding of

---

374 (1996). The *Markman* Court relies on two treatises which state that infringement is found when "the patent claim 'covers the alleged infringer's *product* or process.'" *See id.* (citing H. SCHWARTZ, PATENT LAW AND PRACTICE 80 (2d ed. 1995) and referring to 3 E. LIPSCOMB, WALKER ON PATENTS § 11.2, at 288-90 (3d ed. 1985)) (emphasis added). In this Comment, the term "product" will be limited to non-means-plus-function claims in which a particular device (the "product") is patented and manufactured. Means-plus-function claims can also be described in terms of what product is desired by reciting the means as a method for producing that particular product. *See* 35 U.S.C. § 112 ¶ 6 (1998). To minimize confusion, this Comment will use the term "function" and will avoid the term "product" when referring to means-plus-function claims.

82. *See* IPPV Enters., LLC v. Echostar Communications Corp., 106 F. Supp. 2d 595, 605-06 (D. Del. 2000) (distinguishing literal infringement of means-plus-function language from infringement analysis under the Doctrine of Equivalents).

83. *See IPPV Enters.*, 106 F. Supp. 2d at 605-06; *see also Palumbo*, 762 F.2d at 975 n.4 (describing equivalents under § 112, ¶ 6 and under the Doctrine of Equivalents as not "completely identical notions" because, as one reason, equivalents under § 112, ¶ 6 can lead to "literal infringement"); *see* Dawn Equip. Co. v. Ky. Farms, Inc., 140 F.3d 1009, 1021-22 (Fed. Cir. 1998) (stating that "the separate judicially-created doctrine of equivalents would have no application to those aspects of limitations drawn in means-plus-function form"); Katherine Kelly, *Cases and Recent Developments*, 8 Fed. Cir. B.J. 151, 163 (1998) (stating that means-plus-function claims cannot be expanded by the Doctrine of Equivalents "in the [a]bsence of [a]fter-rising [t]echnology").

84. *See Palumbo*, 762 F.2d at 975 n.4.

85. *See* Graver Tank & Mfg. Co. v. Linde Air Prods. Co., 339 U.S. 605, 609 (1950) (stating that when considering whether two devices are structurally equivalent "[a]n important factor is whether persons reasonably skilled in the art would have known of the interchangeability of an ingredient not contained in the patent with one that was[]"); *see also Palumbo*, 762 F.2d at 975 n.4. (stating the importance of the interchangeability factor in the determination of equivalents is important in both infringement analyses).

literal infringement when an accused device contains the exact structure or a structural equivalent of the patented device, literal infringement of a means-plus-function claim can never be found unless the identical function is performed by the accused device or process.[86] Even when an accused device or process performs the same function as the patented device, there is no infringement of a means-plus-function claim if that function is performed in "a substantially different way."[87] Therefore, literal infringement of a means-plus-function claim requires a finding of identical function and a finding that each limitation is the same as or is the structural equivalent of, that is to say substantially similar to, the accused device.[88]

ii. Infringement under the Doctrine of Equivalents

Infringement under the Doctrine of Equivalents occurs when each limitation of a patent "is met in the accused device either literally or equivalently."[89] For *non*-means-plus-function claims, a court will consider infringement under the Doctrine of Equivalents if literal infringement is not found.[90] If a court finds each limitation of a patent is met literally or equivalently, infringement results because the accused device, upon such findings, is determined to be "substantially" the same as the patented device.[91]

Unlike literal infringement, including literal infringement of a means-plus-function claim (which by statutory definition includes infringement by structural equivalents), the scope of a patent under the Doctrine of Equivalents can include protection from "equivalent" devices after the date the patent was issued.[92] This means that later-developed technologies can

---

86. *See* Johnston v. IVAC Corp., 885 F.2d 1574, 1580 (Fed. Cir. 1989).

87. *See Graver Tank & Mfg. Co.*, 339 U.S. at 608-09 (describing "reverse doctrine of equivalents"); *see also* Scripps Clinic & Research Found. v. Genetech, Inc., 927 F.2d 1565, 1581 (Fed. Cir. 1991).

88. *See Scripps Clinic & Research*, 927 F.2d at 1581; *see also Palumbo*, 762 F.2d at 978; *see also* Laitram Corp. v. Rexnord, Inc., 939 F.2d 1533, 1536 (Fed. Cir. 1991).

89. *See* Cybor Corp. v. FAS Tech., Inc., 138 F.3d 1448, 1459 (Fed. Cir. 1998) (en banc).

90. *See Johnston*, 885 F.2d at 1580 (stating that if each limitation of a claim is not met exactly by the accused device, infringement can be found under the Doctrine of Equivalents if those limitations that were not met exactly are met equivalently); *see also Palumbo*, 762 F.2d at 975 n.4.

91. *See Johnston*, 885 F.2d at 1580-81. "The accused device may then be said to perform substantially the same overall work to achieve substantially the same overall result by substantially the same means." *Id.* Courts use the "insubstantial difference" analysis during infringement determination under the Doctrine of Equivalents when considering if each limitation, if not met literally by the accused device, is met by a substantial equivalent. *See* Ishida Co., Ltd. v. Taylor, 221 F.3d 1310, 1316 (Fed. Cir. 2000).

92. *See* IPPV Enters., v. Echostar Communications Corp., 106 F. Supp. 2d 595,

infringe a patent under the Doctrine of Equivalents.[93] Therefore, "equivalent structure" under the Doctrine of Equivalents is not necessarily synonymous with "equivalent structure" under literal infringement analysis of a means-plus-function claim because literal infringement requires infringement by "structural equivalents" existing at or before the time the patent was issued.[94]

B. *IMS Technology v. Haas Automation*: How the Court Leaped from Structural Equivalency to Contextual Equivalency in Means-Plus-Function Literal Infringement Analyses and Achieved a Sensible Result

1. The Prior Technology Upon Which the IMS Patent Improved, the IMS Patent, and the Invention of Haas.

IMS Technology, Inc. appealed to the Federal Circuit Court of Appeals seeking review of a decision from Federal District Court for the Eastern District of Virginia.[95] The district court had granted Haas Automation's motion for summary judgment of non-infringement of a patent that IMS Technology had been assigned by its parent corporation, Hurco (U.S. patent No. 4,477,754 [hereinafter the IMS patent]).[96] The IMS patent, allegedly infringed by a Haas Technology invention, contains both non-means-plus-function claims and means claims.[97]

---

606 (D. Del. 2000); *see also Laitram*, 939 F.2d at 1539 (stating that equivalents under the Doctrine of Equivalents are not limited to equivalents under § 112, ¶ 6). The Doctrine of Equivalents "extends beyond the literal scope of claims, and permits a patentee to exclude others from practicing later-developed technologies that are equivalent to the claimed device." *IPPV Enter.*, 106 F. Supp. 2d at 605; *see also supra* notes 82-85 and accompanying text.

93. *See IPPV Enter.*, 106 F. Supp. 2d at 606.

94. *See id.* at 605-06; *see also supra* Part II.A.2.b.i. One purpose of the Doctrine of Equivalents is to "'prevent 'pirating' of the patentee's invention.'" Scripps Clinic & Research Found. v. Genetech, Inc., 927 F.2d 1565, 1581 (Fed. Cir. 1991).

95. *See* IMS Tech., Inc. v. Haas Automation, Inc., 206 F.3d 1422, 1425 (Fed. Cir. 2000).

96. *See id.*

97. *See id.* The *IMS Technology* court uses the terms "apparatus claims" to refer to what has been earlier identified in this Comment as structural claims that are complete because they contain limitations describing the structure of the claimed invention within the claim. *See supra* note 30 and accompanying text. Such claims are not subject to § 112, ¶ 6 construction. *See supra* notes 26-31 and accompanying text. The *IMS Technology* court uses the term "method claims" to refer to means claims. *See supra* notes 28-29 and accompanying text. Method claims are those which describe the means for performing the function with no structural limitation within the claim so

The IMS patent describes an invention for controlling a machine tool.[98] The IMS patent is a technological advancement on the machine tool control technology that was employed in industrial settings at the time the patent was issued.[99] Prior to the advent and industrial use of the invention described by the IMS patent, the motion of machine tools was directed by devices that converted punch tape printouts into electrical signals.[100] These punch tape printouts represented numerical code instructions that were pre-programmed from blueprints in a location remote from where the machine tool operated and were fed into the converting device at the time of operation.[101] Therefore, industrial operation of machine tools required a programmer who read the blueprints and generated the punch tape printouts accordingly and an operator who fed the pre-programmed punch tape printouts into the control device of the machine tool.[102] The control device read the punch tape printouts' numerical codes and converted these codes into electrical signals that were sent to the machine tool's motors.[103] As the *IMS Technology* court points out, the pre-IMS patent technology was cumbersome and lacked fluidity of operation.[104]

The IMS patent describes an invention that allows both the programming step and actual operation to be performed by the operator at the machine tool location.[105] In this way, the machine tool control invention is less cumbersome than pre-IMS patent technology. The crux of the IMS patent invention is its interactive programming design.[106] The IMS patent

---

that § 112, ¶ 6 applies and requires the method claim to be construed to cover a corresponding structure and its structural equivalents. *See supra* notes 27-29 and accompanying text.

98. *See IMS Technology*, 206 F.3d at 1426. For example, the invention described by the patent could be used to control a milling machine to cut or drill. *See id.*

99. *See id.*

100. *See id.*

101. *See id.*

102. *See id.*

103. *See id.* The *IMS Technology* court provides a detailed description of the pre-IMS patent technology in its opinion including what programming languages were used to generate the punch tape printouts. *See id.* (providing "EIA standard" and "RS-274-D" as examples of "standard part programming language[s]" used by pre-IMS patent technology control devices). For our purpose, which is to understand the *IMS Technology* court's construction of the IMS patent and the court's means-plus-function infringement analysis, we need only be familiar with the basic operation of pre-IMS patent machine control technology.

104. IMS Tech. Inc. v. Haas Automation Inc., 206 F.3d 1422, 1426 (Fed. Cir. 2000).

105. *See id.*

106. *See id.* In the words of the patentee, IMS Technology, Inc., "the claims were 'intended to focus upon an interactive system in which the operator was sequentially asked questions and via the display the operator in response simply answered the questions rather than have to worry about how to program the device.'" *Id.* at 1433

machine tool control system features a computer system that aids the operator while programming each particular operation.[107] By use of a keyboard, the operator first selects the particular operational task he seeks to perform from a computer screen display of choices.[108] The operator next responds to operation parameter prompts that appear on the screen.[109] These operation parameter prompts seek to solicit from the operator all information necessary to tailor the selected operational task to the blueprint from which the operator is working.[110] The responses to the operation parameter prompts are stored in "data blocks" which are converted to electrical control signals during operation of the machine tool.[111]

The operation program is composed of the "data blocks" that store the responses to the operation parameter prompts. These "data blocks" are saved in a readily "alterable" form of computer memory such as random access memory (commonly known as RAM).[112] The storage of the "data blocks" in an alterable form is a key feature that allows the IMS patent invention to be an interactive programming control system. Use of an alterable memory form sets the IMS patent invention apart from pre-IMS patent technology's use of unalterable punch tape printouts.[113]

In addition to storing the control program "data blocks" in an alterable memory form, the IMS patent also describes how the program may be stored on tape cassettes, a fixed memory format.[114] Control programs can be transcribed *onto* the tape cassettes via tape cassette transport and tape cassette "peripheral interface adapter."[115] Control programs from tape

---

(quoting IMS Technology, Inc.'s statement during the prosecution of the IMS patent recorded in Joint App. at 319). The *IMS Technology* court relies heavily on the interactive programming feature of the IMS patent invention for its construction of the IMS patent's claims and for its infringement analysis. *See id.* at 1433, 1437. The *IMS Technology* court found statements made by IMS Technology, Inc., at the trial court level made for the purpose of "emphasizing the interactive programming technique of the invention[]" important in defining the scope of the term "data block." *Id.* at 1433. The *IMS Technology* court also concluded that the physical structure of the "interfacing means" was insignificant in the context of the whole invention, based on a prior conclusion that "[t]he invention ... permits interactive programming of a machine tool." *Id.* at 1437; *see also infra* note 150 and accompanying text.

107. *See id.* at 1426.

108. *See id.*

109. *See id.*

110. *See* IMS Tech. Inc. v. Haas Automation Inc., 206 F.3d 1422, 1426 (Fed. Cir. 2000). For example, if the selected operational task was drilling, the control system would ask through its screen prompts for parameters concerning the location, diameter and depth of the hole to be drilled.

111. *See id.*

112. *See id.*

113. *See supra* note 100 and accompanying text.

114. *See IMS Technology*, 206 F.3d at 1426.

115. *See id.* The tape cassette "peripheral interface adapter" (PIA) is the device

cassettes can also be transcribed *to* the alterable memory format.[116] The IMS patent does not specify any one alterable memory format for control program storage.[117]

The allegedly infringing Haas Automation, Inc. invention [hereinafter the Haas invention] also allows for interactive programming in the control of machine tools.[118] Like the IMS patent invention, the Haas invention contains an operation parameter prompt system for programming and also features alterable memory capabilities.[119]

The Haas invention uses Quickcode and Conversational Quickcode to aid the operator while programming the machine tool.[120] Under the Haas invention's Quickcode design, the operator chooses from a screen of short operation descriptions.[121] Upon selection of an operation description, the G- or M-code that corresponds with the operation appears on the screen.[122] The operator is then able to edit the G- and M-codes that appear so the control program conforms exactly to the desired operational parameters.[123] The Conversational Quickcode design differs from the Quickcode design because it provides parameter prompts to facilitate the G- and M-code editing.[124]

The control programs generated by the Haas invention's interactive programming system are stored in G- and M-code format.[125] During operation of the machine tool, the programs are converted from the G- and M-codes in which they are stored into a binary format which is then translated into the electrical signals that command the motors of the machine tool.[126]

---

that transfers information from the tape cassette transport to the alterable memory and vice versa. *See infra* notes 166, 172 and accompanying text. The tape cassette transport is the device that transfers information to or from the tape cassette. *See infra* notes 167, 173. Both the tape cassette transport and the PIA are needed to completely transfer information from tape cassette to alterable memory and back.

116. *See IMS Technology*, 206 F.3d at 1426.

117. *See id.*

118. *See id.*

119. *See id.*

120. *See* IMS Tech. Inc. v. Haas Automation Inc., 206 F.3d 1422, 1426 (Fed. Cir. 2000).

121. *See id.*

122. *See id.* at 1426. G-codes and M-codes are series of programming codes that are considered standards in the machine operation programming community. *See id.* During the pre-IMS patent technology era, programmers generated punch tape printouts that would instruct the machine tool according to a blueprint by programming in G- and M- codes that corresponded to the required machine tool commands. *See id.*; *see also supra* notes 100-04 and accompanying text.

123. *See id.* at 1427.

124. *See id.*

125. *See id.*

126. *See* IMS Tech. Inc. v. Haas Automation Inc., 206 F.3d 1422, 1427 (Fed. Cir. 2000).

Some of the Haas invention's control systems use a floppy disk drive as the format through which control programs are stored, while other Haas invention control systems can be connected to a personal computer's memory or other storage device.[127]

2. The *IMS Technology* Court's Analysis: Getting Inside the Minds of the Justices

a. Summary Judgment Granted: The District Court's Analysis and Conclusions

The infringement suit of the IMS patent was tried in the federal district court for the Eastern District of Virginia [hereinafter, the district court].[128] The district court concluded, based on the allegedly infringed claims, that § 112, ¶ 6 was implicated.[129] The district court then construed the disputed terms and claims of the IMS patent according to its understanding of means-plus-function infringement analysis.[130]

The district court found that the structural claim that corresponds with the "interface means" claim (claims 1 and 7) of the IMS patent provides in the written description that the interface function be performed by a tape cassette transport, in particular.[131] Based on this construction of the "interface means" claim, the district court ruled that, as a matter of law, the Haas invention which uses a floppy disk drive to transcribe control programs for operation did not literally infringe the IMS patent.[132] After finding that there was no literal infringement, the district court compared the IMS patent's tape cassette transport device with the Haas invention's floppy disk drive device to determine if the Haas invention infringed the IMS patent

---

127. *See id.* at 1426.

128. *See id.* at 1425.

129. *See id.* at 1428, 1430 (noting that the district court found § 112, ¶ 6 implicated based on how the disputed claims were drafted and that this finding was not challenged on appeal by either party); *see also* 35 U.S.C. § 112, ¶ 6 (1998); *see supra* Part II.A.2.b. (discussing the role of the court in determining as a question of law whether or not a claim is drafted in means-plus-function language and, as a result, whether or not § 112, ¶ 6 controls the scope of the allegedly infringed patent claim).

130. *See IMS Technology*, 206 F.3d at 1428. The *IMS Technology* court found error in the district court's means-plus-function analysis and reversed the district court's grant of Haas's motion for summary judgment. *See infra* Part II.B.3.

131. *See IMS Technology*, 206 F.3d at 1428; *see also supra* notes 114-17 and accompanying text (describing the IMS patent invention and its use of a tape cassette transport as a device to transcribe a control program stored in a fixed tape cassette format to the tape cassette transport which is then transcribed through another device, the peripheral interface adapter (PIA), into the alterable memory format).

132. *See IMS Technology*, 206 F.3d at 1428; *see also supra* Part II.A.2.c.i. (describing literal infringement under means-plus-function patent construction); notes 118-27 and accompanying text (describing the Haas invention).

under the Doctrine of Equivalents.[133] The district court concluded as a matter of law that the Haas invention did not infringe the IMS patent under the Doctrine of Equivalents either, and accordingly granted Haas Automation's motion for summary judgment of non-infringement of the "interface means" claim.[134]

The district court also construed the term "data blocks" of the IMS patent and considered whether or not the Haas invention infringed the IMS patent's display means claim (claims 1 and 11).[135] Again determining that § 112, ¶ 6 applied, the district court found that the IMS patent's display means claim corresponded to the "data block" structure.[136] The district court limited the definition of "data block" to the "specific variables and sequence of inquiries set forth in the written description of the [IMS] patent."[137] Because the IMS patent's "data block" claim does not specify the potential use of G- and M-codes in achieving its data block function, the district court determined that the IMS patent's display means claim did not include the use of G-and M-codes.[138] Based on its construction of the term "data blocks" and its finding that the IMS patent invention's display means claim corresponded to the "data block" structure, the district court concluded that the Haas invention's use of control prompts in G- and M-format did not literally infringe the IMS patent's display means.[139] Likewise, the district court concluded that the Haas invention's use of G- and M-codes in its prompting system did not infringe the "data block" claims of the IMS patent.[140]

---

133. *See IMS Technology*, 206 F.3d at 1428; *see also supra* Part II.A.2.c.ii. (discussing how a court must consider infringement under the Doctrine of Equivalents if the court does not find literal infringement).

134. *See IMS Technology*, 206 F.3d at 1428.

135. *See id.* at 1429; *see also supra* notes 105-11 and accompanying text (discussing how the IMS patent invention displays control parameter prompts to aid the operator in programming and stores the responses to these prompts in data blocks).

136. *See IMS Technology*, 206 F.3d at 1429.

137. *Id.*

138. *See id.*

139. *See id.*; *see also supra* notes 105-27 and accompanying text (describing the IMS patent invention's and the Haas invention's use of computer screen control parameter prompts to achieve an interactive control programming system).

140. *See IMS Technology*, 206 F.3d at 1429. The district court's conclusion that the Haas invention does not infringe the IMS patent even under the Doctrine of Equivalents (irrespective of the fallacies in the district court's legal analysis) is based on a very technical reading of the IMS patent. To conclude that an invention which uses a parameter prompting system to facilitate programming of machine tool controls that differs from the IMS patent merely by use of G- and M-codes and some other minor variation does not infringe the prompting system described in the IMS patent seems illogical. Perhaps that is why summary judgment motions in patent infringement litigation should be approached with great caution. *See, e.g.*, Johnston v. IVAC Corp., 885 F.2d 1574, 1577 (Fed. Cir. 1989).

b. The Appeal, the Reversal, and the Reasons

IMS Technology sought review of the district court's grant of summary judgment motion for non-infringement in the Federal Circuit Court of Appeals.[141] IMS Technology challenged the district court's ruling based on the district court's claim construction and its factual finding of structural non-equivalency for two means-plus-function claims.[142]

i. The Data Block Claim of the IMS Technology Patent.

The *IMS Technology* court agreed with the district court's decision that § 112, ¶ 6 applied to the data block means claim.[143] However, the *IMS Technology* court distinguished between those parts of the IMS patent's data block means claim that were to be construed according to § 112, ¶ 6 and those that were not.[144] The *IMS Technology* court found that the general function of the data block display means claim is to provide an interactive program within the IMS Technology patent's machine control system.[145] The *IMS Technology* court carefully construed the claim's language and determined the specific function of the data block display means claim as that of "'sequentially display[ing] *data block* inquiries.'"[146] Accordingly, the "data block" is not the function, nor is the "data block" the means by which the function is achieved.[147] As a result of its status as neither function nor means (corresponding structure), the circuit court found that § 112, ¶ 6 did not apply to the term "data block."[148] The *IMS*

---

141. *See* IMS Tech., Inc. v. Haas Automation, Inc., 206 F.3d 1422, 1425 (Fed. Cir. 2000).

142. *See id.* at 1428-29. IMS Technology appealed the district court's decisions involving the IMS Technology patent's interface means claim and its data block terminology. *See id.* The federal circuit's analysis for each means-plus-function claim in dispute is discussed separately in the reverse of the order in which the *IMS Technology* opinion addresses them.

143. *See id.* at 1432 (discussing the district court's finding that the method of claim 11 corresponded to the structure of the data block display means of claim 1 and again noting that neither party raised on appeal the issue of whether or not the data block claim was a means-plus-function claim).

144. *See id.*

145. *See id.* at 1434; *see also supra* notes 105-17 and accompanying text (discussing the IMS Technology patent).

146. *IMS Technology*, 206 F.3d at 1432 (emphasis added).

147. *See* IMS Tech. Inc. v. Haas Automation Inc., 206 F.3d 1422, 1432 (Fed. Cir. 2000). The "data blocks" were for storing the display parameters and therefore were neither the function nor display means. *See id.*

148. *See id.* at 1432-33; *see also* 35 U.S.C. § 112, ¶ 6 (1994). *See supra* Part II.A.2.b. (discussing how patent claims can be partially construed as means-plus-

Technology court concluded that the term "data block" is merely part of the written description of the data block display means claim and to limit the means to the specific sequence of variables recited, as the district court did, "would . . . [impermissibly] read a particular embodiment into the claim."[149]

According to the *IMS Technology* court, "[IMS Technology's] invention is an apparatus and method for interactively creating a program for controlling the machine tool."[150] The court went on to conclude that "[i]t is irrelevant to the claimed invention whether a data block is stored as a line of G- and M-code, in a binary format, or in any other format."[151] The *IMS Technology* court further concluded that, based on its findings, the district court erred by limiting the data block term of the IMS patent to the specific set of variables and sequences described in the patent.[152] In so concluding, the *IMS Technology* court found that the Haas invention, though it used G- and M-codes,[153] could be found by the trier of fact to infringe the IMS patent.[154]

ii. The Interface Means Claim of the IMS Patent.

With respect to the claim, IMS Technology first argued that the district court erred by finding the means language of the interface claim to correspond to the tape cassette transport.[155] IMS Technology then argued that even if the district court was correct in finding that the tape cassette transport corresponded to the interface means, the district court erred by concluding that the Haas invention's floppy disk drive was not a structural

---

function claims when appropriate).

149. *IMS Technology*, 206 F.3d at 1433; *see also supra* note 64 and accompanying text (discussing the impropriety of reading preferred embodiments into patent claims).

150. *IMS Technology*, 206 F.3d at 1434. The *IMS Technology* court identified the function of the IMS Technology patent partially on statements made by IMS Technology during its patent application process. *See id.* at 1433. In particular, the federal circuit focused on the following statement: "[T]he claims were 'intended to focus upon a novel interactive system in which the operator was sequentially asked instructions and via the display the operator in response simply answered the question rather than have to worry about how to program the device.'" *Id.* (quoting IMS Technology's statements from J.A. at 319).

151. *Id.* at 1434; *see also supra* notes 106-15 and accompanying text (discussing the IMS Technology patent's use of data blocks and prompts to aid the machine operator in programming the machine).

152. *See IMS Technology*, 206 F. 3d at 1434.

153. *See* IMS Tech. Inc. v. Haas Automation Inc., 206 F.3d 1422, 1434 (Fed. Cir. 2000). The IMS Technology patent did not specify the use of G- and M-codes in its data blocks. *See supra* text accompanying notes 105-17 (discussing the IMS Technology patent).

154. *See IMS Technology*, 206 F.3d at 1434.

155. *See id.* at 1430; *see also supra* note 131 and accompanying text.

equivalent of the IMS patent's tape cassette transport.[156] If the federal circuit was persuaded by either of IMS Technology's arguments, the federal circuit would have to reverse the district court's ruling.[157]

The *IMS Technology* court concluded that the district court had erred in its identification of the function performed by the interface means claims.[158] According to the *IMS Technology* court, the interface means on the IMS patent had one overall function.[159] The IMS patent's overall function is that of an "interface means for transferring a control program and control parameters from an external medium into said alterable memory . . . and medium."[160] The court determined that the Haas invention's system performed the same overall function as the interface means of the IMS Technology patent.[161]

After finding that the Haas invention performed an equivalent function to IMS Technology's patented system,[162] the *IMS Technology* court's next task was to compare the structure of the Haas invention to the structures recited in the IMS patent and their structural equivalents.[163] To identify the appropriate corresponding structure to the interface means claims of the IMS patent, the *IMS Technology* court more closely examined the function of the interface means.[164]

The *IMS Technology* court found that the overall function of interface means could be broken down into two distinct but conjunctive subfunctions.[165] The first function of the interface means is to record a program and its parameters "from an alterable memory form onto an external

---

156. *See IMS Technology*, 206 F.3d at 1430; *see also supra* notes 132-34 and accompanying text.

157. *See supra* notes 72-76, 82-84 and accompanying text (discussing the role of claim construction in patent infringement litigation and discussing when summary judgment of non-infringement is proper). Reversal of the district court's ruling would be required upon the finding of error of the district court's conclusions of law unless the error was harmless. *See* FED. R. CIV. P. 61. While the *IMS Technology* court did not explicitly consider whether or not the error to IMS Technology was harmless, we can assume that an erroneous ruling on a motion for summary judgment for non-infringement is not a harmless error.

158. *See IMS Technology*, 206 F.3d at 1439.

159. *See* IMS Tech. Inc. v. Haas Automation Inc., 206 F.3d 1422, 1430 (Fed. Cir. 2000).

160. *Id.* (quoting the decision of the district court).

161. *See id.* at 1435.

162. *See IMS Technology*, 206 F.3d at 1435.

163. *See id.*; *see also supra* notes 35-40 and accompanying text (discussing the two steps involved in patent infringement analysis of a means-plus-function claim, the second of which is to compare the structure of the accused device with the structure of the patented device); *see also* 35 U.S.C. § 112, ¶ 6 (1994).

164. *See IMS Technology*, 206 F.3d at 1430.

165. *See* IMS Tech. Inc. v. Haas Automation Inc., 206 F.3d 1422, 1430..

medium."[166] The second function of the interface means is to transfer the program that has been recorded by the first function "from the external medium [to which it has been recorded] into alterable memory."[167] Both functions must work together to achieve the overall function of transferring control programs back and forth between external alterable memory and the internal alterable memory or the immediate use of the control computer.[168]

Because both sub-functions are required for performance by the interface means, the *IMS Technology* court sought a structure, or two or more cooperating structures, recited in the IMS patent that would perform both of the sub-functions of the interface means claim.[169] By examining the IMS patent, the federal circuit concluded that the external medium described in the patent consists of the tape cassettes and that the recited alterable memory is the RAM memory of the operating computer.[170] Therefore, to perform the overall function of the IMS patent, the structure(s) that correspond to the interface means claim must record and transfer recorded programs and control parameters between tape cassette, the external medium, and the operating computer's RAM memory, the alterable memory.[171] The *IMS Technology* court concluded that in order to transfer and record between the tape cassette and RAM, both the tape cassette transport and the tape cassette peripheral interface adapter (PIA) must be "part of the disclosed structure corresponding to the 'transferring' function of the interface means."[172] While the recording sub-function of the inter-

---

166. *Id.*; *see also supra* notes 105-17 and accompanying text (discussing the IMS patent's invention and its ability to convert programs stored on tape cassettes (an external memory source which the *IMS Technology* court refers to as the external medium) to the operating computer's memory (an internal alterable memory described by the *IMS Technology* court as the alterable memory) via a tape cassette transport (a device external to and attached to the operating computer)).

167. *IMS Technology*, 206 F.3d at 1430; *see also supra* notes 105-17 and accompanying text (discussing the IMS patent).

168. *See IMS Technology*, 206 F.3d at 1430.

169. *See id.* at 1431.

170. *See id.*; *see also supra* notes 105-17 and accompanying text (detailing the IMS Technology patent); *see also supra* notes 158-61 and accompanying text (discussing the federal circuit's construction of IMS Technology patent and its function).

171. *See* IMS Tech. Inc. v. Haas Automation Inc., 206 F.3d 1422, 1431 (Fed. Cir. 2000); *see also supra* notes 112-13 and accompanying text.

172. *IMS Technology*, 206 F.3d at 1431; *see also supra* note 115 and accompanying text (describing the IMS Technology patent's recitation of a peripheral interface adapter (PIA)). The *IMS Technology* court found that the PIA only transfers between the cassette transport and the operating computer's memory. *See IMS Technology*, 206 F.3d at 1431. Because the transferring function requires transfer between the tape cassette and the computer's RAM, the tape cassette must also be part of the corresponding structure in order to complete a two step transfer, from RAM to tape cassette transfer to tape cassette and the reverse. *See id*; *see also supra* notes 105-17 and ac-

face means requires only the tape cassette transport, the whole function of the interface means is served by the corresponding claims of the IMS patent that describe the tape cassette transport and the PIA.[173] For the interface means claim, the *IMS Technology* court concluded that the district court did not err by including the tape cassette transport as the corresponding structure.[174]

The *IMS Technology* court had already decided that the Haas invention

---

companying text (describing the IMS Technology patent).

173. *See IMS Technology*, 206 F.3d at 1432. The *IMS Technology* court concluded that the recording function of the interface means is performed solely by the tape cassette transport. *See id.* However, because the overall function of the interface means requires recording and transferring, and because the transferring function requires two steps, the proper corresponding structure for the interface means includes both the tape cassette transport and the PIA. *See id.* at 1431-32; *see also supra* notes 112-17 and accompanying text (discussing the two step process of the transferring sub-function of the interface means).

174. *See IMS Technology*, 206 F.3d at 1432. The *IMS Technology* court rejected IMS Technology's argument that limiting the interface means function to the tape cassette transport violates the doctrine of claim differentiation. *See id.* at 1431; *see also supra* note 63 and accompanying text (discussing the judicially created doctrine of claim differentiation and its role in patent construction). IMS Technology argued that the doctrine of claim differentiation precluded a finding that the tape cassette transport is the corresponding structure for the interface means because another claim specifically recites the tape cassette transport. *See IMS Technology*, 206 F.3d at 1431. The *IMS Technology* court dismissed IMS Technology's argument by reasoning that the claim specifically reciting the tape cassette transport was narrower than the interface means because the interface means, as a means-plus-function claim, included structural equivalents to the tape cassette transport. *See id.*; *see also* 35 U.S.C. § 112, ¶ 6 (1994). Because the interface claim was necessarily broader than the other claim which recites the tape cassette transport, construing the interface means claim to include the tape cassette transport did not render the interface means claim the same as the other claim. *See IMS Technology*, 203 F.3d at 1431. The doctrine of claim differentiation is not violated unless two claims are construed so that they are the same as one another. *See supra* note 63 and accompanying text. The court's reasoning seems weak considering the possibility that the claim that recites the tape cassette transport can be infringed by an equivalent under the Doctrine of Equivalents. Arguably, the distinction of the breadth of the two claims is excusable because a means-plus-function claim is literally infringed by structural equivalents while non-means-plus-function claims require application of a separate doctrine to find infringement by equivalents. *See supra* Part II.A.2.c. (discussing infringement of means-plus-function claims). However, the federal circuit's superficial analysis on this issue is understandable under the case law that continuously reinforces the proposition that the doctrine of claim differentiation is not a fixed rule. *See supra* note 63 and accompanying text.

performed the same function as the IMS patent's interface means claim.[175] By comparing the structures of the IMS patent and the Haas invention, the federal circuit concluded that Haas Automation's system did not contain the same structure as the interface means.[176] At this point in the *IMS Technology* court's infringement analysis, only the factual issue as to whether Haas Automation's system contains structures that are the equivalent of the IMS patent's structures remained to be resolved.[177]

c. The Court's Analysis of Structural Equivalency: The Crux of the Opinion

The issue of structural equivalency before the *IMS Technology* court was whether or not a reasonable jury could find that the floppy disk drive of the Haas invention is the structural equivalent of the tape cassette transport and the PIA system of the IMS patent's interface means claim.[178] The federal circuit began its inquiry into whether or not the floppy disk drive is the structural equivalent of the tape cassette transport by analyzing whether the floppy drive performs the identified function[179] in substantially the same way to achieve substantially the same result as the tape cassette

---

175. *See supra* notes 160-61; *see also supra* notes 105-27 (discussing the IMS Technology patent and the Haas Automation invention).

176. *See IMS Technology*, 206 F.3d at 1435 (noting that Haas Automation's invention uses a floppy disk drive while the IMS Technology patent recites a tape cassette transport in conjunction with a PIA). The floppy disk drive of the Haas Automation invention functions to communicate between the external memory of a floppy disk and the internal memory of the control computer and thereby performs the same function as the IMS Technology patent's interface means claim. *See supra* notes 125-27 and accompanying text (discussing Haas Automation's invention and its use of a floppy disk drive through which control programs can be stored and retrieved).

177. *See* IMS Tech. Inc. v. Haas Automation Inc., 206 F.3d 1422, 1435 (Fed. Cir. 2000); *see also supra* notes 75-76, 92 and accompanying text (describing how literal infringement of a means-plus-function claim under § 112, ¶ 6 requires the accused device to either contain the same structure as or the structural equivalent to the structure of the patented claim once the court determines that the two devices perform the same function).

178. *See IMS Technology*, 206 F.3d at 1435 (noting that Haas Automation did not challenge on appeal that its device contains the structural equivalent of the IMS Technology patent's PIA, but only that its device does not use the structural equivalent of the tape cassette transport). If the court finds, as the *IMS Technology* court did, that a reasonable jury could find the two structures are structural equivalents, it must reverse the lower court's ruling of non-infringement as a matter of law. *See supra* notes 35-40 (discussing respective roles of judge and jury).

179. *See supra* notes 162-64 and accompanying text.

transport.[180] The *IMS Technology* court opined that, on appropriate facts, comparisons for the purpose of structural equivalency determinations do not have to focus on physical structure.[181] Instead, the *IMS Technology* court decided that structural equivalency, whether analyzed in terms of substantiality[182] or structure interchangeability,[183] should be determined by considering the "context of the invention . . . just as it is in a doctrine of equivalents determination."[184] By borrowing from the Doctrine of Equivalents, the *IMS Technology* court forged a new dimension into literal infringement analysis of means-plus-function patent claims.[185] The question before the federal circuit became: are the physical differences between the two devices "substantial in light of the role played by the 'interface means' in the claimed invention[?]"[186]

3. The Court's Conclusions and Holding

Under the correct claim construction and infringement analysis—according to its own analysis—the *IMS Technology* court concluded that summary judgment was improper.[187] First, the federal circuit found that the district court had erred in its construction of the data block display means claim.[188] Because the court found it unsuitable to limit the data block to the specific variables described in the patent, the federal circuit reversed the district court's ruling of summary judgment for non-infringement of the data block display means claim.[189] Secondly, while agreeing with the district court's construction of the interface means claim to include the tape cassette transport, found error in the district court's structural equivalency analysis.[190] The *IMS Technology* court determined that the physical characteristics of the Haas invention's floppy disk drive

---

180. *See supra* note 76 (discussing the substantiality test for equivalents).

181. *See IMS Technology*, 206 F.3d at 1436. The court took time to point out that "structural equivalent" is not synonymous with "equivalent structure" and that § 112, ¶ 6 requires only "structural equivalency." *See id.* at 1436 n.3.

182. *See supra* note 76 and accompanying text (discussing the substantiality test for equivalents).

183. *See supra* note 85 and accompanying text (discussing the interchangeability test for equivalents).

184. *IMS Technology*, 206 F. 3d at 1436.

185. *See id. See also supra* Part II.A.2.b.ii. (discussing the Doctrine of Equivalents and its role in infringement litigation involving non-means-plus-function claim).

186. *IMS Technology*, 206 F.3d at 1437.

187. *See id.* at 1425 (finding that "genuine issues of material fact concerning the infringement with respect to some of the accused devices" remain to be resolved).

188. *See supra* Part II.B.3.

189. *See id.* (discussing the construction of the term "data block").

190. *See supra* Part II.B.2.c. (discussing the analysis of the interface means claim).

and the IMS patent's tape cassette transport were unimportant when taken in the context of the entire function of the inventions.[191] Based on its observation that the physical characteristics of the recording-transferring device do not seem to be important to the function of the inventions, the *IMS Technology* court concluded that a reasonable jury could find a floppy disk drive to be the structural equivalent of a tape cassette transport.[192]

III. ANALYSIS

A. The Value of Function in Post-IMS Technology Means-Plus-Function Claims

1. The *IMS Technology* Court's Emphasis on Function

The *IMS Technology* court held that the importance of the physical characteristics of structures is to be considered in light of the function of the whole invention when determining structural equivalency.[193] The court's decision that structural equivalents of means-plus-function claims should be evaluated with respect to the context of the invention[194] emphasized function over structure.[195]

It is interesting that the court approached the task of defining the term "data block" in a manner similar to how it approached the structural equivalency analysis of the interface means claim.[196] Apart from the federal circuit's careful reading of the data block display means claim so that data block was neither function nor structure,[197] the court's analysis of the data block involved identifying the overall function of the IMS patent and

---

191. *See IMS Technology*, 206 F.3d at 1437 (noting that the function of the interface means is to "provide a way of storing programs created using the inventive programming apparatus and process"). The court again emphasizes that the overall function of the IMS Technology patent "is directed to an apparatus that permits interactive programming of a machine tool." *Id.*

192. *See id.*

193. *See id.* at 1435. *But see* Sunrise Med. HHG, Inc. v. Airsep Corp., 95 F.Supp.2d 348, 445 n.743 (W.D. Pa. 2000) (finding the physical difference between the structures of each device "affect the operation of the claimed method itself" and therefore the accused device does not infringe the patent).

194. *See supra* notes 191-92 and accompanying text (discussing the holding of the *IMS Technology* court).

195. The court ultimately decided that physical or structural differences between the patented device and the accused device are unimportant with respect to the common function of the adversary inventions. *See supra* notes 191-92 and accompanying text.

196. *See IMS Technology*, 206 F.3d at 1434, 1435-37.

197. *See supra* notes 143-51 and accompanying text (discussing the analysis of the data block display means claim).

the relative importance of the definition of "data block."[198] While the *IMS Technology* court's approach seems appropriate because terms within a patent must be construed with respect to the whole document,[199] this is a noteworthy illustration of the predominance of function in the court's infringement analysis.[200]

2. Is the Emphasis on Function, as Opposed to Structure, in a Court's Determination of Structural Equivalents Appropriate in the Literal Infringement Analysis of Means-Plus-Function Patent Claims?

Section 112, ¶ 6 of the United States Patent Code[201] was drafted to statutorily address the issue of patent claims that were being written as combinations of separate elements "brought together under a single cooperative principle."[202] The invention of such claims, now known as means-plus-function claims, is said to "lay in the combination itself, [or in] the relationship or association of elements."[203] So defined, it is easy to argue that a means-plus-function claim is really a patent on the organization and inventive interaction between elements of the means and less so the structure of each individual component of the means.[204] If this definition is accurate, it is the integration of many steps and structural components to achieve a goal that makes the means-plus-function claim what it is.

Organization and association among elements, though part of the "means" (which is structural by definition), which are more closely tied logically to the function than they are to physical characteristics of the individual devices.[205] Inventive association is a comprehensive idea that covers more than the physical and structural details of each piece of the whole. While developing an inventive association, the patentee surely has the desired function and result in mind while piecing together the elements

198. *See supra* notes 150-51 and accompanying text (discussing the court's conclusion that the program used in the data block was unimportant to the patent's function).

199. *See supra* Part II.A.1 (discussing patent construction).

200. *See infra* Part III.B (discussing the greater weight of function in the analysis).

201. 35 U.S.C. § 112 (1998).

202. Hofmann & Heller, *supra* note 2, at 259-60.

203. *Id.* at 279.

204. Common sense tells us that if the invention sought to be protected by a § 112, ¶ 6 claim is in the combination or relationship of elements used to achieve a function, the structure of each component must be less important than the overall function the organization has been designed to accomplish. It seems the patentee would start with the desired function in mind and find a suitable combination of components to achieve that goal. While non-means-plus function claim inventions must begin in a similar way, the structure serving a particular result, means-plus-function inventions differ because it is not the individual structure that is novel but instead is the whole system of interacting structures that is unique. *See id.* at 259-60, 277.

205. *See supra* note 204.

that will work together to achieve the ends.[206] If this is true, the *IMS Technology* court's emphasis on function over structure finds support in the intention of the drafters of § 112, ¶ 6.[207]

3. Does the Borrowed Contextual Equivalency Test Achieve the Distinct Goals of Section 112, Paragraph 6?

Even if the emphasis on function in means-plus-function claims is appropriate, has the *IMS Technology* court correctly injected the Doctrine of Equivalents use of context to determine equivalency into the structural equivalency determination for literal infringement of means-plus-function claims?[208] The Doctrine of Equivalents, as a distinct doctrine,[209] is applicable in non-means-plus-function infringement suits when the court fails to find literal infringement.[210] While §112, ¶ 6 includes an equivalency clause, equivalency under means-plus-function claims has a purpose separate from that of the Doctrine of Equivalents.[211] Speaking generally, the structural equivalency clause of §112, ¶ 6 is designed to "'benefit the inventor's genius and not the scrivener's talents.'"[212] The structural equivalency clause does so in infringement litigation of means-plus-function claims by limiting the scope of the claim to structures that match the overall invention.[213] The effect of the scope limit inherent within the equivalency clause is to protect other inventors from patents that, due to their clever drafting, would provide sweeping protection to the clever patentee of a technically outmoded device.[214] Alternatively, there are instances in which structural equivalency

---

206. *See supra* note 204.

207. *See infra* notes 244-54 and accompanying text (discussing the *IMS Technology* court's emphasis on function); *see also infra* text accompanying notes 260-62 (discussing the goal of the drafters).

208. *See supra* Part II.B.2.c (discussing the *IMS Technology* court's opinion and its decision to borrow from the Doctrine of Equivalents to determine whether the Haas Automation invention could be a structural equivalent of the IMS Technology patent's interface means claim); *see also supra* Part II.A.2.c.ii (discussing infringement under the Doctrine of Equivalents for non-means-plus-function claims).

209. The *IMS Technology* court is not the first court to have borrowed from the Doctrine of Equivalents. *See* Hofmann & Heller, *supra* note 2, at 233 (discussing how courts have borrowed from the Doctrine of Equivalents when interpreting the meaning of "equivalents" under §112, ¶ 6).

210. *See supra* Part II.A.2.c.ii (defining the Doctrine of Equivalents and its application).

211. *See* Hofmann & Heller, *supra* note 2, at 234, 280-81.

212. *Id.* at 280 (quoting Autogiro Co. of Am. v. United States, 384 F.2d 391, 399-400 (Ct. Cl. 1967)).

213. *See id.* at 292.

214. *See id.* at 280; *see also supra* notes 82-83 and accompanying text (discussing how, unlike the Doctrine of Equivalents, structural equivalency for means-plus-function claims does not cover later developed technology).

provides *more* protection to the patentee than would be provided by a mere literal interpretation of the written means-plus-function claims.[215] To truly capture the intent of the drafters, structural equivalents must be determined so that they comply with the "internal logic" of the patent.[216]

Despite the different purposes behind the Doctrine of Equivalents and the structural equivalency provision of § 112, ¶ 6, the contextual equivalency test borrowed from the Doctrine of Equivalents seems to be a well-suited application to means-plus-function claims. The contextual equivalency test requires an allegedly equivalent structure to be viewed in the context of the invention's function.[217] The "internal logic" of a means-plus-function claim is closely linked to the patented invention's function because means-plus-function claims are designed to protect the organization used to achieve the purported function.[218] As a result, a contextual equivalency analysis aids in determining whether the accused structure fits within the "internal logic" of the patented invention.

4. Does the IMS Technology Court Inappropriately Blend the Two Steps of the Equivalency Determinations for Patent Infringement?

Infringement analysis of means-plus-function claims involves two steps that the courts have traditionally kept separate.[219] As previously discussed in this Comment, the first step involves determining whether or not the accused device performs the same *function* as the claimed device.[220] The second step requires a finding that the accused device performs the function in the same *way* to achieve the same result.[221]

The *IMS Technology* decision allows function to leak into the second step.[222] Such an analysis not only gives greater weight to function,[223] but also blurs the distinction between the two steps.[224] Whether the *IMS Tech-*

---

215. *See supra* note 79 and accompanying text.

216. *See supra* notes 202-03 and accompanying text. *See also* Hofmann & Heller, *supra* note 2, at 282-83 (discussing the modern theory of claim construction that requires claims to be "interpreted as [their] unambiguous scope permits consistent with the patent's internal logic").

217. *See supra* notes 193-94 and accompanying text (discussing the contextual equivalency analysis).

218. *See supra* notes 202-03 and accompanying text (stating that the purpose of means-plus-function claims is to patent a combination or organization of components which achieve a certain function or result).

219. *See supra* Part II.A.2.c.i. (discussing the two step analysis and the authority behind it).

220. *See supra* note 74 and accompanying text.

221. *See id.*

222. *See supra* notes 193-95 and accompanying text.

223. *See supra* note 195 and accompanying text.

224. *See supra* notes 190-91 and accompanying text. By its contextual equiva-

nology court's contextual analysis is appropriate depends partly upon the reasons courts have maintained the divide between function and means, way, or structure.

With respect to means-plus-function claims, the second step seems more to be an addition to the first than a distinct element.[225] Due to the legislature's and courts' fear of over-broad patent claims, performance of the same function is insufficient alone to establish infringement.[226] In the process of interpreting § 112, ¶ 6, the courts have developed the second step to ensure that a patent will not provide protection of function alone.[227]

Because the courts seem to have developed the two-step approach to determine infringement of means-plus-function claims as a judicial method by which to enforce the scope clause of § 112, ¶ 6, the *IMS Technology* court's decision is inappropriate on grounds of blending the two steps of the traditional analysis. Quite to the contrary, the contextual equivalency test is supported by the statutory purpose of linking function to matching structures in means-plus-function claims.[228]

B. Function over Structure: Is this what the Drafters Intended? Is this Where Means-Plus-Function Patent Law Wants to Be?

1. Does the Contextual Analysis of Structural Equivalents for Means-Plus-Function Claims Disrupt the Desired Balance of Policy Consideration?

While the drafters of § 112, ¶ 6 intended to address means-plus-function claims as a distinct category of patent claims,[229] the drafters also intended for means-plus-function claims to comport with the underlying policy of all patents.[230] The drafters of § 112, ¶ 6 seem to have recognized the importance of allowing flexibility in drafting when the purported function and organization of the invention is its most novel aspect.[231] However, a patent must serve purposes other than that of convenience to the inventor.[232] A

---

lency test, the *IMS Technology* court addresses function again at step two of the analysis. Step two has traditionally involved only a structural comparison with function being addressed separately in step one.

225. *See* Hofmann & Heller, *supra* note 2, at 280.

226. *See id.* at 247, 251. Courts routinely invalidated patents that claimed only a function without any structure for impermissible over-breadth. *See id.* at 247. Section 112, ¶ 6 requires means-plus-function claims to contain a structure that corresponds to the claimed function. *See* 35 U.S.C. § 112, ¶ 6 (1994).

227. *See* Hofmann & Heller, *supra* note 2, at 279.

228. *See supra* notes 202-03 and accompanying text (discussing the legislature's purpose in drafting § 112, ¶ 6).

229. *See supra* notes 202-03.

230. *See* Hofmann & Heller, *supra* note 2, at 246-52.

231. *See id.* at 232.

232. *See id.* at 248-52 (discussing the notice-providing and progress-promoting

patent must not be so broad that it fails to provide adequate notice to other inventors and stands in the way of progress by covering infinitely many devices and equivalents.[233]

A structural equivalency analysis that focuses too much on function may undermine the importance of structure limitations[234] and lead to judicial construction of means-plus-function claims that would render these claims impermissibly broad.[235] To decide if the contextual equivalency analysis for means-plus-function developed by the *IMS Technology* court poses a threat to the balance desired by the drafters of § 112, ¶ 6, we must examine how the *IMS Technology* court defines the scope of the context in which structural equivalents are to be evaluated.[236]

2. What Level of Functional Emphasis Should We be Striving for in a Contextual Analysis of Structural Equivalents?

The proper functional emphasis is that which strives to maintain a desirable balance among policy considerations.[237] Which values are more important: the patentee's interest in protecting his intellectual property from other inventors and competitors;[238] the right of other inventors to know what they are precluded from inventing so they can avoid costly litigation and waste of their mental energy;[239] or, the goals of Congress to promote the progress of science according to its constitutional mandate to do so?[240]

---

goals of patent law).

233. *See id.*

234. *See infra* notes 254-61 and accompanying text (discussing the drafters' intent to avoid overbreadth of means-plus-function claims by limiting means claims to a corresponding structure and its structural equivalents); *see also* 35 U.S.C. § 112, ¶ 6 (1994).

235. *See* Hofmann & Heller, *supra* note 2, at 252 (discussing the early development of the legal doctrine of overbreadth and its applicability to patents that claimed "mere use, result or function"). *See, e.g.*, Holland Furniture Co. v. Perkins Glue Co., 277 U.S. 245, 257-58 (1928) (holding that a patent claiming the product to be achieved, a starch based glue, without claiming a device or mechanism, is overbroad and invalid); Gen. Elec. Co. v. Wabash Appliance Corp., 304 U.S. 364, 370 (1938) (holding a patent invalid for overbreadth even though it provided adequate notice because it claimed only the function of the invention).

236. *See infra* Part III.C.1. (discussing the scope of the context of the *IMS Technology* court's contextual equivalency analysis).

237. *See* Hofmann & Heller, *supra* note 2, at 242.

238. *See* Kundupoglu, *supra* note 17, at 41.

239. *See* Hofmann & Heller, *supra* note 2, at 241.

240. *See* Hofmann & Heller, *supra* note 2, at 248; *see also* U.S. CONST. art. I, § 8, cl. 8. Article I, section 8 clause 8 states that one of Congress' enumerated powers is "[t]o promote the Progress of Science and useful Arts, by securing for limited Times to Authors and Inventors the exclusive Right to their respective Writings and Discoveries." *Id.*

A proper contextual analysis for the determination of structural equivalents for literal infringement of means-plus-function patent claims must emphasize function without minimizing the physical structural analysis to a point at which the means-plus-function claim is a patent on function only.[241] A greater emphasis on function better serves the interests of the patentee.[242] Alternatively, a greater emphasis on structure better serves the interests of the accused inventor.[243] The next section of this Comment illustrates how these countervailing policy considerations interact and affect the determination of the proper functional emphasis in an effort to clarify the scope of the *IMS Technology* court's structural equivalency analysis.

C. The Scope of the "Context" in Which Post-IMS Technology Structural Equivalents are to be Taken for the Purposes of Equivalency Infringement Analysis

1. The *IMS Technology* Court's Definition of "Context."

The *IMS Technology* court defined the context in which structural equivalents should be evaluated for means-plus-function claim infringement when it stated that differences in the physical characteristics of structures must be "substantial in light of the role played by the 'interface means' in the claimed invention" for the physical differences to be important in the structural equivalents determination.[244] What is the "claimed invention" to which "the role played" by the interface means must be compared? The answer lies in the definition of the ambiguous "claimed invention" and the "role played."[245]

---

241. *See supra* notes 234-36 and accompanying text (discussing the propriety of emphasizing structure in means-plus-function claims); *see also supra* notes 226, 231, 258 and accompanying text (discussing the historic judicial and legislative animosity towards claims that appear to be patenting function only). Some commentators do not believe that means-plus-function claims should be limited to the recited corresponding structure and equivalents. *See* Hofmann & Heller, *supra* note 2, at 286-87 (arguing that the scope limitation is needlessly restrictive and introduces uncertainty into means-plus-function claims and should be eliminated because other general patent law rules are sufficient to ensure that the means-plus-function claim is clear, concise and adequately narrow to curb the fear of overbreadth).

242. *See supra* note 231 and accompanying text (discussing the flexibility of the means-plus-function claim).

243. *See infra* notes 255-61 and accompanying text (discussing the scope limiting purpose of the corresponding structure).

244. IMS Tech., Inc. v. Haas Automation, Inc., 206 F.3d 1422, 1437 (Fed. Cir. 2000); *see also supra* notes 190-92 and accompanying text (discussing the *IMS Technology* court's legal conclusions about evaluating structures in the context of the function).

245. *See supra* notes 22-23 (discussing how a patent is composed of individual claims and elements).

An obvious answer would be that the "claimed invention" must be the entire patented invention—the interactive computer programming system of the IMS Technology patent[246]—because the *IMS Technology* court separately addressed the "role played" by the interface means.[247] If this is the solution, the "role played" by the interface means might be the function of the *interface means claim*, or it might be the function of the *specific corresponding structure* of the interface means claim.[248] Alternatively, the "claimed invention" may be addressing the function of the *interface means claim*, not the function of the entire invention. In this case, the "role played" by the interface means must be addressed to the function of the interface means' *structure*.

2. How Does the *IMS Technology* Court Define the Scope of its Contextual Evaluation?

The *IMS Technology* court's conclusion that the physical characteristics of the conflicting devices could be found to be structural equivalents by a reasonable jury,[249] turned on the court's factual finding that the structural differences were unimportant to the function of the entire patented invention.[250] The *IMS Technology* court's equivalency analysis on the facts of the *IMS Technology* case necessarily defines the "context," which the court defined in terms of the "claimed invention," to involve a comparison between the individual disputed claim and the function of the entire invention.[251] This, however, does not completely resolve the ambiguity in the *IMS Technology* court's contextual approach to structural equivalency analysis.

The question remains unresolved: how should the "role played" by the

---

246. *See supra* note 161 (discussing the *IMS Technology* court's finding that the entire function of the IMS Technology patent and the Haas Automation invention is to create an interactive machine tool programming system).

247. *See infra* Part II.C.2. The *IMS Technology* court ultimately concluded that the "context" is in reference to the function of the entire invention. *See infra* text accompanying note 250.

248. *See supra* notes 244-45 and accompanying text. Because means-plus-function claims require a corresponding structure or means for the recited function, the "role played" must be the function of the interface "means." *See supra* Parts II.A.1 & II.A.2.b (discussing the drafting and construction of means-plus-function claims).

249. *See supra* Part II.B.3 and accompanying text (discussing the *IMS Technology* court's conclusions).

250. *See supra* notes 191-92 and accompanying text. The *IMS Technology* court identified the function of the IMS Technology patent to be that of "provid[ing] a way of storing programs created using the inventive programming apparatus and process." IMS Tech., Inc. v. Haas Automation, Inc., 206 F.3d 1422, 1437 (Fed. Cir. 2000).

251. *See supra* notes 191-92, 250 and accompanying text.

means be defined?[252] Unlike the function of the "claimed invention," the *IMS Technology* court does not specifically characterize the "role played" by the means in its opinion.[253] Is the court concerned with the function of the interface means claim with respect to the overall invention or is it concerned with the role that the particular corresponding structure is able to serve with respect to the overall invention?[254]

3. How Should the Scope of the Context be Defined?

The appropriate scope of the contextual equivalency analysis for literal infringement of means-plus-function claims must accommodate the policy and purposes behind § 112, ¶ 6 and patent law generally.[255] The two main purposes of patents are: to "promote progress of science and the useful arts"[256] and to give notice such that the "public may know what they are prohibited from doing during the existence of the monopoly[.]"[257] Overbreadth has been traditionally seen as an evil with the power to undermine both the progress-promoting and the notice-providing goals of patents.[258] Patents that claim only function without reciting structure have historically been held invalid as overbroad.[259]

In balance with the foregoing policy considerations, means-plus-

---

252. *See supra* note 244 and accompanying text. The means claim at issue in the *IMS Technology* case was the interface means. *See also supra* Part II.B.2.c.ii. and accompanying text.

253. *See supra* notes 244-47 and accompanying text.

254. *See supra* note 247 and accompanying text (discussing the possible meaning of the "role played" by the means claim). While it may not be obvious, these two definitions are not one in the same. The function of the means claim is a much broader definition that would cover any structure that can perform the same function. The role that the particular structure is able to serve is narrower and focuses more on the physical limitations of the corresponding structure to which the means claim of the patent is limited. The distinction may not be outcome determinative for a given fact pattern, but the distinction is legally significant when viewed in light of the policy behind § 112, ¶ 6. *See also supra* notes 234-35 and accompanying text.

255. *See supra* notes 229-33 and accompanying text; *see also* Hofmann & Heller, *supra* note 2, at 239-40 (discussing the historic purposes behind the enactment of § 112, ¶ 6).

256. Blanchard v. Sprague, 3 F. Cas. 648, 650 (C.C.D. Mass. 1839); *see also* Hofmann & Heller, *supra* note 2, at 248.

257. Brook v. Fiske, 56 U.S. 212, 214-15 (1853); *see also* Hofmann & Heller, *supra* note 2, at 252.

258. *See* Hofmann & Heller, *supra* note 2, at 255-56; *see also* cases cited *supra* note 235.

259. *See* Hofmann & Heller, *supra* note 2, at 255-56; *see also* cases cited *supra* note 235.

function claims are supposed to provide flexibility to the inventor.[260] Since the enactment of § 112, ¶ 6, inventors can claim and protect their inventive functions subject to scope limitation.[261] The scope limitation inherent in the corresponding structure requirements of § 112, ¶ 6 serves to protect the progress-promoting and notice-providing goals of patent law.[262]

To best preserve a balance among the countervailing policy considerations underlying patent law, the proper scope of the contextual equivalency analysis used by the *IMS Technology* court must not be overbroad.[263] How the "role played" by the means is defined is central to the issue of whether or not the *IMS Technology* court's structural equivalency analysis leads to overbreadth of means-plus-function patent claims.[264] Of the two definitions proposed in this Comment,[265] which seems to be a better safeguard against overbreadth[266] and patenting of function only?[267] The *IMS Technology* court pushed function to the forefront when it defined the contextual equivalency analysis for structural equivalents with respect to the function of the overall invention.[268] A means-plus-function claim must be grounded partially in structure or it would be a patent on function only.[269]

---

260. *See* Hofmann & Heller, *supra* note 2, at 259.

261. *See supra* notes 226-28 and accompanying text.

262. *See* Dabney, *supra* note 53, at 293-94; *see also supra* notes 201-07 and accompanying text.

263. *See supra* note 258 and accompanying text (discussing how a patent must not be overbroad). The *IMS Technology* court's contextual equivalency analysis contradicts policy if it is not interpreted to avoid overbreadth of means-plus-function claims. However, even if the contextual analysis is interpreted so as to avoid overbreadth to means-plus-function claims, it still may not be appropriate. The *IMS Technology* analysis must also be evaluated with respect to other considerations such as "certainty." *See infra* Part III.E.

264. *See supra* notes 245, 252 and accompanying text. The *IMS Technology* court's decision must comport with the policy behind patent law generally and means-plus-function claims in particular to be a model infringement analysis strategy worthy of following. Because the *IMS Technology* court did not provide a concrete definition for the "role played" by the means claim, which is half of the reference for the court's structural equivalency analysis, the "role played" should be interpreted so that the *IMS Technology* opinion is consistent with the underlying policy considerations.

265. *See supra* note 248 and accompanying text (discussing two possible definitions).

266. *See supra* notes 248-49 and accompanying text (discussing how courts and scholars have regarded overbreadth as a force capable of deteriorating the notice-providing and progress-promoting purposes of patent law).

267. *See supra* note 259 and accompanying text (discussing the historic judicial intolerance for claims that attempt to patent function only).

268. *See supra* notes 195, 237-41 and accompanying text (discussing the effect of *IMS Technology* court's scope of its structural equivalency analysis).

269. Section 112, ¶ 6 allows means claims to describe the function to be patented without reciting structure. *See* 35 U.S.C. §112, ¶ 6 (1998); *see also supra* Part II.A.1.

Limiting the context to the role that the particular structure is able to serve with respect to the overall invention forces structure into the equation. Alternatively, defining the context in terms of the function of the means claim with respect to the overall invention obscures the structural requirements of § 112, ¶ 6.[270]

Accordingly, the "role played" by the means claim in a means-plus-function structural equivalency determination must be the role of the particular coordinating structure recited in the patent.[271] So defined, the contextual equivalency analysis properly limits means-plus-function claims to structure without undermining the spirit of the means-plus-function claim.[272] To conclude otherwise would be to run the risk of interpreting the *IMS Technology* court's contextual equivalency analysis to allow for means-plus-function claims to be patenting function only.

In application, this Comment's interpretation of the *IMS Technology* court's contextual equivalency analysis limits structural equivalents to those that can perform the function of the recited corresponding structure in substantially the same way as the corresponding structure. The alternative definition would allow any device capable of performing the same function as the recited function of the means-plus-function claim to be considered a structural equivalent.[273] Therefore, after a court has determined that the same function is performed by the claimed device and the accused device,[274] a court following the *IMS Technology* court's opinion

---

If the importance of the corresponding structure is too greatly minimized by the contextual equivalency analysis' functional emphasis, the means claim will essentially cover any structure that can perform the recited function. With the loss of their structural restrictions, means-plus-function claims would be construed to patent function only.

270. *See supra* notes 255-62 and accompanying text (discussing the scope limiting structural requirement of §112, ¶ 6).

271. *See supra* notes 56-67 and accompanying text (discussing the construction of a means-plus-function patent claim).

272. *See supra* notes 202, 260 and accompanying text (discussing how § 112, ¶ 6 was enacted to allow patentees flexibility in drafting patents for inventions that are inventive organizations, arrangements and interaction among structural components).

273. *See supra* note 269.

274. *See supra* note 74 and accompanying text (discussing the first step of means-plus-function literal infringement litigation). It can be assumed that the function of the particular corresponding structure serves the function of the means-plus-function claim. *See supra* note 56-67 and accompanying text (discussing the construction of means-plus-function claims). As a result, the differences between the two proposed definitions of the "role played" by the means claim do not affect the determination of whether or not an accused device performs the same *function as the means-plus-function claim.* The definition of "role played" by the means claim pertains only to the determination of structural equivalents and not to the first step of determining whether the same function is performed. If this Comment's interpretation of the *IMS Technology* court's contextual infringement analysis were misapplied (i.e., mistakenly

would then commence the contextual equivalency analysis to determine if the accused device is a structural equivalent of the patented device.[275]

The contextual equivalency analysis begins with the relative importance evaluation portion of the *IMS Technology* court's contextual equivalency analysis.[276] The relative importance evaluation involves inquiry into how important the physical characteristics of the corresponding structure of the disputed means claim are with respect to the overall function of the patent.[277] If the chosen corresponding structure is of small importance, relative to the overall function of the invention, a wider range of physical structures will be able to fulfill the role of the corresponding structure of the means-plus-function claim.[278] Of course, the reverse is also true. That is, if the particular corresponding structure is of great importance to the overall function of the invention, perhaps even affecting the overall function, fewer devices will have physical characteristics similar enough to the patented device to fulfill its role in the overall invention.[279]

For the purposes of litigation, the contextual equivalency analysis can be performed to determine infringement in a particular case by evaluating how well the physical characteristics of the accused device would serve the particular role of the corresponding structure of the disputed means claim.[280] Would the overall function of the patented invention be disrupted

---

applying the definition with its explicit structural limitations of "role played") to *both steps* of the infringement analysis, no harm should result.

275. *See supra* note 74 and accompanying text (discussing the two steps involved in patent infringement litigation of means-plus-function claims). According to the *IMS Technology* court, contextual analysis is the proper way in which to determine the second step of the infringement analysis: deciding whether the function is performed in substantially the same way. *See supra* notes 192, 244 and accompanying text.

276. *See supra* note 248 and accompanying text (discussing the two proposed definitions); *see also supra* notes 249-51 and accompanying text (discussing the context in which the structural equivalency must be evaluated according to the *IMS Technology* court's contextual equivalency analysis).

277. *See supra* notes 265-72 and accompanying text (discussing this Comment's interpretation of the context in which the *IMS Technology* court's structural equivalency analysis must be made); *see also supra* notes 191-92, 244 and accompanying text (discussing the *IMS Technology* court's conclusions).

278. *See supra* notes 204, 260 and accompanying text (discussing the desired flexibility of means-plus-function claims when the invention is unique for its organization or particular combination more than for its individual structural components). This is also the *IMS Technology* court's ultimate conclusion when it decided that the Haas Automation device, a floppy disk drive, could be considered a structural equivalent of the IMS Technology patent's tape cassette transport. *See supra* notes 191-92, 244 and accompanying text (discussing the *IMS Technology* court's conclusions).

279. *See supra* note 254 and accompanying text. This is the conclusion of the *Sunrise Medical* court. *See infra* notes 314-18 and accompanying text (discussing the *Sunrise Medical* court's conclusions).

280. *See supra* notes 244-48 and accompanying text (discussing the contextual

if the corresponding structure of the means claim was replaced by the accused device? Short of disruption, how well can the accused device serve the overall function of the invention?

For the purposes of drafting a means-plus-function claim, the contextual equivalency analysis employed by the *IMS Technology* court is more problematic.[281] Unlike infringement litigation, during which the structure of the accused device is known, patent drafting requires a futuristic determination of what structures the courts are likely to find equivalent.[282] By patenting his invention, a patentee seeks to protect his invention by disabling a competitor from "mak[ing], us[ing], sell[ing], offer[ing] to sell, or import[ing] a product containing all the limitations recited in [the patentee's] claim."[283] A valid patent must also provide adequate notice to all others as to what is being claimed as protected from duplication.[284] How does the contextual equivalency analysis employed by the *IMS Technology* court affect the patentee's ability to protect his invention and provide adequate notice by use of means-plus-function claims?

D. Does *IMS Technology* Import Additional Uncertainty Into Structural Equivalency Analysis?

1. The Problem with Uncertainty in Patent Claims

One problem with the means-plus-function claim is that it is, by statutory definition, somewhat indefinite.[285] The indefiniteness of a means-plus-function claim is a product of its permissible drafting which allows means claims to stand independent from structure as long as the means claims correspond to some structure recited within the patent.[286] During

---

analysis). Without explicitly stating that it is doing so, *IMS Technology* court performs the contextual equivalency analysis in this way when it bases its legal conclusion on its factual finding that the differences in physical characteristics between the Haas Automation device and the IMS Technology patent's device are insignificant. *See* IMS Tech., Inc. v. Haas Automation, Inc., 206 F.3d 1422, 1437 (Fed. Cir. 2000). The *Sunrise Medical* court performs the analysis in the same way. *See* Sunrise Medical HHG, Inc. v. Airsep Corp., 95 F. Supp.2d 348, 445 (W.D. Pa. 2000).

281. The scope of the patent's protection and the outcome of patent infringement litigation depend heavily on the patent document. *See supra* notes 48-55 (discussing judicial construction of patent claims during infringement litigation).

282. Means-plus-function claims must be drafted with an eye towards how the courts will employ the contextual equivalency analysis to provide the patentee adequate protection in the event of infringement and litigation. *See infra* notes 289-95.

283. Kundupoglu, *supra* note 17, at 41.

284. *See* Hofmann & Heller, *supra* note 2, at 252.

285. *See id.* at 236.

286. *See supra* Part II.A.1 (discussing the drafting of means-plus-function claims); *see also supra* Part II.A.2.b (discussing the construction of means-plus-function claims during infringement litigation).

patent infringement litigation of a means-plus-function claim, the parties can argue over what structure is the corresponding structure.[287]

Uncertainty in means-plus-function claims implements policy concerns about the notice-providing goal of patents as well as a patentee's right to protect his invention.[288] How does a means-plus-function claim provide notice to other inventors of what is protected if the corresponding structure often must be judicially determined as the result of the patentee instigating infringement litigation?[289] Furthermore, how can a patentee be guaranteed that his invention is protected by his patent's means-plus-function claims? Such questions illuminate the level of uncertainty inherent in § 112, ¶ 6 claims.[290]

Uncertainty is problematic and undesirable because it "makes difficult the assessment of likely outcomes of patent infringement litigation."[291] Naturally, inventors would like to avoid having to be dragged into court to defend themselves against patentees' claims of infringement. Inventors, competitors and importers[292] need "to make crucial decisions with relative predictability and comfort as to what infringes a patent without going through a lawsuit."[293] Inventors, competitors and importers must be able to ascertain the scope of a patent to safeguard themselves against infringement liability.[294]

One of Congress' purposes in enacting patent legislation is to foster the development of technology.[295] An important aspect of patent law that advances the goals of development is the "right of the public to design around a patented invention."[296] Uncertainty about the scope of means-plus-function claims, limited by the corresponding structure, makes in-

---

287. *See* Hofmann & Heller, *supra* note 2, at 236.

288. *See supra* notes 257, 260 and accompanying text.

289. *See* Hofmann & Heller, *supra* note 2, at 240. Hofmann and Heller argue that § 112, ¶ 6 "provides a statutory exception to the general requirement that the claim provide notice to all of the proper scope of the patent protection." *Id.* at 235. *See also supra* note 257 and accompanying text (discussing the notice-providing role of patents).

290. *See* Hofmann & Heller, *supra* note 2, at 240. "[Paragraph 6] provides uncertainty as to the construction of a means-plus-function limitation." *Id.*

291. *Id. See* Cybor Corp. v. FAS Tech., Inc., 138 F.3d 1448, 1455 (Fed. Cir. 1998). The *Cybor Corp.* court treated "intrajudicial certainty" among factors important in means-plus-function infringement determinations. *See id.* "Indeed...the Supreme Court endorsed this court's role in providing national uniformity to the construction of a patent claim . . . ." *Id.*

292. *See* 35 U.S.C. § 112, ¶ 6 (1998).

293. Hofmann & Heller, *supra* note 2, at 241.

294. *See id.*

295. *See supra* note 256 and accompanying text.

296. Hofmann & Heller, *supra* note 2, at 241 (stating that the "right of the public to design around a patented invention is one of the most important benefits that justify awarding the patentee exclusive rights to the invention").

venting "legally risky" which can ultimately "impede[] the ability and willingness of potential inventors and competitors to develop in the area of the patent."[297] According to some, "arbitrary and inconsistent application of [equivalency doctrines] contribute[] to [the] uncertainty [inherent within means-plus-function claims]."[298]

2. How Does the *IMS Technology* Court's Contextual Analysis for Structural Equivalents of Means-Plus-Function Claims — A Judicial Application of an Equivalency Doctrine[299]— Affect the Level of Uncertainty Inherent in § 112,¶ 6 Claims?

The *IMS Technology* court's emphasis on function, generally, and this Comment's interpretation of the *IMS Technology* court's contextual equivalency analysis, in particular, tend to provide greater certainty to means-plus-function claims.[300] First, consistent with the purposes of the drafters of § 112, ¶ 6, a functional emphasis allows a patentee to patent with greater ease what § 112, ¶ 6 allows him to patent: the organization and interaction of individual components.[301] As a result, the patentee knows that his innovative organization is judicially viewed as the essence of his means-plus-function claims and will serve as the focus of a court's infringement analysis. Secondly, this Comment's interpretation requires that the contextual equivalency analysis' relative importance comparison be made with respect to the functions that the particular corresponding structure recited can perform.[302] This interpretation provides the patentee with a surer idea of the protective scope of his means-plus-function claims.

---

297. *Id.* at 241-42.

298. *Id.* at 242.

299. *See supra* text accompanying notes 184-85.

300. *See supra* Part II.A.1 (discussing the IMS Technology court's structural equivalency analysis and its effect of emphasizing function); *see also supra* Part II.B.2 (discussing this Comment's interpretation of ambiguities within the *IMS Technology* court's holding and how this interpretation best balances the countervailing policy concerns). According to Hofmann and Heller, the scope limiting corresponding structure requirement of § 112, ¶ 6 should be eliminated all together. *See* Hofmann & Heller, *supra* note 2, at 234-35 (arguing that "constricted definition[s] of scope under the statute cast[] considerable doubt as to the utility of means-plus-function patent claims for protecting an invention"). While this Comment does not go so far as to recommend legislative elimination of the scope clause of § 112, ¶ 6 or judicial abandonment of a means claim's link to its corresponding structure, it finds that the *IMS Technology* court's emphasis on function is supported by the purposes and policy behind § 112, ¶ 6 and endorses the *IMS Technology* court's less "constricted definition of scope." *See supra* Part III.A.2.

301. *See supra* note 260 and accompanying text.

302. *See supra* Part III.C.3 (discussing this Comment's recommendation for the proper scope of the contextual analysis for structural equivalents of means-plus-function claims).

The patentee can evaluate the extent of his patent's protection by evaluating the corresponding structures he has chosen to recite and the relative ease with which these structures can be interchanged with others and still adequately perform the overall function of the invention.[303]

The inventor, competitor or importer is provided more competent notice by a means-plus-function claim due to the greater certainty provided by the *IMS Technology* court's contextual analysis and its recommended scope. A court will not reach the contextual equivalency analysis for structural equivalents unless it first finds that the accused invention performs the same function as the patented invention.[304] Therefore, an inventor will not infringe a means-plus-function claim if his invention does not perform the same function as that recited in the means-plus-function claim. Furthermore, an inventor can tell with certainty whether his invention performs the same function as that recited in the disputed means-plus-function claim.[305]

To avoid infringing a means-plus-function claim, the *IMS Technology* court's contextual analysis first requires an inventor to understand the overall function of the invention.[306] The inventor must also understand the function of the particular means-plus-function claim at issue.[307] One would assume that an inventor that is designing a device to perform the same function as the one described in the means-plus-function claim must be able to appreciate the overall function and the function of the individual means-plus-function claims at issue.

This Comment's interpretation of the proper context in which structural equivalents are to be evaluated may be a more challenging analysis for an inventor to undertake to keep from designing an offending device.[308] However, the recommended context, while potentially more difficult to utilize than other possible interpretations of the *IMS Technology* court's

---

303. *See id.*

304. *See supra* notes 72-74 and accompanying text (discussing the two steps of patent infringement determinations, the first of which is to determine if a common function is performed).

305. *See supra* notes 288-90 and accompanying text (discussing how the uncertainty inherent within means-plus-function claims is attributable to the room for argument over what is the proper *corresponding structure* for the disputed means-plus-function claim, but not disputes about function).

306. *See supra* notes 249-51 and accompanying text (discussing the *IMS Technology* court's analysis).

307. *See id.*

308. *See supra* Part III.C.3 (discussing this Comment's recommendation). The recommended analysis is arguably more challenging because it requires the inventor to appreciate the particular functional limitations of the corresponding structure of the means claim, as opposed to just the function of the means claim. *See id.* However, it is not unreasonable to expect "one having ordinary skill in the technology of the invention" to appreciate the particular functional limitations of the means claim's corresponding structure. *See supra* note 21.

contextual analysis,[309] is not only the most appropriate in terms of policy[310] but also adds greater certainty to the scope of a means-plus-function claim's scope of protection.

To avoid infringing a means-plus-function claim, an inventor must identify the inventions' function and determine which recited structure corresponds to the means claim at issue irrespective of the *IMS Technology* court's holding.[311] The corresponding structure determination is difficult because it is subject to varying views and arguments.[312] Even when the inventor's determination of the corresponding structure is consistent with the patentee's understanding of the corresponding structure, the inventor must also determine what structures run the risk of infringing the patent as structural equivalents.[313]

Under the proposed contextual analysis, an inventor will have to determine the relative importance of the function of the corresponding structure compared to the overall function of the invention.[314] This should be easily within the realm of the inventor's abilities because he can assess the relative importance of the structure of his own device with respect to the common function of the overall inventions.[315] If the corresponding structure recited in the patent can be interchanged for his own device without affecting the overall function of his invention, his device will likely be found to be a structural equivalent under the contextual analysis.[316]

Obviously, structures that are easily interchangeable will broaden the protection of the means-plus-function claim.[317] However, while breadth of protection is desirable from the patentee's viewpoint, it is in the patentee's interest, as an inventor himself, to select and recite structures that most adequately serve the function of the means-plus-function claim and the overall function of the invention.[318] In doing so, the patentee is voluntarily

---

309. *See supra* Part III.C.3 (discussing this Comment's recommendation).

310. *See id.*

311. *See supra* Part II.A.2.b (discussing the general construction of means-plus-function claims); *see also* 35 U.S.C. § 112, ¶ 6 (1998).

312. *See supra* notes 288-90 and accompanying text (discussing this uncertainty within the corresponding structure requirement and its susceptibility to dispute).

313. *See supra* Part III.C.3 (discussing this Comment's recommendation).

314. *See id.*

315. This analysis is the reverse of the comparison the courts and adverse parties must undergo during infringement litigation to determine and argue whether or not the accused device is the equivalent of the patented device. *See supra* notes 276-81 and accompanying text.

316. *See supra* notes 265-72 and accompanying text (discussing the recommended contextual analysis for structural equivalents).

317. *See supra* Part III.C.3 (discussing this Comment's recommendation); *see also* notes 255-72 and accompanying text (discussing the effect of equivalents on the breadth of a patent's scope).

318. This statement assumes, commonsensically, that a patentee's most important goal is the design of his invention, not just to design for the sake of blocking out de-

limiting himself to the recommended structural equivalency analysis.

While the identification of the corresponding structure is still a large area of uncertainty, the contextual analysis for determining structural equivalents recommended by this Comment adds certainty to what will be deemed a structural equivalent. Under this Comment's recommended interpretation of the *IMS Technology* court's contextual analysis, structural equivalents of means-plus-function claims must closely perform within the particular functional limitations of the corresponding structure when those functional limitations affect the overall function of the invention.[319] The function of the corresponding structure is limited by the physical characteristics of the device. As a result of its link to the physical characteristics of the particular corresponding structure recited, the recommended analysis gives the patentee less room for argument in litigation, and the inventor better overall notice as to what are functional equivalents of the corresponding structure.

E. The *Sunrise Medical*[320] Court's Application of Post-*IMS Technology* Means-Plus-Function Theory: Analysis of an Associated Case

1. The findings, the analysis, and the conclusions of the *Sunrise Medical Court*

In addition to addressing accusations of infringement of several patented claims, the *Sunrise Medical* court analyzed literal infringement of the "timing means" claim under § 112, ¶ 6.[321] The "timing means" claim describes a means of predetermining certain operating parameters of a respirator[322] that are used to control when and for how long oxygen is delivered to the patient.[323] The patented device uses a fluidic circuit to perform the timing means' function while the accused device uses an electronic circuit.[324] Both of these circuits transmit signals in response to timing parameters that help control the timing of oxygen delivery to the patient.[325]

---

velopment of vast areas of technology.

319. *See supra* Part III.C.3 (discussing this Comment's recommendation).

320. Sunrise Medical HHG, Inc. v. AirSep Corp., 95 F. Supp. 2d 348 (W.D. Pa. 2000).

321. *See id.* at 445.

322. *See id.*

323. *See id.* at 379.

324. *See id.* at 445. The fluidic circuit uses a variable orifice resistance as compared to the electronic resistance of the accused device. *See id.* The fluidic circuit also uses an elastomeric balloon as compared to the electronic capacitance counterpart found in the accused device. *See id.*

325. *See id.* at 376 (discussing that the order in which the circuits operate is the same because the accused device, like the patented device, uses the timing means each time an inspiration (breath in) is "sensed" to control the delivery of oxygen to the

The *Sunrise Medical* court noted that the fluidic and electric circuits are not "structurally similar,"[326] in other words, the respective devices are not equivalent structures, so the next step in determining infringement under § 112, ¶ 6 is to evaluate whether the two devices are structural equivalents.[327]

The *Sunrise Medical* court found that the two types of circuits, while noting that they could be equivalents in some contexts, could not be deemed structural equivalents for these particular inventions.[328] The *Sunrise Medical* court identified the primary objective of the overall invention as that of "provid[ing] a respirator apparatus and method of operating same wherein respirating gas is supplied to a patient substantially at the beginning of an inspiration and for a time period thereafter *which is a fraction of the duration of inspiration.*"[329] The *Sunrise Medical* court then determined that the *timing means* has great importance to the overall function of the invention and made note of the structural differences between the two devices.[330] In a significant footnote, the *Sunrise Medical* court clarifies its conclusion.[331] Claiming to distinguish the *Sunrise Medical* case from the *IMS Technology* case, the court found that the physical difference between the structures of the two timing devices "affect the operation of the claimed method itself[]"and therefore the accused device could not be deemed to be a structural equivalent of the patented device.[332]

2. How the *Sunrise Medical* Case Inadvertently Applied the *IMS Technology* Court's Contextual Analysis

Though purporting to distinguish *IMS Technology* on the facts, the *Sunrise Medical* court actually applied a contextual analysis for structural equivalents of means-plus-function claims.[333] The *Sunrise Medical* court first speaks generally of the importance of the timing means with respect to the overall invention.[334] In footnote 743, the court clarifies that it is

---

patient).

326. Sunrise Medical HHG, Inc. v. AirSep Corp., 95 F. Supp. 2d 348, 445 (W.D. Pa. 2000).

327. *See supra* Part II.A.2 (discussing patent infringement litigation of means-plus-function claims).

328. *See Sunrise Medical*, 95 F. Supp. 2d at 445 (quoting IMS Tech. Inc. v. Haas Automation, Inc., 206 F.3d 1422, 1435 (Fed. Cir. 2000) for its general proposition "two structures that are equivalent in one environment may not be equivalent in another").

329. *Id.* (emphasis added).

330. *See id.*

331. *See id.* at 445 n.743.

332. *Id.*

333. *See supra* notes 328-30 (discussing the *Sunrise Medical* court's analysis of structural equivalency); *see also supra* notes 249-51 (discussing the *IMS Technology* court's contextual analysis).

334. *See supra* note 330 and accompanying text.

interested in the effect on the overall function of the particular *structures* that have been chosen in each device to carry out the function of the timing means.[335] By doing so, the *Sunrise Medical* court applied a contextual analysis analogous to this Comment's recommended interpretation of the *IMS Technology* court's structural equivalency analysis.[336]

3. The *Sunrise Medical* Case and this Comment's Interpretation of the *IMS Technology* Court's Contextual Analysis of Structural Equivalents for Means-Plus-Function Claims

The *Sunrise Medical* case illustrates how this Comment's interpretation—which is careful not to eliminate physical characteristics from a contextual analysis for evaluating structural equivalents—can have a practical, not merely a theoretical, effect on means-plus-function claims.[337] A contextual analysis that requires structural equivalency to be evaluated with respect not only to the overall function of the invention but also to the particular functional limitations of the chosen corresponding structure can in fact be outcome determinative.[338] Had the *Sunrise Medical* court disregarded the functional limitations of the two devices and instead focused only on the function of the timing means, the fluidic circuit and the electronic circuit, both of which generally perform the timing function, could be deemed to be structural equivalents.[339] As can be seen from this example, a contextual analysis that disregards the functional limitations of corresponding structures would give means-plus-function claims a much greater scope and would curtail the public's ability to design around the patent.[340]

The *Sunrise Medical* court's analysis also illustrates the need for guidance in structural equivalency determinations for means-plus-function claims. By distinguishing *IMS Technology* on the facts while simultaneously applying a similar contextual analysis, the *Sunrise Medical* court

---

335. *See supra* notes 331-32 and accompanying text.

336. *See supra* notes 268-70 and accompanying text (discussing this Comment's recommended interpretation of the *IMS Technology* court's contextual analysis). The *Sunrise Medical* court did not spell out its definition of contextual analysis (quite to the contrary, it disclaimed *IMS Technology*) but clearly finds the structure of the devices to be important. In this way, the *Sunrise Medical* court's analysis supports this Comment's recommended approach.

337. *See supra* notes 255-63 and accompanying text (discussing the importance of structure for achieving a desirable balance among policy considerations irrespective of whether in application the link to structure would have an effect on the outcome of any one particular infringement case).

338. *See supra* notes 329-33 and accompanying text.

339. *See supra* notes 321-27 and accompanying text (discussing the two devices of the *Sunrise Medical* litigation).

340. *See supra* notes 323-28 and accompanying text (discussing the policy considerations of patent law).

exposes its confusion.[341] Courts, inventors, importers, retailers, patentees, and the public need a clear standard by which each can determine, avoid and protect from infringement of means-plus-function claims. This Comment's recommended interpretation of the *IMS Technology* court's contextual analysis should be applied in all structural equivalency determinations of means-plus-function claims.

CONCLUSION

The *IMS Technology* court made progress towards a better understanding of the statutory requirements for means-plus-function claims. Although the courts have once again borrowed from the Doctrine of Equivalents, a distinct doctrine with different policy considerations, evaluating structural equivalents of means-plus-function claims to determine literal infringement by use of a contextual analysis serves the underlying purpose of § 112, ¶ 6.[342] While the *IMS Technology* court's contextual analysis leans in the right direction, its analysis contains ambiguities.[343] This Comment recommends an interpretation of the *IMS Technology* court's contextual analysis that allows flexibility in patenting when the invention is in the organization of and interaction between structural components[344] while at the same time, provides notice to the public and promotes the advancement of technology.[345]

The recommended contextual analysis, which is the *IMS Technology* court's analysis as interpreted by this Comment, involves an evaluation of the functional limitations of the particular corresponding structure recited in the patent and the effect those limitations have on the overall function of the invention.[346] If the importance of the structure of the means claim with respect to the function of the overall invention is small, the functional limitations of the particular corresponding structure will have little or no effect on the overall function. As a result, a greater range of structural variation will be tolerable, or in other words, a greater range of structures will be considered as structural equivalents.[347]

As illustrated by the *Sunrise Medical* case, a contextual analysis can have a determinative effect and it is therefore important that this Com-

---

341. *See supra* notes 255-62 and accompanying text (discussing the *Sunrise Medical* court's analysis).

342. *See supra* notes 329-33 and accompanying text (discussing the purpose behind the enactment of § 112, ¶ 6).

343. *See supra* notes 201-03 and accompanying text (discussing the ambiguities of the *IMS Technology* court's contextual analysis).

344. *See supra* notes 244-54 and accompanying text (discussing the purpose behind the enactment of § 112, ¶ 6).

345. *See supra* notes 256-62 and accompanying text (discussing the policy considerations of patent law).

346. *See supra* Part III.C.3 and accompanying text (discussing this Comment's recommendation).

347. *See supra* notes 276-80 and accompanying text.

ment's recommended contextual analysis be used in order to maintain a desirable balance among policy considerations.[348] The *Sunrise Medical* court's analysis also indicates a need for a clear and enumerated test by which courts, patentees and others can determine whether certain devices will be deemed structural equivalents later down the road in infringement litigation.[349]

The contextual analysis recommended in this Comment maintains a proper balance of policy considerations—protecting the patentee, providing notice, and promoting advancements in technology, and offers guidance in the area of means-plus-function claims necessary to increase certainty in means-plus-function claim drafting and litigation.

Shenanne Tucker

348. *See supra* notes 338-41 and accompanying text.

349. *See supra* note 345 and accompanying text; *see also supra* note 1 (discussing the appropriateness of describing the *IMS Technology* court's contextual analysis as a mutation partially because of the shadow of skepticism it casts). This author is of the opinion that skepticism is not synonymous with close-mindedness but instead is a state of mind that spawns thorough analysis, which in turn, spawns understanding. Understanding is a noble goal and a powerful thing to gain.

SYMPOSIUM

Translating Science into Law: Lessons from Doctors, Judges, and Lawyers about the Use of Medical Evidence in the Courtroom

INTRODUCTION

The relationship between law and medicine is serendipitous at best, despite the fact that it is difficult to imagine a criminal, environmental, or personal injury case that does not include empirical evidence. Although science increasingly permeates our courts, an overwhelming majority of the legal community has little knowledge of even basic scientific language or methodology. The courtroom can be a difficult venue in which to gain a true understanding of how scientists develop theories, collect data, draw conclusions or reject alternative explanations. As a result, judges and jurors involved in science-based litigation often find themselves struggling to evaluate the validity of scientific data and its relevance to the case at hand. Equally troubling is the fact the legal process, with its strict procedural constrainsts can impede the effective communication of scientific concepts. When efforts to communicate fail, legal decision-makers are left with Hobson's choice; they must either reject the science that they do not fully understand or accept it on faith.

This Symposium issue offers thoughts from experts in the medical, judicial, and law enforcement fields addressing the troublesome relationship between law and science. Most importantly, these experts offer pragmatic solutions designed to improve our ability to understand and use medical science in the courtroom. These suggestions provide guidance for judges, practicing attorneys, medical experts, and members of the legal and scientific and academic communities.

Some of the issues discussed in the following pages include the following: How can non-scientists overcome their fear of science? Can we begin by correcting common misunderstanding about the scientific process? What do judges need to understand about medical expert testimony when making admissibility decisions? Can scientists offer specific guid-

ance to the bench in the form of concrete suggestions or be serving in the role of the neutral expert? Scientists and legal practitioners differ in their use of language, in the methodologies they employ, and in their goals. Can the recognition of these distinctions aid us in improving the comprehension and use of scientific evidence in our courts? Whom can we enlist from medical and legal communities to work together to craft creative, practical solutions aimed at both preventing the misuse of science and achieving superior integration of science into the legal process? The questions raised in this issue of the *New England Law Review* highlight the critical relationship between law and science. In all cases involving empirical evidence, good law depends on good science. Recently it has become popular sport for scientists and lawyers to blame each other for legal decisions that misapply basic principles of science. However, the recent proliferation of academic literature and discourse expounding on the interdisciplinary chasm has done little to educate judges and lawyers who must grapple with real scientific evidence in real cases. If we know that judges and juries cannot make accurate legal decisions if invalid scientific evidence distorts their understanding of the facts, then we should recognize that this practical problem demands a practical solution.

Science in the Courts

Robert L. Park, Ph.D.*

It is a special pleasure to be asked to speak at the New England School of Law. About two Sundays after the September 11 terrorist attack, there were full-page ads in *The New York Times* and *The Washington Post* from organizations expressing their grief, or patriotism, or asking for donations to help the victims and their families. There was even a full-page ad by the SPCA to raise funds to find suitable homes for the pets owned by deceased victims of the attack. They were simply doing what they do best.

Of course, there was a certain amount of commercialization in some of these ads. General Motors had an ad that stated "Let's Keep America Rolling." As though it's your civic responsibility in this time of crisis to go out and buy a gas guzzling SUV. That was alright. However, there was one full-page ad that really struck me. It was an ad for "The Endowment Funds for World Peace." The "Endowment Funds for World Peace" was seeking contributions of one billion dollars. "Surely, there are a lot of Americans," the ad said, "who must have a billion dollars. But, if you cannot donate that much, perhaps you could simply advance it to us, and we will pay you back." However, before you give your billion dollars, you might want to know to whom you are giving it, and what they are going to use the money for. You would actually be giving it to the Maharishi Mahesh Yogi, spiritual advisor to the Beatles and the founder of Transcendental Meditation.

What he wanted to do, he said, was pure science. He wanted to create a consciousness field for the whole world: a consciousness field that would induce peace among all people. To do this, he needed a trained core of

---

* Dr. Park is a Professor of Physics at the University of Maryland and Director of the Washington Office of the American Physical Society. He was the Edgar Lewis Marston Fellow at Brown University in 1960. In 1974, he was appointed Professor of Physics and Director of the Center of Materials Research at the University of Maryland in College Park. In 1978, he became the Chair of the Department of Physics in Astronomy. Professor Park is the author of *What's New*, a controversial weekly electronic commentary on science policy issues. He is also a regular contributor of opinion articles in major newspapers, including the *New York Times* and the *Washington Post*, and a frequent guest on radio and television news programs. He is also the author of the book *Voodoo Science: The Road from Foolishness to Fraud.*

40,000 transcendental meditators. And, when he said trained, he meant the best. And, by the best, he meant yogic fliers. You might be wondering what a Yogic flyer is. Actually, this organization had been to Washington two years earlier during the trouble in the Balkans. They had gone to Washington to raise money to create peace in Kosovo. And the way they were going to do it, was again with a trained core of transcendental meditators. They were going to meditate in unison. In this case, they thought about 7,000 would suffice.

In the new call, they needed 40,000, because they were going to spread peace all over the world. But, 7,000 were presumably enough for the Balkans. And, right there in the National Press Club, they gave a demonstration of yogic flying, and I just must show you their announcement. Here they were -- yogic fliers demonstrating the mechanics of creating peace. Excellent photo opportunities; the poster said, camera crews were welcome, and all you see are these three young fellows who look like they are sort of hovering. I could not resist this. My office is in the press building, so all I had to do was go upstairs to watch it.

The person leading the whole thing, was Dr. John Hagelin. Well, Dr. Hagelin is a Physicist. He has a Ph.D. in theoretical physics from Harvard University, was Summa Cum Laude from Dartmouth, and at one time worked at the Stanford Linear Accelerator, and published many respectable papers. But the first rule you have to keep in mind is that there is no claim so preposterous that a Ph.D. scientist cannot be found to vouch for it. Remain alert for this one. Hagelin is now the Chairman of the Physics Department at Maharishi University. He explained that meditating to create a peace field is very scientific.

You may ask -- why now? Why now when science is on such a role? There are real breakthroughs in medicine and technology everyday. And yet, we are also seeing more nonsense labeled science. At any rate, I went to the yogic flying demonstration. And right there, a block and a half from the White House, they spread mattresses out on the floor, and 10 to 15 fit looking young guys came in and assumed the lotus position on the mattresses. I would demonstrate the lotus position for you, but I fear that I could not get up again. They assumed the lotus position, and began meditating, and the rest of us were cautioned, to remain perfectly silent so as not to break their concentration. Suddenly, one of the young men levitated. I do not mean that he floated up and out the door or something. He just sort of popped up a few inches and thumped back down on the mattress. You know, I cannot prove it, but I would guess that he followed a parabolic trajectory in his brief launch into space. And a moment later, another one levitated, and then another, and another -- they went thump, thump, thump, back to the mattress. It began to look like corn popping. And right then, I knew that Madeline Albright had made a mistake. She

really should have supported these guys. The sight of 7,000 yogic flyers bouncing around would have rendered the Serbian troops helpless with laughter.

We now see scientific claims everywhere. That's because science is successful. The more successful science is, the more opportunities there are to mislead the public in the name of science. Consider the best selling book by Deepak Chopra, *Ageless Body, Timeless Mind: The Quantum Alternative to Growing Old.* Well, I will take almost any alternative! But, I can tell you that Deepak Chopra knew absolutely nothing about quantum mechanics. The few lines in his book about quantum theory seem to be lifted directly out of Stephen Hawking's "A Brief History of Time," as far as I can tell, was the extent of Chopra's knowledge of quantum mechanics.

We live in a very perilous time right now. The most recent threat is bio-terrorism. Just last week, there was a hearing of the House Government Reform Committee, chaired by Dan Burton of Indiana. Now, Dan Burton as you may know, is a strong advocate of alternative medicine. Alternative medicine refers to untested medical procedures that are not widely accepted in the medical community. There are many of these alternative therapies, and Representative Burton (R-IN) likes them all. A couple of years ago, Burton held a hearing on alternative medicine, in which the lead witness was none other than Jane Seymour, who plays the title role in Dr. Quinn Medicine Woman on TV. The congressman may be having trouble separating fact from fiction

In this hearing, Dan Burton was concerned about bio-terrorism. And, rightly so. We are all concerned about bio-terrorism. But, what worried him the most was that we might run short of medicine if these terrorist attacks continue. He said we might run short on things like Cipro and Doxycycline, two antibiotics commonly used to treat anthrax. Or we might run short of vaccine against small pox - all of these things. But, Dan Burton thought there was a solution. Alternative medicine could substitute if regular medicine ran short.

One of Burton's first witnesses was none other than Stephen Strauss, a medical doctor and also the Director of the Center for Complementary and Alternative Medicine at the National Institute of Health. He is a highly respected medical scientist. There is an interesting story behind this. When that center was first created, Congress picked the directors. And the previous director of that office was a man named Wayne Jonas. But when Harold Varmus left as the director of NIH, his last official act was to persuade Stephen Strauss to take over as the head of the Center. He was now there to testify on the possible use of alternative medicine in coping with a terrorist attack, but he did not say what he was expected to say. In his eloquent, and forceful prepared remarks, he asked whether alternative measures for promoting health, do anything more than prey on people's

fears, and distract them from taking more prudent steps to protect themselves. In fact, it may not even be prudent to combine natural products such as herbs with antibiotics, because of the possibility that the herbs will interfere with the proper action of the drugs. Well, I think that he stated it very well. There had been suggestions from all over the country that you could you could treat Anthrax with various alternative methods, with herbal medications or even homeopathic medications. But in fact, even the dietary supplement "Energia" had its spokesman out, warning the public not to use these natural products, in the treatment of Anthrax. Fearing, no doubt, lawsuits that might ensue. These are untested treatments. One of them that is popular these days is Coloidal Silver, which has been on the list of toxic substances for a long time. At any rate, the testimony of Stephen Strauss had a powerful impact.

There was another witness at the same hearing, whose testimony was more to the liking of Representative Burton. This was the former head of the NIH Office of Alternative Medicine. His name is Wayne Jonas. Wayne Jonas is the author of an "authoritative" book on homeopathy. So what in the world is homeopathy? Homeopathy is a treatment that dates back about 200 years. It was developed by Samuel Hahnemann in Germany. The idea is that you can treat illnesses with something that might produce the same symptoms in a healthy person. But you do this with extremely dilute solutions of the substance. It is not a totally nonsensical idea: it is kind of what we do when we inoculate people against diseases. You give them a little bit of the disease agent. The question is, how little?

I looked in Wayne Jonas' book, to see what common illness homeopathy might be effective for. The list included just about everything that afflicts us. So I asked my assistant to go and buy me a homeopathic remedy. And she said, "for what?" And I said, "it doesn't matter." What she came back with was pills for menstrual cramps. Actually, I had not been suffering from menstrual cramps. The pills looked like aspirin tablets, but were a little smaller.

Now, am I concerned that I am taking things that might be bad for me if I do not actually have menstrual cramps? Well, not really. It tasted a little sweet. In fact, it was lactose, milk sugar. It was nothing but lactose. If I looked on the package, it tells me that the active ingredient is "pulsatilla," an herb that produces cramps. Not to worry. It says that the concentration is 12x. What does 12x mean?

When Hahnemann was treating patients over 200 years ago, he used a method of sequential dilution. The "X" means that it is diluted one part in ten. And the number twelve means that you repeat that twelve times. So, if I do it once, I have one part in 10. If I do it twice, I have one part in 100: Six times, one part in a million, and nine times, one part in one billion. And, according to this menstrual cramp box, we were going to do

this twelve times. A millionth of a millionth. Now, there is not a chemist on earth that could detect it, if it was there.

I then looked up some of the other standard homeopathic cures, and there was one that I liked. This was for diaper rash. Now, what do you suppose that you use to treat diaper rash? Well, you use an herb called "rhus-toxicodendron." And, what do you suppose that is? Poison ivy. Baby has a rash? Give baby a little bit of something that causes a rash, namely poison ivy. Fortunately for baby, it was diluted 30C. What does it mean to be diluted by 30C? Well, "C" means one part in 100. And, the thirty means we are going to repeat it thirty times. Now, that is one part in one followed by 60 zeros. That would exceed the dilution limit. In other words, a pill is unlikely to contain even one molecule of the substance. One molecule in the Atlantic Ocean would be much too concentrated.

I also picked up something in the drugstore for the flu. Go to your drugstore and you can find this on the flu counter. It is called Oscillococoxinum. The box says, "For the symptoms of flu, fever, chill, or body aches and pains." And what is Oscillococoxinum? It is actually derived from duck livers, but it's no threat to the duck population. I looked on the side of the box, and it told me that the concentration is 200C. Now, I can excuse Hahnemann. Two hundred years ago, we did not know Avogadro's number. We do not need to go into detail here, but this number allows you to calculate how many atoms or molecules are in a known amount of any substance. So, you can give me water, tell me how much it weighs, and I could tell you how many molecules are in it. This is really great, it's a very valuable thing to know. Hahnemann and Avogadro were in fact, contemporaries. But the truth is that even Avogadro did not know Avogadro's number. He knew there should be such a number, but it took another 50 years before other scientists determined what Avogadro's number is. Avogadro's number allows us to calculate how many molecules are left in a particular dilution. When you get down to one molecule, you've reached the "dilution limit."

You cannot just dilute something indefinitely, because things are made up of atoms and molecules. So I get down to the point where there is only one molecule or atom of this substance left. After the next dilution, there is unlikely to be a single molecule left. I can go on with dilution after dilution after dilution. So, its important to decide whether you are exceeding the dilution limit of the substance. In the case of Oscillococoxinum, the concentration is 200C. Well, that means one part in 100, repeated 200 times. That would be one part in one followed by four hundred zeroes. There is no reason why you should know, but does anybody have a guess as to how many atoms there are in the universe? Well, astronomers know that number pretty well. It is about one followed by eighty zeros. Oscillococoxinum has vastly exceeded the dilution limit of the universe. So there is

clearly something wrong here. Help me figure out what this means. This is a billion dollar business, a multi-billion dollar business, over all the world. It is a big business within the United States and an even bigger business in Europe. This suggests that there is a lot of room here for litigation. People who take these homeopathic medicines think they are curing themselves. It is a wonder that a few of the purveyors of homeopathy have not been successfully prosecuted.

The whole picture of Science and the Law changed in 1993 with the Supreme Court's *Daubert*[1] decision, which you people are more familiar with, than I am. Basically, the Supreme Court ruled that federal judges have a responsibility to act as gatekeepers, to make sure that juries do not hear scientific nonsense. The previous standard of scientific evidence was that testimony had to be "scientifically accepted". Under the *Daubert* rule the evidence must be scientifically gathered. Science is open to change. That's the good news. The good news is, that science is open. When you get better theories, or you do better experiments, you re-write the textbooks and throw the old ones away.

That is what gives science its legitimacy. The validity of science depends on this simple, self-correcting mechanism. When we get better science, we throw the old ideas away. The Supreme Court recognized this concept. The condition the Supreme Court put in the *Daubert* decision was that scientific evidence had to be gathered, by scientifically valid procedures in order to be used. This was a rather novel concept. That allows the possibility of evolving scientific concepts. What you do not want to do, is to lock in an old idea, since the new ideas are responsible for the progress that we see all around us. In fact, many of my colleagues were pessimistic that judges would adhere to the *Daubert* decision.

A few months ago, however, I was invited to Cold Spring Harbor on Long Island by the Federal Judicial Council to talk to a group of federal judges. They wanted to know about a practice that was beginning becoming more and more common. Judges were creating panels of experts who have no connection with the case that they are presiding over. These experts advise the Judge on what testimony can be admitted. And it is working very well. There have been some very important decisions made on this basis.

One of the first things the federal judges asked me about was the Mass Tort blitz that was attempted during the power line/cancer controversy. A few years ago, power lines were thought by many to cause cancer. We now know they do not cause cancer. That was recognized first by the California Supreme Court which issued a ruling based on information

---

1 Daubert v. Merrell Dow Pharm., Inc., 509 U.S. 579 (1993).

from the National Academy of Science, the American Physical Society, and by a group of prominent scientists and Nobel Laureates who wrote amicus briefs for the Court. The Court decided this was a controversy that belonged to the regulators and not in the courts. This essentially ended the debate on power lines and cancer. In fact, out of all of the suits on health effects from power lines, there was never a successful suit linking power lines to cancer. We see the same debate repeated today in cell phones. I fear cell phones every time I see one being used by the driver of car, but I don't worry about cancer.

So, I set out to develop a set of warning signs that we might use to guide us in similar situations. These are only warning sings and not absolute rules. The first warning sign is that a discovery is pitched directly to the media. Now if I come up with a discovery, the first thing that I am going to do is talk to my colleagues and see if they can point to a mistake in my measurements or reasoning. I will then probably present my work at a departmental colloquium. And, if nobody in the department jumps on me, I will probably give a paper on my new discovery at a scientific conference and if I get away with that, I will probably submit my discovery to a scientific journal which will send it out to a bunch of anonymous reviewers who would like nothing better than to tear me apart, and I go all the way through the process. We rely on our colleagues to keep us straight in science, and this is very important. That check is missing when a discovery is announced first to the press.

A second warning sign, is that there is said to be a powerful establishment, often involving the media, suppressing a new discovery. Some of you are too young to remember this, but, back in the early '70's, there was an inventor by the name of Sam Leach, who claimed to have invented an automobile that runs on water as fuel. He drove that car across the United States and disappeared. I do not mean that he disappeared literally. In fact, he retired a rich man. The President of Budget Rent-a-Car and a few other investors had invested millions of dollars in his invention. Sam Leech even took a reporter from Newsweek on his drive across the United States in this car that supposedly ran on water.

But, there is no car today that runs on water. Whatever happened? Well, Sam Leech says his invention was suppressed by the oil companies, and there are a lot of claims like that. Discoveries that are covered up by powerful interests. That is one of the warning signs that I was talking about; the discovery and the suppression of the discovery. Who knows what the oil industry would have done to keep cars from running on water?

Another warning sign is that information is withheld that would allow other scientists to replicate the results. This violates the basic rule of science. Science is open. We submit our publications, our ideas, our find-

ings to the scrutiny of other scientists who can build their reputations by tearing us down. And, we do this freely, because this is the process that gives science its legitimacy.

Another warning sign is that the evidence for a scientific discovery relies heavily on anecdotal evidence. Anecdotal evidence consists of personal stories: Uncle Fred took such and such a medicine and he got out of his hospital bed and threw away his crutches. You hear these stories all the time and the public is very strongly persuaded by such anecdotal evidence. A good example of anecdotal evidence is the magnet craze. If you want to see magnets being used, go to a golf course. I don't know what it is about golf courses. Lawyers go to golf courses more than they go to work. I am not a golf expert, but, 20 years ago, just about everyone on the golf course was wearing copper bracelets. They said that wearing copper bracelets suppressed arthritis and made their joints move more freely. Copper bracelets must have stopped working. But you can still buy copper bracelets at golf pro shops and I see them occasionally on the golf courses.

More common today, is wearing magnets. The wearing of magnets also started on golf courses, but, now it has spread everywhere. I have some problems with this. I look at this ad for a particular magnetic therapy device, and in the photo on the right, I see a woman wearing a magnet on her wrist. This troubled me. When I walk, my wrist passes right by my hip pocket and my hip pocket is where I have my credit cards. This ad says that the magnet is 800 gauss. Now, 800 gauss is plenty strong enough to erase my credit cards. This company couldn't have wanted my credit cards erased. So, I went out and bought a magnetic bracelet. It really was very handsome. It consisted of a nice blue velvet strap, which goes around the wrist, and then, the magnet is in a little Velcro pouch that attaches to the velvet wristband. So, now that I have got it around my wrist, this presumably is helping whatever problem was affecting my joints. The claim was that the magnet attracts blood because blood has iron in it. Let's talk about this.

The iron in blood is in the chemical form of hemoglobin, which is not ferro-magnetic. It turns out to be diamagnetic. That means blood is actually repelled slightly by a magnet. Next, I checked to see if the magnet was really that strong. I measured it right at the surface and just like the magnet company had said, the magnet was 800 gauss. But, less than an eighth of an inch away from the surface, I was unable to measure any magnetic field at all. There was a reason for this.

One of the neat technologies that we all use and never even think about is a refrigerator magnet. Refrigerator magnets are a fairly new development. This is clever stuff. You have this refrigerator magnet that looks like your local pizza delivery truck with their phone number on it and you

post a Dilbert cartoon on your refrigerator with it. If you ever notice, it has a painted side with the phone number and the other side, which goes against the refrigerator, is black. If you try to put it on backwards, with the black side out, it does not stick. How do they do that? How do they make a magnet that is only magnetic on one side? In fact, if you've ever tried putting more than one piece of paper under it, all of a sudden, the magnet falls off the refrigerator. They do not weigh much. The first piece of paper I put held up pretty tightly. But, as I added more and more paper, the paper and then eventually the magnet fell off. Well, take a look at my therapy magnet that I bought, placed next to a paperclip. A paperclip is a good test for a magnet. It sticks to a magnet quite strongly. But now, when I put my therapy magnet on the velvet strap, the way they show you to use it, the side that would be next to my body there has no magnetic field at all. Now, where did it go? Well, this is how refrigerator magnets are made. They are made in narrow strips, with alternating north and, south poles. And so, right at the surface, the field is very strong, but a short distance away, the north pole cancels out the south pole. A very short distance away, there is really no magnetic field at all.

Most people I talk to in the public are convinced that if people are out there in a billion dollar business, selling magnets, these magnets must work. However, these magnets have such a short range that the field may not even extend through the skin. Little, if any, magnetic field reaches the muscles and joints where the injury is. People just cannot believe that there is not a federal agency that would stop people from selling these magnets if they don't work, or would prevent people from selling worthless sugar pills.

Incidentally, the osillicoxinum in this box, consists of six doses. It cost me $11, exactly what I paid for a flu shot. Now, which one, the sugar pill or the flu shot, is going to help me the most? These sugar pills get around federal regulations quite freely however.

This is a full-page ad from USA Today. Full-page ads do not come cheaply and this is the most read newspaper in the world. Here are these happy, smiling people, the picture of health. They seem to be saying they are happy because they take Vitamin O. What in the world is Vitamin O? You cannot read the words, so I blew some of them up. Here is what it said: "It is so safe that you can drop it into your eyes. It is so natural it contains the most abundant element on earth. It is so effective you can spend hours reading the unsolicited testimonials of those who have used it with dramatic results." But, how does Vitamin O do this? "The bloodstream absorbs the Vitamin O and carries the pure oxygen straight to your cells and tissues, where it maximizes your nutrients, purifies your bloodstream, and eliminates toxins and poisons. All the processes necessary to prevent disease and promote health." But, what is in this Vitamin O? The

amazing thing to me, is that the ad tells you exactly what it is. "Vitamin O contains stabilized oxygen molecules in a liquid solution of sodium chloride, salt, in distilled water, to be taken orally as a supplement: usually 15-20 drops, two to three times a day." Well, what are these "stabilized oxygen molecules" the ad talks about? It is ordinary oxygen; the stuff you are breathing right now. The question is, how much of my oxygen am I going to get from 20 drops of water? Well, codfish get their oxygen from water. For humans, it is known as drowning. In fact, you need only two numbers. First is the solubility of oxygen in water. You can get about 8 milligrams of oxygen in a liter of water. How much do we use? Well, just sitting there, unless you are particularly agitated at me, you probably are consuming 15 to 20 milligrams per second. So, how much water am I going to have to drink? I am going to need a couple of liters every second. And that would be if I could absorb the oxygen from my gut, which I simply cannot. That is why we have lungs. The lungs do a fantastic job of it. I mean, if I was a trained athlete and I was performing at a high level, I would be using perhaps 130 milligrams per second. We humans are very good at this. We have the right apparatus up here in the lungs to do this, but when I put it into my gut, it is going to wind up in my bladder. It is not going to end up in my bloodstream. It would not matter if it did, because there is such a tiny amount of it. Well, so what happens? I published this information in my column and Science Magazine picked it and published it in their magazine. Someone from National Public Radio read Science and I was invited to talk about it on National Public Radio.

And, where was the Food & Drug Administration? Well, the Food & Drug Administration could not touch it, because it was a natural supplement. The 1994 Dietary Supplement and Health Education Act, passed by our Congress, prevents the Food & Drug Administration from requiring proof of either efficacy or purity for natural supplements. So, when you buy these dietary supplements, you have no assurance that they work or even that you are getting what it says on the package. It is kind of shocking.

There is another agency that should look after us and that is the Federal Trade Commission. The Federal Trade Commission stepped in and charged the company selling Vitamin O with false advertising. The company was shut down, and, I was feeling like I had finally accomplished something. However, in the mail one day recently, I received a flyer, advertising Vitamin O. They were no longer advertising in USA Today and they had changed the name of the company and changed the address to the next block. Well, it was obviously the same company, selling the same product. And, in fact, as I searched the web, I found at least five companies that were selling what amounted to Vitamin O. Essentially, they were all selling ordinary water, at prices like $40 dollars an ounce. And, they

even sell this to the hockey teams. I don't know if it's gotten to football yet, but in hockey, the teams all use oxygenated water to give the players an extra boost. It may send them back to the men's room prematurely.

The most important discovery in medicine is not antibiotics, nor vaccination. The greatest achievement of medical science is the randomized, double-blind test. Medicine was not a science until the double-blind test was invented. The fathers of this nation, for example, all believed that bleeding was a cure for almost anything that bothered you. Thanks to the double-blind test, we now know that it is not only not helpful, but it is harmful.

Another warning sign is that a discovery is said to be credible, because it has endured for centuries. This is my favorite on this whole list. It's the old Ancient Wisdom myth. Somehow, before people knew that blood circulated, or that there was a germ-theory disease, before they knew any of these things, they had these magical cures that we modern people have lost somehow. Well, I am waiting to be shown one of these that actually works. But there are, in fact, hundreds of health devices, methods, therapies, on the market right now, that rely on this idea of Ancient Wisdom. Homeopathy, for example. It was discovered, if discovered is the right word, about two hundred years ago. Now, that is pretty recent compared to some ancient wisdom. Things like acupuncture, are thousands of years old. And, sound medical evidence that they work is still missing.

There are millions of people that swore by bleeding for example. The discovery was made in isolation. You know, this is not the tricky one. We have this myth of the lone genius laboring all by himself in his little workshop in the attic. There he makes some wonderful discovery. But, this is a myth that is not really impossible. It is certainly possible that some lone inventor somewhere will make a really important discovery. But, by and large, science is a very closely-knit fabric. We all depend for every step we make in science, on things that are learned by our colleagues. Yet another warning sign is that no plausible theory at all is offered for a miraculous discovery. That again is the one thing we hear about most in the medical field.

I do not know if this has helped you in what you are here to discuss today. But, we all have to be concerned that the public is being misled by false scientific claims. Scientists tend to think we would not have a problem if we had a more scientifically literate public. I ask, what it is that you want society to know? I mean, I have trouble keeping up in my own field. Most scientists do because progress is being made so rapidly today. The idea that the public should know the solubility of oxygen in water, for example, is totally unrealistic. I did not know what the solubility of oxygen in water was until I looked it up. But, we cannot expect the public to even know where to look. We live in an age of experts. And, it is the

responsibility of the scientific community, to speak out on these issues. I think we have neglected our responsibility to a shocking extent. Scientists are busy people, they enjoy what they do, and focus on their own research. They do not want to take the time to get involved in every false claim.

Sometimes, however, scientists avoid getting involved for fear of being sued. My book was held up for a couple of months by threats of lawsuits from people who didn't like what the book said about them. They had obtained prepublication review copies. Oxford Press then hired lawyers to go through the book, and satisfy themselves that every line could be documented.

Scientists, by and large, are not paid enough to risk lawsuits. My wife probably wishes I wouldn't take that risk. It certainly inhibits scientists from speaking out on issues that are very important to the public. Scientists will have to learn to work with the legal profession. A lot of this depends on you, the future lawyers. Take care. Thank you.

Using Court-Appointed Experts

The Honorable Timothy Hillman*

Good Morning. Following Professor Park with Federal Rule 706[1] is like following Drew Carey with a great business meeting. I will do my best. I wanted to speak with you about something I think is very important, in light of all the scientific information that we are being asked as courts to digest. That is, consideration of courts and counsel in using Federal Rule 706,[2] which allows for court-appointed experts. At the risk of being very dry, I want to speak a little bit about the mechanics of Rule 706.[3] I also want to talk about some practical uses of the rule, as well as some instances where I have considered its possible uses, although I have not personally used it. But, frankly, it's coming, and it's going to happen sooner, rather than later.

The general authority for this Rule has long-standing effect; Professor McCormack traces it back to the 14th Century. As long ago as 1962, however, our federal courts said that the court's discretionary appointment of an impartial expert, in this case, a medical expert, is an equitable and forward-looking technique for promoting the fair trial of a lawsuit. It is now well accepted that the trial judge is not a mere umpire at the trial. Indeed, there may be circumstances in which he or she may have a duty to seek impartial assistance in order to enlighten the jury and themselves, because the issues have been confused by partisan presentation.

The courts have moved forward since 1962. The mechanics of Rule 706[4] tell a lot about the ability of the court to use it, and the ability of counsel to recommend it to the court. First of all, the court may have its own motion, along with a motion of any party, to enter an order to show cause for why an expert should not be appointed.[5] The court may ask for

---

* Judge Hillman has been a Judge of the Massachusetts Superior Court in Worcester since 1998. From 1991 to 1998, he was an Associate Justice and Presiding Justice in the Gardner District Court and the Presiding Justice in the Worcester District Court. He has prior experience in private practice, the Worcester County District Attorney's Office, and as Town Counsel and City Solicitor. Judge Hillman is a 1973 graduate of Suffolk University Law School. He received his bachelor's degree from Coe College in 1970. Judge Hillman currently teaches Law & Psychiatry at the Massachusetts School of Law.

1. FED. R. EVID. 706.
2. *Id.*
3. *Id.*
4. *Id.*
5. FED. R. EVID. 706(a).

nominations.[6] More importantly, the court may ask for a limitation on the scope of the expert's particular ability, or for a particular focus on the areas under which the court or the finders of fact need enlightenment. The Rule also provides that the court must inform the expert and the parties, in writing, of the scope of their duties. Of course, this writing must be filed with the clerk of the court.[7] This process of narrowing down, shows the finders of fact what the court specifically wants. Is it judicial touring? Does it do anything for the jury? Is it some sort of camp order, or something along those lines? A witness can be deposed by either party and can be called by either party.[8] A neat little part of Rule 706 says, when the rule is raised by either party, the expert witness may be cross-examined.[9] In other words, the expert being called for an adverse party, under the rule, may be cross-examined by different techniques. One of the problems with this practice is addressed within the rule in as clear a method as possible. First, court-appointed experts are paid from funds provided by law in criminal cases, or in civil cases where somebody invokes a privilege.[10] In civil cases, you have to wait to see what constitutes a privilege, and you may want to get a court-appointed expert if there are funds available for indigents that may be used. In civil cases, the cost of the expert is as the court directs, thereafter, designated as cost. The court may tell the finder of fact that the person is an expert. For those of you who practice in Massachusetts, the Massachusetts Rule of Criminal Procedure 41,[11] allows for the appointment of interpreters. It also puts the words "or expert" in the rule.[12] There are no decided cases on it to date. When mental health competency issues exist, the court has a responsibility regarding these issues, although this is not spelled out in Rule 41.[13] This is when court-appointed experts come into play, even though we do not call them that. For those of you who practice in Massachusetts, there is civil authority for such appointments. In basic cases, the Superior Court judges have needed to buy supplies, such as tape recorders. This example of the overall superintendent's powers of the court is somehow dovetailed in the Rule 706[14] analysis. There is no Massachusetts ruling on Rule 706,[15] however, its power is present, apparent, and I see no problem with it.

---

6. *Id.*
7. *Id.*
8. *Id.*
9. FED. R. EVID. 706(a).
10. FED. R. EVID. 706(b).
11. MASS. R. CRIM. P. 41.
12. *Id.*
13. *Id.*
14. FED. R. EVID. 706.
15. *Id.*

Rule 706 has many uses.[16] The use of this rule that bears the most scrutiny, and the greatest potential for use, is the use of court-appointed experts as technical advisors to the court.[17] One example that demonstrates the many potential uses of Rule 706[18] deals with a particular patent selling issue. In Worcester, there is a bio-medical center. As a result, there are many bio-medical cases that arise. I currently have a dispute before me over a settlement agreement on a nuclear transferred-derived cell. I have no idea what a nuclear transferred-derived cell is. Frankly, the issue is actually over the settlement agreement. At some point, however, I told the court that I am going to have to learn what a nuclear transferred-derived cell is. I think under certain appropriate circumstances, I have the power to appoint an expert to explain what it is, and use the explanation in application to the settlement agreement. Another use of a court appointed expert is as a judicial tutor on esoteric subjects, which dovetails with the last use of Rule 706.[19] Finally, the last use of the rule, with regard to court appointed experts is with respect to witnesses.[20] When using the rule in application to witnesses, you have to be careful.

We recently started trying these applications of Rule 706[21] in Worcester, which are purported to be the first cases in the United States involving such uses of the Rule. I do not know if I have any authority for this, other than what the lawyers have told me, but we have a number of cases from younger people who were hemophiliacs who were transfused in the early 1980s and mid-1970s with HIV-infected blood. Those cases are now coming forward. The issues before the court are ones of informed consent, and when the companies knew or should have known about the effect of the blood. In addition, the lawyers were pretty good, and they did not just tell me the authority of Rule 706,[22] but on their own, they basically designated the plaintiff's expert as the historical expert to explain the progression of the disease. The expert went in, and as awful as it was, it was helpful to the finder of fact, and it laid a foundation. Then, we were able to go back and argue about who knew what, when and what they did. As a foundation, it was very helpful. I do not know whether the plaintiff and the defendant shared the expense of this expert, but they both agreed on having the historical expert, and they both had their points that they wanted to get from the expert. As a result, there was very little cross-examination. It was mostly direct examination and it was very helpful.

---

16. *Id.*
17. FED. R. EVID. 706(a).
18. *Id.*
19. FED. R. EVID. 706.
20. *Id.*
21. *Id.*
22. *Id.*

The problems: Problems are fairly common with Rule 706[23] and court-appointed experts, which rely on the common law tradition against courts getting involved in the fact-finding process. This way of thinking is traditional and long-standing. It is ingrained in the culture. But, times have changed. In fact, when Judge Lauriat is here this afternoon, I hope he gets into some innovations that he and the Superior Court have been doing. I use the word "innovations" because I am almost ashamed to tell you that they are only recently allowing jurors to take notes and ask questions during trial. These are considered innovations in Massachusetts. The importance of these new innovations like court-appointed experts and juror note-taking is that we are now getting more and more involved in the fact-finding process, in order to make sure that the jury is searching for and reaching the truth. As mundane as that may sound, this is what they are supposed to be doing. Jurors taking notes and asking questions is inevitable, if we want jurors to be involved in this search for the truth. Judge Lauriat has been a real pioneer, and I followed along in his lead. Our Supreme Court in Massachusetts has allowed juror questions, even in an assault and battery with a dangerous weapon case, and there were nearly forty questions asked.[24]

There are two problems. The first problem is the legitimate fear of ex-parte communications. The second problem is the expense. It is obviously expensive to ask an indigent plaintiff, or some plaintiff in a medical malpractice case or a criminal defendant, who does not have many resources, to share in the expenses. Let me give you some brief examples. We recently had an awful mess, involving a junkyard out in Gardner and the issue over cleaning it up. The defendant was clearing in violation of a zoning ordinance. The junkyard clearly needed to be cleaned up, and I considered appointing a court-appointed expert to help the Court in deciding what goes to the dump, what is good, what needs to go where and assessing the cost. All of this information was successfully determined through the use of the expert.

Court-appointed experts and asbestos litigation seem to be common too. Judge Lobel has a session over in Middlesex devoted primarily to this type of litigation, in which court appointed experts prove to be key players. Complex medical cases also seem to lend themselves to the issue of court-appointed experts. One very well known and lengthy litigation involved breast implantations, in which the judge who consolidated and tried the cases, used a detective used him well, and with approval of the federal courts. Then, of course, there is the novel, scientific evidence that Profes-

---

23. *Id.*

24. *See* Commonwealth v. Burgos, 739 N.E.2d 717 (Table) Mass. App. Ct., 2000. Dec. 8, 2000.

sor Park talked about. For example, the Genome Project in Worcester dealing with bio-meds, and all of the scientific projects that go on in this Commonwealth are perfect for using experts under Rule 706.[25] I strongly urge you to consider this rule when practicing as attorneys. There is plenty of foundation for it. I know I am going to consider using it, so you can beat me to the punch and get brownie points if you suggest it yourself before the court. Let me just end with a comment from the Federal Rules on Rule 706[26]: The practice of shopping for experts, the venality of some experts, and the reluctance of many reputable experts to involve themselves in litigation, have been matters of deep concern. Though the contention is made that the court appointed expert acquire an aura of infallibility to which they are not entitled, the trend is increasingly to provide for their use. While experience indicates that actual appointment is a relatively infrequent occurrence, the assumption may be made that the availability of the procedure in itself decreases the need for resorting to it. The ever-present possibility that the judge may appoint an expert in a given case must inevitably exert a sobering effect on the expert witness of a party and upon the person utilizing his services.[27]
Thank you.

25. FED. R. EVID. 706.
26. *Id.*
27. FED. R. EVID. 706 advisory committee notes.

The Expert Witness

Brian Holmgren*

There were three professionals who were debating the oldest profession in the world: a surgeon, a physicist, and a lawyer. The surgeon said he believed that surgery was the oldest profession in the world because on the sixth day, God took Adam's rib and created Eve, and that must have been done by a surgeon. The physicist said yes, but before God created Adam and Eve, he created the heavens and earth out of chaos, and that was done by a physicist. The lawyer then piped up, "who do you think created all the chaos?" The fact of the matter is that the law has created its own form of chaos when it comes to the subject of expert witnesses. The issue of the day is, at least as far as I am concerned, do experts advance the truth furthering process of the trial? That is what they are supposed to do. Do they really do that?

There are many expert witnesses for whom it can be said - "if you pay them, they will come." We call these the "Field of Dreams" experts and they do come to court on a very frequent basis. The law contributes to this process, and this process will be the focus of my remarks.

How does the law help to create this body of chaos for us? First of all, we have to understand the jury selection process. Or, actually, I should refer to it as the jury de-selection process. We do not actually go out and pick the jurors we want. We go out and get rid of the jurors that we do not want. We begin this process by getting rid of the jurors who know anything about the parties, the subject matter of the litigation, or the technical issues which are the subject of the litigation. Basically, we ensure a fair

---

* Brian Holmgren is an Assistant District Attorney General with the Davidson County District Attorney General's Office in Nashville, Tennessee, where he is assigned to the Child Abuse Unit. He previously served as an Assistant District Attorney in Kenosha County, Wisconsin, where he was the Director of the Sensitive Crimes Unit. Between November 1995 and July 1999, Mr. Holmgren was a Senior Attorney with the American Prosecutors Research Institute's National Center for Prosecution of Child Abuse. He is currently on the Board of Directors of the American Professional Society on the Abuse of Children. Mr. Holmgren also serves on the National Advisory Panel on Shaken Baby Syndrome and the Board of Advisors for the Civic Research Institute's *Sex Offender Law Report*. He is a contributing author and editor to the third edition of the National Center's highly acclaimed manual on the *Investigation and Prosecution of Child Abuse Cases*.

jury, by ensuring an ignorant jury. This is then balanced against the criteria for qualification of an expert witness. An expert witness is defined in the law as somebody who can assist the fact-finder in the process of advancing the so-called search for the truth. If we get rid of anybody on the jury panel that knows about the issues in the case, and if the qualifications of the expert are defined as being someone who has knowledge beyond the scope of the juror's knowledge, then someone who can advance the juror's knowledge even just a little bit can be qualified as an "expert." That is the definition of an expert witness in court under our laws. Mark Twain said over a hundred years ago: "the jury system puts a ban upon intelligence and honesty, and a premium upon ignorance, stupidity and perjury."[1] Those words are no less true today, than they were one hundred years ago.

How else does the truth remain hidden when we get to court? First of all, we have this very low threshold for the admissibility of expert witnesses; therefore, it is very easy for the expert witness to make their way into court. Generally, when courts talk about the qualifications of the expert, the courts suggest that an expert's lack of qualifications go to the weight of their testimony, rather than to their admissibility. Expertise is often established by the fact that the expert has appeared in court several times before, which begs the question: if they did not belong there in the first place, then why does it weigh as part of their credentials later on? The expertise of an expert may be limited in his or her practice to a specific area but, once they get into court, that expertise may be broadened ten-fold. The expert may offer opinions on things that go well beyond his or her specific area of expertise. While we would hope that the rules of law would somehow preclude this possibility, in fact, it is exactly the opposite.

So, we get into court and we have these wonderful new rules that say the courts are supposed to act as gatekeepers regarding the admission of expert witnesses. We are supposed to determine whether experts are or are not qualified. We are also supposed to determine what they should or should not be able to say in court. Well, I have news for you folks. All of you went to law school because you did not want to learn anything about science, right? This is a fundamental truth. So now, we are asking you when you become lawyers to know and ask the right kinds of questions of experts that qualify them or disqualify them as experts. We ask judges who do not have training in this particular area to make these important decisions. It does not work so well.

Another problem is that the Federal Rule of Evidence regarding expert

1. MARK TWAIN, ROUGHING IT ch. XLVIII (1872).

witnesses, Rule 702,[2] is not adopted in every state. A lot of states do not even have a specific rule on expert testimony. Those that do have such a rule labor under different standards in its interpretation, such as the *Frye*,[3] *Daubert*,[4] and *Kumho*[5] standards. Some states, however, do not have such standards. Rule 702 was recently revised to tighten up the doors a little bit on expert testimony and yet, a substantial number of the jurisdictions have not adopted the revised Federal Rule 702,[6] even though they may have Rule 702 in their books.

Then, we get into the issue of who is the "best expert" under the Rule? For example, in the field of child abuse we have nurse practitioners who do a lot of the sexual abuse examinations of children and who are then asked to offer testimony about their examinations. Are they the best experts in that situation, or do we need a medical doctor in order to be able to provide expert testimony?

Another problem is that the ethical standards for practitioners, especially in the health-related field, are vastly different than they are for psychologists or other experts. It is much easier for practitioners in medicine to come forward and offer testimony that would, by other standards, in other professions, be considered unethical. But, they make their way in. I am singling out medical doctors to a large extent because this is where I see the bigger problems, at least in terms of the criminal field.

Now, in the civil arena, we have a little bit of a balancing to some of these problems because attorneys are allowed to depose witnesses. They are allowed to hire their own experts, and they have a little bit more access to funds, generally, than we do in the criminal field. In the criminal field, a lot of times we do not even know what expert is coming, or what they

---

2. "If scientific, technical or other specialized knowledge will assist the trier of fact to understand the evidence or to determine a fact in issue, a witness qualified as an expert by knowledge, skill, experience, training, or education, may testify thereto in the form of an opinion or otherwise." FED. R. EVID. 702 (1983) (revised 2000).

3. Frye v. United States, 54 App.D.C. 46, 293 F. 1013 (1923).

4. Daubert v. Merrell Dow Pharm., Inc., 509 U.S. 579, 113 S.Ct. 2786, 125 L.Ed.2d 469 (1993).

5. Kumho Tire Co. Ltd., et al. v. Carmichael, etc., et al., 526 U.S. 137, 119 S.Ct. 1167, 143 L.Ed.2d 238 (1999)

6. "If scientific, technical or other specialized knowledge will assist the trier of fact to understand the evidence or to determine a fact in issue a witness qualified as an expert by knowledge, skill, experience, training, or education, may testify thereto in the form of an opinion or otherwise, if (1) the testimony is based upon sufficient facts or data, (2) the testimony is the product of reliable principles and methods, and (3) the witness has applied the principles and methods reliably to the facts of the case." FED. R. EVID. 702.

are going to say, until they arrive at the courtroom door. This practice, of course, has many vices, not the least of which is its allowance of experts who really ought not to be there to say things that really ought not to come forward in court.

There is another problem in the criminal field which is that there is a judicial bias towards permitting expert testimony by the defense and conversely a bias toward the exclusion of expert testimony by the prosecution. This comes from several different sources. One is the presumption in favor of allowing the defendant, as part of his due process rights, to present a defense. Second, is a concern on the part of the judiciary who fear reversal if they exclude expert testimony and that decision is reviewed on appeal. The prosecution, by contrast, generally does not have the right of appeal if their experts are excluded, so it is easier to rule them out. A third factor, that has substantially contributed to this problem, is the tendency on the part of many prosecutors to push the envelope when seeking convictions by calling in experts who are going to advance their theory of the case. Prosecutors, by calling in their experts in an improper manner, cause these cases to go up on appeal. We then get case law rulings against such use of expert testimony that impacts future cases. Therefore, it is incumbent on somebody in my profession to make sure that in every case that I litigate I do the job right, so that I am not creating that first case of adverse precedent for all the other practitioners across the country.

I want to give you two examples of how this plays out in the real world. Recently, there was some litigation in a federal jurisdiction in the South over the testimony of a proposed expert who wanted to come in, and basically testify that he had examined, or run an assessment technique on an alleged sex offender, and that basically this sex offender did not show any deviant arousal towards children. Now, the expert that came to satisfy the admissibility requirements in a *Daubert* hearing in this federal court, happened to be the very person that developed this particular assessment technique. So this expert had a vested financial interest in the outcome of the court's ruling, because he could then say basically, and he did say ultimately, that the Federal Courts had accepted his technique under the *Daubert* criteria. Now, mind you, this gentleman had served on the Board of Directors at the largest professional organization dealing with sex offender treatment and assessment in the country. The Code of Ethics for that particular profession says unequivocally, that no expert witness should come to court, and basically tell the court that an offender does, or does not, have a pre-disposition to engage in this behavior, in terms of the guilt-innocence phase of the trial. So, this man directly violated this Code of Ethics. He came before the court, offered this opinion, and got the technique accepted even though there was well-published literature that said that nobody should do this and that we do not have reliable information in

this regard. His testimony was accepted and this expert now has this court's decision posted on his website. So the defense wins a victory and they get this technique admitted into the trial courts, even though this assessment technique has little scientific reliability in terms of the evidentiary proposition for which it is being offered – that the defendant is not likely to have committed the alleged sexual offense. In fact, the evidence is totally unreliable on this point.

Contrast that scenario with a case that I was involved in recently, which involved the asphyxial deaths of three children by a single caretaker. In order to support the proposition that this woman killed these children by smothering them and leaving no traces, we brought in a statistician who examined the probability that three deaths to children of this age could have occurred by chance in that situation. The statistician basically came to the conclusion that the probability of that occurring exceeded the world's population. This was very reliable statistical proof that this woman smothered these three kids. The court, hearing that proof, clearly made a determination that it was reliable and that it satisfied all of the *Daubert* criteria; but, the court excluded the testimony, at least in the case in chief, on the basis that it would be overly prejudicial to the defense to permit this evidence to come before the court and before the jury. So, on the one hand, very reliable information is excluded, even though it is scientifically sound, and on the other hand, unreliable information is admitted. This dilemma is one that we face in the courts every day.

We contribute to this problem in a couple of other ways. We ask the expert, when they are coming into court, to offer information that is going to assist the court in the determination of what should be admitted - what should and should not be admitted. Well, think about this ladies and gentlemen! If the irresponsible expert is coming to court and being asked to basically offer information about whether they should, or should not, be permitted to testify, the reliability of the information that the court is going to receive to make that determination is inherently suspect. So, we have a big problem. The problem has to do with the issue of whether this information is going to be accurate, or not, given that we are dealing with irresponsible experts.

On the other side of the coin, we also have many responsible experts who come to court. These individuals are being asked to go through the same process, but they are also being asked to try to train and educate judges, juries, and lawyers, who do not know anything about that subject material, to make appropriate determinations. We are not well suited in the law to make that determination, certainly not in a criminal field where we have to make these determinations fairly quickly. It is rare that judges would take a day, or two days, out of their schedules, in the middle of a trial or prior to trial, to make a determination of whether or not the par-

ticular expert's testimony is, or is not, reliable. Rather, they would put it before the jury and let the appellate court sort it out later on. Again, this is another bias. If we do not know who these experts are, what they are going to say, and what it is that they are relying on, we certainly are not going to be in a position to cross-examine, or examine them in an appropriate fashion. Further, the rules in many jurisdictions do not permit prosecutors for the State to find this out by deposing defense experts that are going to testify. Therefore, we are often flying blind in many of those circumstances.

In the jurisdiction where I prosecuted in Wisconsin, we had an open rule on the admissibility of expert witnesses that was supposedly counterbalanced by a more liberal rule in terms of cross-examination. However, we did not have any discovery rules for the expert witnesses. It was wonderful! I could question the expert all day long if I wanted to, as long as the judge was tolerant. Still, I did not have a clue what he was going to say before he actually got there. Under these circumstances, you learn to try a lot of cases by the seat of your pants, which is not very good for furthering the truth-seeking function of the courts.

Daubert and *Kumho* have also created a rule of thumb that says that one of the things we are going to look to is whether or not the science is published in a peer review journal. That is all well and good, but there are lots of publications out there who uphold very low thresholds and standards for admitting particular scientific evidence. For example, there is a well-known defense expert who likes to go around the country and talk about how children can receive very serious injuries from short falls. One of the articles that he cites is an article that was published out of England in 1993. It surveyed twenty-eight children, a very small sample size, as you can well imagine. The conclusion of the study was that, somehow, Caucasian babies and African-American babies had different rates of sustaining subdural hematomas when they fell. The conclusion one draws from this data is that there must be something physiologically different about Caucasian and African-American babies to account for these findings. This is a peer review published, scientific article that gets cited in a court of law, and it is absolute garbage. There are other studies and examples just like this one, that permeate their way into the so-called truth-furthering function of the courts.

The final point that I want to make is that we see lots of different kinds of experts in the courtroom. When you look to apply, or try to apply, the different criteria from *Frye*, *Daubert* and *Kumho* to some of these problems, and the different types of expert witnesses, things just do not quite mesh. We have social scientists and psychologists that frequently come in and offer testimony in courtrooms. We have chemists, physicists, and accident-reconstructionists that come to court and base their opinions on

more concrete forms of science. Then we have physicians and doctors who base their opinions on different things. Across the spectrum, we are asked to determine: can the expert offer opinions on the ultimate issue? Can the expert offer opinions that deal with the issue of providing information from which the jury can then extrapolate other information, even though the expert is not giving an ultimate opinion?

The rules, as they apply to these different experts, are difficult to interpret. They vary because of the manner in which a social scientist and a psychologist can testify, and the manner in which a doctor can testify may be very different, yet the rules of evidence are the same for both. Moreover, we read case law decisions that seem to apply different standards to those different individuals in ways that do not make a whole lot of sense. So, it is very difficult for us to try to put together and fashion out a consistent rule of law that will apply to every individual expert across the board.

Yet, that is exactly what *Kumho* and *Daubert* seem to tell the courts that they have to do. Now, I find it ironic that in the dissenting opinion in the *Daubert* decision, Chief Justice Rehnquist basically said: look, I do not like this rule, because I do not think that the judges should be in a position of being amateur scientists and determining the reliability of information. We are not trained to do that. We are not skilled at doing that. We ought not to be placed in that position of making our decisions based on having to learn all this science in the first place. Later on, one of the other justices, in a different decision determining the admissibility of polygraph evidence, which, as you all know, has historically been excluded, in a concurring opinion[7] said: well, I think that we ought to allow all kinds of expert testimony to come in because we allow all other kinds of unreliable evidence to come before the court. We have eyewitness testimony all the time, which is really unreliable. We permit that with no restrictions whatsoever. So why not permit experts to come in and offer opinions on polygraph evidence, and let jurors sort it out?

So, on the one hand, one Justice of the United States Supreme Court is saying: as judges we should not have to figure this out, and we are not really equipped to make decisions on the reliability of scientific evidence and expert testimony. On the other side, another Justice of the United States Supreme Court is basically saying: let us just give it to the jury, because they were too stupid to get out of jury duty, and they certainly can figure out this material. So that is the quagmire that all of you are about to enter into, those of you who are studying to be lawyers. Those of you who already are lawyers, you know what I am talking about and dealing with.

---

7. United States v. Scheffer, 523 U.S. 303, 118 S. Ct. 1261 (1998) (Kennedy, J., concurring).

This is the dilemma that we face with the expert witness in court.

Thank you.

Evidence-Based Medicine

Dan Mayer, M.D.*

Thank you very much. First of all, I just want to thank you all for inviting me here, and letting me talk about "Evidence-Based Medicine," which, over the years, I have become very passionate about. The way I begin my course, is by telling the medical students that everything I learned in medical school has changed, even to the point of saying that anatomy has changed. That is not to say that anatomy is different, but, now-a-days, we can image parts of the body with MRIs and see things that we could never see before without cutting people open. What I am going to talk to you about today is "Evidence-Based Medicine." I will tell you what it is and give you some examples of how it is used in medical malpractice cases. Then, I will talk a little bit about what the EBM process, or "Evidence-Based Medicine" process is, and how it works.

In the past, medicine was based on what a bunch of gray-haired experts believed and, since now I have gray hair, I can count myself among them. Basically, based on what "we" said medicine should be, "we" determined how medical practice should occur. I call this "Eminence-Based Medicine." In the law, this translates that the person with the biggest and heaviest CV (curriculum vitae) is going to have the most sway with the jury. The jury may, or may not, have gray hair.

However, over the last ten years, in medicine at least, there has been a radical change in the way that we have done things. A lot of this has to do with the changes in the way in which we are going to finance and manage health care. Physicians are being asked to justify their costs; everything from what pills we prescribe, to what tests we order, to how many tongue-blades we use. These are all going to be scrutinized under the microscope and many will ask, why are you doing it this way? Why are you not doing certain medical procedures in a more effective and less costly way?

---

* Dr. Mayer is a Professor of Emergency Medicine at Albany Medical College, where he teaches a four-year course in Evidence-Based Medicine to medical students. He is also an attending physician in the Emergency Department of Albany Medical Center Hospital. Dr. Mayer is a member of the Society of Academic Emergency Medicine, a member of the Association of American Medical Colleges and was named "Physician of the Year" for the Regional Emergency Medical Organization region of New York State in 1993. He has previously served on the National Association of EMS Physicians and the Regional Emergency Medical Organization. Dr. Mayer has a great deal of experience as an expert witness and had addressed the issue of evidence-based medicine and error prevention at a number of national conferences.

This line of reasoning stems from studies that were done in the 1970's, which showed there was a huge variability in the way medicine was practiced in the New England area. Specifically, the studies looked at contiguous counties in the New England area, and noticed that hysterectomy rates varied by as much as three-fold from one county to another. The people who did the study asked, why was it that some obstetricians were not taking out enough uteruses? A similar difference in rates applied to men and prostate surgery. In fact, there was a three-fold variation in prostrate surgeries across contiguous counties. Obviously, there were different people doing the surgery. Maybe there are some obstetricians who just need to make an extra car payment, or something like that. I do not want to put down obstetricians, but there was no good or rational reason as to why there was this variability from one area to another.

When physicians were asked about why they did things, they replied that it was because that was how things had always been done. Nowadays, this would not be a good enough answer, because we need to use well-done clinical studies. Dr. Park, I want to thank you for mentioning the greatest contribution to medicine, "the randomized clinical trial." We need to have well-done studies to show that a medical practice or procedure is worth doing and to justify a particular course of action.

Let me tell you a little bit about what I do. At Albany Medical College, which is a small medical school, we have a new curriculum. In the last nine years, the curriculum required that all students take a four-year course in "Evidence-Based Medicine." As you have heard, this is unusual. Most schools fail to have required any courses on "Evidence-Based Medicine," let alone a four-year course. The goal of the course is to teach the principles of critical thinking and analysis to medical students who have been accused of only learning a bunch of facts and then simply administering those facts. Well, we must prove that this method is wrong, because physicians can, and do, think critically. Part of this includes interpreting and applying the medical literature in the most accurate way.

The first question one should ask is how the particular evidence will be useful in a courtroom. For example, in one particularly heart-breaking case that I was involved in, a physician's attorney asked me to review the case of a forty-year-old man who was a gainfully employed construction worker that had became disabled due to a stroke, the consequence of an apparent minor neck injury. He was seen in the Emergency Department and admitted to the hospital. Approximately twelve hours after being admitted to the hospital, the patient had a stroke. It was caused by an unusual internal injury to one of the arteries in his neck. The plaintiff felt that, had the doctors used more diligence and care in diagnosing his symptoms, the stroke could have been prevented, thereby preventing the plaintiff's current disability. I was approached by the defendant's attorney, and asked for my medical opinion on

whether I thought this unusual internal injury could have been foreshadowed. Since I had never encountered such a case, I began researching the medical literature to find some standard of care that would support early diagnosis for such a particular injury. However, I was unable to locate such a standard of care in over two hundred cases, dating back to the early 1900's. In my diligent efforts, I discovered that approximately 95 percent of patients with this injury were diagnosed *after* they suffered a stroke.

I disclosed my findings at trial. Although the plaintiff's two expert witnesses had lengthier CV's than me, they both had never seen a similar case that was diagnosed before the appearance of a stroke. The jury concluded that there was no negligence and that the doctor had done the best she could under the circumstances. Even though I am not an attorney, I believe the outcome was fair, since the jury was well-informed. If I had been approached by the plaintiff, I would have gone through the same procedural methods as I had for the defendant and discovered the same end result. Therefore, my conclusion would have been that the plaintiff lacked the necessary requirement to pursue a legal suit against the doctor. Since there was no applicable standard of care, there could be no negligence. All of this objective research and analysis may have prevented a trial altogether. This was not a circumstance of medical negligence. This was a case of bad luck.

What is this "Evidence-Based Medicine" process, hereinafter called "EBM"? It is basically a five-step process also known as the Educational Prescription. It uses this alternate name, because of the reality that physicians are truly life-long learners. So, we tie this process of evaluating the medicinal thinking to the process of being up-to-date, and keeping your skills on the cutting edge.

The first step is to craft a clinical question. The question has to be one that you can answer regarding what the patient's problem is. Is there a way of diagnosing internal injuries to the carotid artery after blunt neck trauma?

The next step, and this is crucial, is to be able to search for the medical literature that is going to answer that question. This is something in medical school that we call "Informatics." The search engine that we use is something called "MEDLINE," which is equivalent to WestLaw, the search engine for those in the legal profession. In Albany Med., this is taught as a separate course. We wanted to teach our students to become good researchers and good users of the available Informatics resources that are out there. The idea is to find all of the studies that may answer this clinical question.

The third step is to select the studies that look the best. This is done by reading the part of the study called "The Abstract," which is when the person writing the article describes the article in a brief summary. This summary tells you what the results are. From this abstract, you get a sense of what the study is going to propose, and how strong the results might be.

Once the best studies have been selected, the fourth step is to critically appraise the studies. In the critical appraisal step, we are looking at two things. The first, and probably the most important, is the validity. This is where issues of junk science come in. A good evidence-based medicine practitioner will be able to look at a study and determine how valid that study is, what the potential sources of bias are in that study, and how those sources of bias may affect the results. For example, a study of injuries among twenty-eight kids falling off a bed, or a study on short vertical falls, that I actually looked at before coming here, that had about 167 kids in it, were both riddled with potential sources of bias. What it essentially shown, not very well, but reasonably well, was that kids who had serious injury were more likely to be victims of child abuse, than were kids who did not have serious injury. I suspect the study that had only twenty-eight kids was probably much weaker in terms of its methodology.

The next step in the EBM process is to assess the strength of the results: how much more likely are kids with serious injury to be victims of child abuse than kids who did not have serious injury? In this case, it was something like 6.5 times, which is a strong result, as opposed to 1.5 or 2 times, which would be a weak result. This requires that physicians understand and be able to interpret the statistics that are involved in the study.

Finally, the last step, and the trickiest one, is to apply the results to an individual patient. Remember, these studies are done on samples of populations: the twenty-eight kids and the 167 kids, and not individual patients. The studies are not done on individual patients to see what is going to happen to that individual person. It is what we expect to find statistically, in a large group of people. What the good practitioner of EBM does, is combine the evidence with their own clinical expertise, and their own experiences with the values of their patients, to determine what the next best step is for the therapy or diagnosis of their patient.

Let me now tell you about another case in which I applied EBM. This was a case of a dead young mother of three. She was a 32-year-old woman who developed abdominal pain a little less than a week after delivering her third child. She was seen in the Emergency Department. Based upon an incorrectly obtained urine sample, the diagnosis of a urinary tract infection was made. She was sent home with some oral antibiotics. Two days later, she died. She died of something called "Child Bed Fever." "Child Bed Fever," as a diagnosis, disappeared in the early 1900's. In the late 1800's it was noted that "Child Bed Fever" was a result of doctors not washing their hands well when going from the autopsy room to the delivery room. This is sort of a relic, as far as the diagnosis goes. It is also called puerperal sepsis, or Group A Beta Hemolytic streptococcal sepsis. Regardless of what it is called, it is not something you want to be diagnosed with.

The defense expert was a world-renowned expert in Critical Care Obstetrics, and as the expert witness for the plaintiff, I was highly intimidated by his impressive credentials. He proceeded to point out that there was a causation flaw. Although he agreed with my deposition, that the urinalysis was incorrectly obtained, and thus, the diagnosis was missed, he believed, based on his own expertise, that the patient was so sick, she would have died regardless of the treatment that had been given. After reading his deposition, I looked at the bibliography from his CV, and found a study he had done on the treatment of women with "Child Bed Fever" and puerperal sepsis. In this particular study, 8 of the 10 women who were newly treated, survived. I thought to myself, maybe my patient had been sicker. Perhaps there was some truth to his contention, that she would not be able to survive, even if she had gotten treated at the time of the Emergency Department visit. I looked over the methods of the study and found that, in order to get into the study, the patient had to exhibit certain signs and symptoms. In this case, the decedent did not even meet the necessary criteria to be included in the study. Essentially, she was not sick enough. As a result, his very own literature contradicted his deposition testimony. It was ironic that this renowned expert would impeach himself, compromise his integrity, and testify in contradiction to research that he had conducted. It was also surprising that an attorney would depend on his testimony without question.

I recognize that it is asking attorneys to know a great deal about science, and I do not know if this attorney relied solely on this one expert's opinion. Many of the attorneys I work for usually have two or three experts review each case. Typically, one expert has good courtroom presence and speaks well, and the second expert simply supports his testimony and analysis of the case. To me, this makes a lot more sense than relying on one witness. It was only because I looked so closely at the evidence and the methodology behind the evidence, that I was really able to evaluate it, and help the plaintiff win the case.

EBM, "Evidence-Based Medicine," is a great leveler. It makes medicine, and the whole process of gathering evidence, much more democratic. Today, there is no more Eminence-Based Medicine. This really represents a paradigm shift. This older paradigm was based on pure sciences, and the idea that, if a therapy probably will not hurt, we ought to try it. Essentially, it is based on the idea of the older physicians telling the younger ones what to do. The new paradigm is best exemplified by the idea that you have to first prove that something is going to work, and that it is beneficial, before you do it. The definition that is given to "Evidence-Based Medicine" is "the

conscientious, explicit and judicious use of current best evidence in making decisions about the care of individual patients."[1]

The Court can use experts to decide which evidence is going to be admitted in two different ways. The first way is to simply just accept any expert, self-proclaimed or otherwise, which could result in a distortion of the data and incorrect conclusions from scientific studies that are of dubious quality. Alternatively, we can employ experts who explicitly use the EBM process, in order to demonstrate how good the science is.

I hope that I have been able to provide you with some information on EBM and how it can be used to obtain the most powerful and useful data for a jury, thus, arriving at a just result.

Thank You.

1. David L. Sackett, William MC Rosenberg, JA Muir Gray, R. Brian Haynes, W. Scott Richardson, *Evidence-Based Medicine: What It Is and What It Isn't* (This article was based on an editorial from the British Medical Journal, January 13, 1996, (BMJ 1996; 312: 72)) http://cebm.jr2.ox.ac.uk/ebmisisnt.html (last visited on 03/11/00)

Medical Evidence in the Context of Child Abuse Litigation

Robert Reece, M.D.*

Paul Kennedy,[1] you are not going to have to keep peace between the doctors and the lawyers here, because part of their common ground is what I am here to talk about. We need to work together. My speech focuses on the qualifications of expert witnesses. It is very important to the medical and judicial professions to come to grips with the need to qualify expert witnesses since there are many issues to be dealt with in the courtrooms. I was in Iowa a couple of months ago when I heard an expert witness who has testified in over forty-six Shaken Baby Syndrome cases for the defense. This expert makes $100,000 a year simply by testifying in these types of cases. This should not happen, however; we cannot do any-

---

* Dr. Reece is a Clinical Professor of Pediatrics at Tufts University School of Medicine and the Director of the Institute for Professional Education at the Massachusetts Society for Prevention of Cruelty to Children (MSPCC). Dr. Reece has worked as a clinician, teacher and researcher in child maltreatment since the early 1970s. He has served on numerous government advisory boards and commissions that address child abuse and neglect. Dr. Reece was the Program Chair for the Child Abuse Section of the American Academy of Pediatrics from 1992 to 1996. He has served on national boards and executive committees for the American Professional Society on the Abuse of Children, Prevent Child Abuse (America), and the National Children's Alliance. Dr. Reece is currently the Chair of the Section on Child Abuse and Neglect for the American Academy of Pediatrics.

1. Lieutenant Paul Kennedy was another speaker at the Symposium. He is a Detective Lieutenant with the Providence, Rhode Island Police Department and is currently Commander of the Youth Services Bureau. Lieutenant Kennedy is responsible for all child abuse investigations and supervises and oversees all investigations involving suspects under the age of 18. He also is responsible for developing and implementing programs to combat juvenile delinquency. Lieutenant Kennedy serves on the Child Protection team at Hasbro Children's Hospital, which investigates all child physical and sexual abuse in Rhode Island. He has been an investigator for half of his law enforcement career and currently focuses on youth crime, child abuse, and prevention/intervention programs designed to assist the youth of the city of Providence.

thing about this particular expert and many others like him. We need to work with the judiciary to prevent these well-paid witnesses from testifying in court. As lawyers and doctors, we need to consider several issues before we go into the courtroom. First, we need to consider the timing of the injury. If you are in a criminal case, you must be able to place the perpetrator. You need the correct perpetrator with the correct victim. If you cannot do that, you may not have a case.

I am going to speak a little bit about timing, with regard to bruises, head injuries, and fractures. Keep in mind that there are certain characteristics that are truly abusive characteristics. It is not all guesswork, since there is a substantial amount of literature about child abuse, and I myself have edited a couple of books on this topic, and many of the experts in the field have contributed. Carole Jenny and Betty Spivack have both written chapters in these books. A couple of years ago, I focused on particular literature in attempts of finding out how many articles have been written about Shaken Baby Syndrome alone. Over 400 articles were published two or three years ago. I am sure that it is up to five, maybe even 600, articles today. However, not all of these are wonderful articles. Only about one hundred of these are very good articles.

I think there are characteristics based on certain lesions that we see in child abuse. The origin of these characteristics can be seen as abuse by virtue of either the location or the pattern. However, this is not always 100 percent accurate. You must look at timing, to determine the age of bruises. We used to think we knew more than we do today. We know that the coloration of bruises does not really tell us what we used to think it did. That is, we used to think that if the bruises were of red or blue coloring, they were two or three days old. If it was a different color than red or blue, it was five to six days old. The reason we know now that it is wrong is because several studies have been done that have looked at the characteristics and the coloration of bruises over periods of time. Using Evidence Based Medicine, people have now taken time to look at bruises and take photographs on a regular basis to monitor changes in the bruises and what causes these changes.

The location of bruises make them difficult to evaluate. For example, if there were bruises around the eye, where there is very loose tissue, this would appear to be a different kind of a bruise from one on the thigh or buttocks. It is important to evaluate the depth of the bruise, the location of the bruise, and what is underlying the bruise. Consider, if there is highly vascular tissue there, or if there is just bone. The bruises will look different depending upon these considerations.

Another consideration is the age of the child. The younger the child is, the more likely an injury is going to be of an abusive nature, especially under one year of age.

Next, there is the issue of skin complexion. Different kids have different skin complexions. They range from extremely white to extremely dark. These considerations affect the characteristics of a bruise.

Our knowledge about the characteristics of bruises has changed as new studies have emerged. With superficial bruises, usually you can see yellowish tinges in about three days. Does that help very much in most of these cases? Not much. Deeper bruises may turn yellow in seven to ten days. Loose skin allows blood to be seen sooner. Blood collection in certain locations may not be appreciated for some time. For example, I have a case that I have to go see today in our hospital where a child has a subgaleal or sub-scapular hematoma, which means bleeding under the scalp. This can be present but go unnoticed for one or two days. Then, suddenly, a mother will be holding or feeding her baby and notice this 'mushy' feeling on the side of the head for the first time, anytime between one to three days after the initial injury. This injury could be inflicted or accidental, depending on the history provided.

There are a couple of other things that need to be said about bruising in children, especially in children under the age of one, or more specifically, under the age of nine months. There was an article published by Ken Feldman in Seattle entitled, "Those Who Don't Cruise Rarely Bruise." The article suggested that children who are not up and around are not going to be bruised since they simply are not doing the things children do to get bruises. Therefore, a six-month-old baby with bruises has to be considered a possible victim of an inflicted injury, since their movement is so limited.

We are also aware that bruises with any yellow color are greater than eighteen hours old. Red, blue and purple can be seen any time from an hour to resolution. This information does not help us very much.

What about head injuries? How many of you think that Louise Woodward did it?[2] Well, let me put it a different way. How many of you think that Louise did not do it? Well, I am here to tell you she did it. That trial was four years ago, and one of the issues in that case was the timing of the subdural hematoma. That same defense expert that I mentioned in Des Moines, Iowa was also the defense expert testifying to essentially the same thing in this case. You can tell what the expert is going to say by looking at where they have testified before and what they said in those

2. Commonwealth v. Woodward, 427 Mass. 659 (1998).

prior testimonies.

The timing was really an issue. When did the incident happen, and who was with the child at that time? We also know that the Shaken Baby Syndrome, or Shaken Impact Syndrome, is another issue that will be challenged in court. Academics can argue about the difference between shaking alone and shaking plus impact. The argument is one that appears in courts often, since it is something that they can fixate on and talk about. Essentially we are talking about acceleration/deceleration injuries. Impact is going to cause a greater deceleration injury. In all of these shaking cases there is deceleration at the end of the arc of the swing. Shaking versus impact is not actually a controversy; it is simply a topic to talk about and a topic that is used to mislead people in the jury box. Timing is the key, but where do we get the information about timing?

Information can be attained from the child's clinical history. Carole Jenny mentioned the "discrepant history," when the history does not match the physical condition in front of you. An example of a discrepant history is a situation of a child falling off a couch. The "killer couch," what Bob Kirschner called it, does not constitute sufficient force for a significant head injury. Falling off of a couch does not typically cause significant injury to the head, with a couple of exceptions such as epidural hematoma. History is the really important consideration. Does the history fit what you see? If it does not, then you must question how such an injury could have occurred.

I have mentioned the Howard Study in the United Kingdom, which compares Caucasian and African children. It is a terrible study not only from a racist standpoint, but also from the standpoint of its methodology. There is a lot literature studies out there about falls in childhood. Some studies about falls in childhood are very good, some are less valid, and some are poor. We look at the good literature and studies to winnow out those things that we need to use in court and be able to identify the mediocre studies. The other side may bring up the mediocre studies if the conclusions fit their defense theory. You should be able to say, "that is a poor study because of a, b, c, and d." Perhaps the statistics are invalid. The conclusion, in the majority of valid literature on falls, is that short falls, usually under four feet to six feet, do not cause life-threatening injuries. The other sources of information are the anatomic and physiologic data that we collect during the clinical examination.

Clinical assessment tells us a lot about these children. I will use the

Matthew Eappen case[3] again, since it is such an illustrative case in so many ways. This is a case about a child who, at 11:30 in the morning, was eating canned peaches or vegetables. Louise Woodward attested to this fact while she was on the stand. Woodward also said that she was talking on the telephone with a friend, who said she could hear Matthew Eappen laughing and gurgling in the background, as Louise Woodward was feeding him. This was 11:30 in the morning, and at 3:00 o'clock in the afternoon, that same child was near death in the Emergency Room at Children's Hospital. There is no way that suddenly the child would collapse and die as a result of a previous injury that occurred two to three weeks before, as claimed by the defense. The proximity of that timing tells us that the fatal injury had to have happened within that three to four hour period.

The other source of information is from accidental trauma cases, for example, kids injured in motor vehicle crashes or being picked up off the pavement after a six-story fall from an apartment building. We look at how quickly their signs and symptoms develop and compare them to a possible abusive head trauma. Based on that information, we can say that the kinds of injuries they receive are very similar and can occur from a particular accident. We know when the accident occurred, and we can monitor the onset of signs and symptoms of those injuries.

An autopsy can often tell us if there is acute or old cranial blood. However, you have to be careful not to listen to everything that the pathologists say. Sometimes, they have biases too. An autopsy also can tell us what other lesions were present. While we may direct our attention to head lesions, there may be other lesions that we do not suspect. Sometimes, they cannot be detected clinically or by an x-ray at the time we look at the child. We use a lot of different kinds of imaging modalities, such as X-Rays and CT scans, and we can give some estimate of timing; yet, it has never been terribly helpful in pinning down the actual time of the event.

How about the timing of fractures? We do know that there are certain stages of healing that fractures go through. Skull fractures, however, do not heal in the same way as long bones since they do not produce calluses. It takes between seven and fourteen days for calluses to form. In this photograph, the arrows show a callus formation in a rib fracture. If you see that, you know the injury is at least seven days old if in an infant, or ten days old, if in an older child. This helps to pinpoint the timing of the injury.

I would like to talk about the position of lesions. The three most impor-

3. Commonwealth v. Woodward, 427 Mass. 659 (1998).

tant things about a lesion are location, location, and location. It is like real estate. Location is important because certain kinds of lesions appear on certain parts of the body due to accidents, whereas other types of lesions appear because of inflicted injuries. This is not an exact science but, by and large, you often see two- to three-year-old children with bruises on their shins or lower arms or on the forehead. For example, if they fall off a big wheel, they can get a big bump on their forehead. The hips, elbows, and other bony prominences are where we see bruises from accidents. Abusive bruises are usually in those areas unlikely to be injured accidentally. It is hard to imagine a way a child could fall and bruise his or her upper arm, although it can happen. Bruises on the interior thighs, tongue, cheeks, buttocks, and genitalia are more suspicious. Most injuries to the male genitalia are intentionally inflicted. Another area where suspicious injuries appear, that usually are inflicted and not accidental, are on the ears and neck. When falling, you fall and hit the top of your head or you bring your shoulder up to protect your ear. In a fall, ears are seldom hurt.

This is a picture of a child who has been slapped. You can see the parallel lines between the fingers or between the slap marks, but very few other things. We have location and patterns here. It is hard to say that it is anything but an inflicted injury. These are upper arm bruises resulting from the child having been held, tightened, or some other abusive act.

Injuries to the neck are very difficult to do accidentally. The neck is not a very accessible spot. The usual excuse for this is that the child has gotten tangled up in a window shade or a Venetian blind shade or something similar. That is where Paul Kennedy's group comes in. In order to rule out abuse, Paul Kennedy's group goes to the house to find out whether there are in fact any Venetian blinds.

These are posterior bruises. Rarely can one think of any type of accident that would produce these kinds of bruises.

There are two kinds of fractures that I want to mention. One is the posterior rib fracture, and the other is what is called "classic, melaphyseal lesions," or CML. By and large, in children under the age of three, posterior rib fractures are almost always inflicted injuries. The reason we have posterior rib fractures is because of compression and rotation of the rib arc beyond the level of the spine. When you see the Kleinman's Abuses series, eighty-eight percent of the kids that do have posterior rib fractures have them in that location because of the levering effect of the transverse process on the spine.

These are bucket-handle fractures. You have probably heard of these. These are the classic melaphyseal lesions. The disc of the bone is frac-

tured by pulling or even by twisting forces. Bucket-handle fractures are seldom seen in any other kind of situation. Other characteristics of fractures pointing to child abuse are multiple fractures and fractures in varying stages of healing. Fractures of the pelvis are hard to acquire unless they are inflicted.

This is a child who was tied in the crib. The child was restrained with strings around the ankles, which resulted in lacerations. These are coat hanger marks. Notice the slap mark on the child's face and lesions on the bottoms of the child's feet. Children do not get lesions on the bottom of their feet or on the palms of their hands accidentally.

We were talking about sources of medical evidence. I want to echo what has been said a couple of times before -- peer review is an interesting process. Peer review entails sending out the article to two, three, or sometimes four people who are allegedly in the field of practice that you are in. I say "allegedly" because a lot of times the article is sent to people who do not know much about child abuse. After a period of review, the article is sent back and either accepted as is, rejected outright, accepted with modification, or can be submitted to another journal.

The problem is that peer review is not always as unbiased as it might seem. There are some journals that we admire more than others, such as the *Journal of Pediatrics* and the *New England Journal*. The LANCET might also be considered in this category. There are eight or ten journals that are really very good. I think that all of us at one point have had articles rejected that we feel should have been included because, often times, we see articles printed that we believe probably should have been rejected. It is not a perfect process, which is what they say about the law, but it is better than any other. That is why peer review of these articles is so important. I think you need someone to consult with if you are dealing with a child abuse case, or any medical literature in a case for that matter. You need someone that you can go to and ask, "Will you help me with this?" Otherwise, you risk the potential of getting into trouble in court.

Thank you.

Judicial Perspective on the Presentation of Medical Evidence

The Honorable Peter Lauriat*

Just a couple of quick comments. First, I would like to address some of the issues that Lynn Rooney[1] raised, which I certainly echo. Early in my career as a Judge, I learned that Medical Examiners, as expert witnesses, absolutely love to show slides, photographs, and bodies to the jury. I think the reason is that the jury is the first live audience they have ever had. They only deal with dead bodies. It is remarkably difficult to prevent a medical examiner from showing slides of a mutilated face or a slide with the scalp pulled down over the face so that you can see the back of the brain or the back of the head. An attorney has to monitor his medical examiner. We, as judges, know only enough about medicine from medical malpractice cases to be eminently dangerous. We attempt to diagnose our family members every day. We usually get it wrong. But it is a lot of fun.

I would like to address this issue from the perspective of the trier of fact. How does this affect the trier of fact? How is it presented? And what does the trier of fact do with all of this? I say this when cases are tried without juries from time to time. However, the real triers of fact are jurors. These are people from every walk of life. An attorney will have to

---

* Judge Lauriat has been a judge of the Massachusetts Superior Court in Boston, Massachusetts since 1989. Prior to appointment to the bench, he practiced law in Boston for 17 years, specializing in civil and criminal litigation. Judge Lauriat is a 1971 graduate of University of Chicago Law School. Judge Lauriat has taught at Harvard University Law School, Stanford Law School, and the National Judicial College. He has written and lectured nationally on litigation and jury trial issues and techniques. Judge Lauriat is an editor and contributor to several publications, including the *Massachusetts Jury Trial Benchbook, Jury Trial Innovations in Massachusetts*, and the *Massachusetts Deposition Practice Manual*. He is also co-author of the Massachusetts Practice Series Volume 49: *Discovery*. Judge Lauriat is a member of the Massachusetts Supreme Judicial Court's Jury Management Advisory Committee.

1. Lynn Rooney was another speaker at the Symposium. Ms. Rooney is the Deputy First Assistant in the Middlesex District Attorney's Office. She has presented at numerous conferences on the issues of child abuse and domestic violence, including the National Conference on Shaken Baby Syndrome and the Massachusetts Guardian Ad Litem Association.

find twelve jurors and convince them of his or her particular side of the case.

In Massachusetts, we have a Confidential Jury Questionnaire that we give to each juror to bring to court on the day that that juror has been summoned and will be considered for jury duty. On the form there is a box in which the potential juror is to provide us with any information that may be relevant to his or her ability to be an impartial juror. Here are some examples of the responses we have received. In 1991, a juror wrote: "I am prejudiced against repeat criminal offenders, able-bodied Welfare recipients, ambulance-chasing lawyers, Judges that don't work full days, career politicians and others whose actions reflect their warped views of ethics and individual responsibilities." The next juror's response reflects a point that David Siegel[2] made. This juror wrote: "I believe that justice is determined by how good a lawyer is. Not by guilt or innocence. I find most lawyers distasteful, and their antics would sway my opinion more than the evidence. I resent my tax dollars going to fund Public Defenders for career criminals." The third juror simply wrote: "I am psychic." These come into the court everyday, and I love reading them. Now you know what to say if you want to get out of jury duty. It is time to turn our attention to why we are actually here.

The goal in presenting medical evidence, or any evidence, in the courtroom is to persuade the trier of fact, in most cases, the jury. The attorney has to persuade them that the evidence being presented is credible and worthy of belief. In light of these facts, all of the evidence persuades the jury to decide the case in your favor. It is easy to say, but not easy to do.

Medical evidence, which we probably see most frequently, often means using experts, or presenting experts. It means finding an expert who is outstanding in his or her field. It is usually someone who is well-credentialed, a teacher, an explainer, articulate, friendly, outgoing, and comfortable with an audience. You should have a witness who relates easily to others, and does not become upset when challenged. Now, if anyone here says that he or she has ever seen such an animal, not a zebra, but an expert, I would suggest that you are not worthy of belief. At least in my twelve years of experience as a Judge, like the unicorn, that expert probably does not exist.

We have to deal with what real life is all about, and what real experts are all about. They are critically important in many kinds of cases. I will give you one example of a child abuse case that I tried a few years ago in which the victim was old enough to testify. She was probably about nine,

---

2. Professor David Siegel was a moderator at the Symposium. He teaches Criminal Law, Criminal Procedure, Criminal Advocacy, Clinical Evidence, Evidence and Mental Health Issues in the Criminal Process at New England School of Law.

ten or eleven years old. The accusation against her father was sexual abuse and child rape. The problem was that for a period of nine months after he was out of the house, the child did not mention to anybody the fact that she had been raped, or sexually assaulted by her father. Eventually, an event triggered her memory, and she told her mother what had happened nine months earlier. It was prosecuted by the Middlesex County District Attorney's Office. They knew that it would be a large question mark in the eyes of the jury as to why this young lady, with the abuser gone from the house, for nine months, simply did not say anything to anybody about this incident. The District Attorney's Office brought to the witness stand, a child psychologist, who, for the first time, for this jury, explained post-traumatic stress disorder or Rape-Trauma Syndrome. The witness testified to the factors that could trigger post-traumatic stress disorder. However, the expert was not allowed to relate the factors to that specific child. The father absolutely denied it ever happened. The jury found the father guilty of sexual assault. I am convinced, by my own observations, and in talking with the jury, it was the expert whose testimony was believable, credible, made sense, and explained the phenomenon to the jury. Juries do not like being told by experts what they already know or believe they already know. They feel that they have some competence in thought. However, you need experts for those fields where the jury really is not in a position to make a reasoned judgment based on knowledge, training, or experience. That is where the experts will be most helpful.

I would like to now focus on preparing the expert for the jury. You have to prepare your expert, not only for cross-examination, but for direct examination, to be sure that the expert witness can speak English to the jury and present the case to the jury in a way that is not only understandable but persuasive. As Lynn Rooney has suggested, the use of blow-ups, diagrams, chalks, charts, photographs, anatomical models, slides, videos and sometimes Power Point actually is wonderful. Now, one thing that I have learned over the years in a courtroom is that whenever you try to use a mechanical device it will never work. You should be prepared for it not to work when you plug it in. You should ideally have your witness, especially your expert witness, take the stand after a break. This allows you some time during the break to determine if all of the equipment is working. You should understand, for example, in Suffolk County, Boston, we operated some courtrooms on direct current until a few years ago. If you had any kind of electrical device that you wanted to plug in and use during expert testimony, it had to be plugged in with a 300 foot extension cord. It ran from our courtroom, out the window, down three floors, into the basement, and into a plug that was actually alternating current, like the kind that we all have. About once every two weeks, someone downstairs

accidentally kicked the plug out, and you would have to call a recess to plug the cord back in.

So they are experts, and they tend to use it more than anyone else. They actually make the kind of presentations that would be useful for the jury. In addition, if you can make copies of your exhibits, and once they are in evidence, or if they are charts, you get the court's approval, give a copy to each juror to be kept in a notebook by that juror. You can provide the notebook at the start of a trial, in which you can put the key exhibits as they are admitted into evidence, or you can put pre-marked exhibits for a civil case or a criminal case in the notebook. Prior to the start of trial, each juror should be given the notebook with appropriate judicial instructions. Depending on the nature of your case, the notebook can also contain a glossary of terms that are going to be used by your experts. Usually you can get opposing counsel to agree to a glossary of terms that will be used by both parties.

It is also helpful to have a list of the testifying witnesses in the notebook. You can put their curriculum vitaes in the notebook, providing they are a reasonable size. Curriculum vitaes ordinarily will go into evidence, if both sides agree. Either both sides agree, or nothing is admitted into evidence. However, it does decrease the length of time that you need to qualify the witness. It is still to your benefit to qualify the witness. You want the jury to understand what the witness is all about. It is important for the jurors to have the witness' curriculum vitae and other accomplishments in a notebook for the duration of the trial.

Also include a section for note taking since judges are increasingly allowing jurors to take notes. Judges have allowed it for twenty years now, although, it may actually be closer to ten years in the Superior Court. Most judges allow jurors to take notes. It is important that they be allowed to take notes and remain engaged throughout the trial of the case. As I think Lynn Rooney said earlier, you do not want the jury to have a glazed look or fall asleep during the trial. You want to keep them engaged in the cases for as long as possible and, hopefully, throughout the trial of the case, and this is one way to do it. You put something in their laps and give them a pencil or a pen to write with. You allow them to take notes. They can take notes on the very exhibits that you now have in evidence, or you can give them each a copy. As witnesses are testifying about particular evidence, you can have a juror refer to some tab or a particular document. They can make their own notes on the document. After the case is over, I talk to the juries to find out if it was helpful to have a notebook to take notes. I also ask if it was helpful to have copies of the evidence in the notebook and, most of the time, the answer is yes. The notebook is critically important for the jury.

A second approach is to have one or two binders with exhibits. This is

useful in an abuse case with a lot of medical records, as all the records will usually be put into evidence. I have urged lawyers, and sometimes ordered lawyers, at the start of a trial, when they know what exhibits and medical records are going into evidence, to prepare a notebook of at least one or two thick volumes. The first thing they should do is to get the volumes into evidence. As soon as they are into evidence, another set of those volumes is sent off to the jury room. I tell the jury that they are going to be in the jury room for the balance of the trial. The jurors are told that they can view the notebooks during any break, or any point in the morning, when they first come in, and before the trial begins, as well as after trial ends for the day. I have found this to be very important for the jurors and their comfort level when taking on the awesome responsibility of deciding someone's fate or deciding the serious issues between plaintiffs and defendants in civil cases. The jurors are hungry for information. The jurors want to know that they made their decision on the basis of the greatest possible amount of information and had the chance to read and digest it. Think of this as a technique for allowing jurors to understand what your expert is saying.

You can also include excerpts from learned treatises, if the parties agree that it is a standard text in the field. It has to be a subject that the experts from both sides are going to testify about. It is important that you alert your experts that you are going to allow the jurors to ask the witness questions. This was approved by the Massachusetts Supreme Judicial Court,[3] in a case decided this past spring. I recommend it to you because the Supreme Judicial Court agreed with what a number of us have been doing for a number of years. Each witness who testifies in a case, civil or criminal, is being told, that at the end of your testimony, before you step down from the witness stand, we are going to allow the jurors to write out any questions they may have of you. The jurors are going to pass those questions to the judge who will share them with the lawyers at side bar. The judge will rule, based on the Rules of Evidence, whether those questions from the jurors are admissible. The judge tells the jurors that they are not expected to know the Rules of Evidence but their questions will be governed by them, as are the lawyers' questions. Then, with respect to those questions the judge decides are in compliance with the Rules of Evidence, the judge is going to read the questions to the witness. The questions can be directed to an expert or a lay witness. The judge asks the witness to answer the question as best as he can, and if he cannot, simply say that he cannot answer the question.

When the witness has answered all of the questions that the jury has

3. *See* Commonwealth v. Britto, 433 Mass. 596 (2001).

asked, I will briefly allow the lawyers from both parties to follow up on those questions in the area that is covered by any of the questions. This procedure has a really terrific effect on the trial of the case. The questions keep the jury absolutely engaged in the trial of the case. Allowing questions also alleviates a lot of frustrations that jurors have when they feel that there are unclear areas. The questions help when a juror may have not been paying attention to a critical part of the case. They sometimes want to know things that the lawyers may not have thought about, but which, in retrospect, or upon reflection, are important aspects of a witness's testimony. Jurors now have a chance to at least present the questions to the court, and most often to the witnesses themselves.

I have found, in my four to five years of experience with this technique, that experts are the ones who get most of the questions. It is a little daunting for the first party, for example the Commonwealth, or the plaintiff in a civil case, because their experts first take the heat. There are a lot of questions, and some can be fairly pointed. The expert may feel that he or she is being picked on, or that the jury has already decided how they will rule based on the flavor of the questions. I tell them to please be patient. You will find that the jury will ask every expert the same kind of penetrating questions because they are hungry to get to the bottom of this. They know their responsibility. It is to decide this case, and they want to do it with as much information as possible. They want to be as clear as they can about the information they are receiving from the experts and be able to satisfy themselves that it is appropriate -- that they make this decision on this basis. As a result, it is becoming a common and very useful technique.

I am using it now in a civil case that I am trying at the moment. By yesterday, we were up to about 142 questions from the jury. We are in the tenth day of a six day trial, and it is still not finished; not because of the jury, but because of the lawyers. However, the jury is involved. It is a very boring case, so I will not bore you with the details. Since the jurors are involved and engaged, they are committed to the trial in this case. They are committed to seeing it through. They are committed to coming out with what they think is the right answer.

Again, in my experience, juries rule appropriately ninety-nine percent of the time. They can see through and understand the experts. They can apply their own life experiences, common sense and good judgment to what experts say, as well as to what witnesses say. In all the years that I have been a judge, I think that I have taken only two cases away from the jury. I have agreed with a jury decision in every other case, at least when motions have been made to take the case away from a jury or enter a judgment notwithstanding the verdict of a jury. I am completely sold on the jury system. The jury system is a great opportunity for witnesses and

lawyers to use their persuasive powers and get a real sense of what justice is and how justice is done. Those are the things on which I would like to concentrate.

I would like to make two additional comments with respect to expert witnesses. In Massachusetts, for years we have relied upon the hypothetical question as the vehicle for getting an expert witness to give his or her opinion about a particular issue. If you and your expert know how to do it, and if you can remember all the facts for the record that the expert must assume before giving his or her opinion, hypothetical questions are probably a pretty good technique for securing that expert's opinion. If you are like most of us, however, you will not remember all the facts. That will mean more time at side bar with the judge and the opposing counsel objecting to whatever it is that you are about to ask the expert because you left out or added some crucial fact.

In Massachusetts, that is no longer required. Hypothetical questions have not yet gone the way of the horse and buggy, but they are not required.[4] You can now do what has been done for years in Federal Court. You can ask the expert witness what he or she did after being engaged in the particular assignment. If, based on that assignment, they form an opinion or come to a conclusion, you may ask the witness what that opinion or conclusion is, and upon what it is based. It makes more sense, and it is more understandable for the jury. I found, in the end, it is quicker and more beneficial for all concerned.

The last comment about experts, for those who will be practicing in Massachusetts, is that it is no longer appropriate for a lawyer to ask a judge to qualify a witness as an expert witness. You may have seen or have been taught that after the witness's qualification as an expert has been presented to the jury you turn to the judge and say, "Your Honor, I ask that you qualify the witness as an expert in the field of child abuse." The answer I will give you is no. That is clearly stunning to some of these lawyers who do not know what to do when the judge says no.

The Supreme Judicial Court in *Commonwealth v. Richardson*,[5] essentially held that it is not appropriate to have a judge qualify an individual as an expert. It is for the jury to decide whether the individual is an expert. It is a better practice not to have the judge put his or her signature on the witness as an expert witness. It is better to simply leave it to the jury. Therefore, if you are practicing in front of me, please do not ask me to qualify a witness as an expert. Simply go through the qualifications and move right to your presentation. Make your presentation as straight forward as possible so that the witness will be persuasive.

---

4. *See* Department of Youth Services v. A Juvenile, 398 Mass. 516 (1986).
5. 423 Mass. 180 (1996).

Finally, in Massachusetts, lawyers are not allowed to talk to jurors after a case is over. Likewise, expert and lay witnesses are not allowed to talk to jurors after a case is over. Only judges are allowed to talk to jurors after a case is over. We guard that prerogative jealously. I think lawyers should be allowed to talk to jurors after a case is over. If you try a case before me, and you and opposing counsel agree, I will let you both talk to the jurors after a case is over, in my presence in the courtroom. In that way, you learn what the jury is thinking about the kinds of things, such as, presentation, which witnesses you called or did not call,what was important to the jurors and what was important to the lawyers.

Let me paraphrase some of the things that jurors have said to me after the case. They often assume that experts are hired guns and, of course, are going to testify for the side that calls them. There is already a healthy skepticism building in the system based upon the fact that one side or the other calls an expert witness. Cross-examination almost invariably involves inquiries about the fee paid to the expert. If they can get away with it, opposing counsel will ask, "You were paid for your testimony were you not?" The quick expert will say, "No, I was paid for my time." It could be the next expert who will say, "Yes, I was paid for my testimony," and you are seeing dollar signs hovering above the witness's head. That kind of thing is standard cross-examination of expert witnesses on bias. The jury understands that experts probably are always hired guns. So, the jurors go into the courtroom with a degree of healthy skepticism.

The next thing the jurors say is that experts are impossible to understand. It is becoming less true now because jurors are allowed to ask questions. It also helps that experts are being better prepared and are able to use equipment and techniques that assist them in explaining to the jury what is going on. However, jurors have often simply shaken their heads and said that they do not understand the expert. They say that they understand it the way that lawyers would like them to.

The last thing they say is that they did not need an expert to tell them this. They knew it all along, and they are fairly bright and capable people. Jurors feel they do not need an expert to tell them about identification or fingerprinting issues. They understand those issues. This in part may relate to Massachusetts having this one-day, one-trial system where there are no exemptions from jury duty. Those of you who live or reside in Massachusetts will be called as jurors. If you are over eighteen and under age seventy in Massachusetts, you will be called to serve as a juror in the Commonwealth. I have been called three times even though I'm a judge. I have not yet been in the jury box, but I have come close once. I did a survey last year at a judicial conference. Out of eighty judges, about forty-four have been called to jury service in the Superior Court, and fourteen have served. We had one judge who was sworn in on a Friday,

and on a Monday, he was called for jury duty and sat for three weeks in a first-degree murder case. It was his first experience as a judge, or shall I say as a juror. After the one-day, one-trial system was put into place in 1986, one of the first judges called as a juror was a District Court Judge named Arthur Sherman, who presided in Cambridge. After his experience as a juror, he had two comments that he would tell anybody and everybody who would listen. One is that the other jurors paid no attention to the fact that he was a judge. The second is that he did not understand the judge's jury instructions. The point of the one-day, one-trial system, which is now in effect in many other States, is that you get a terrific, well-educated, very bright and capable jury. You cannot start with the assumption that the jury knows nothing and needs to be educated. Finding that happy medium is difficult to do. It is well worth the effort, and you will find that you can effectively persuade the jury.

Thank you very much.

COMMENT:

Stenberg v. Carhart: Have the States Lost Their Power to Regulate Abortion?

I. INTRODUCTION

Abortion is an issue that this country has struggled with for more than 150 years.[1] The decision of whether or not to permit abortions was left up to the states until 1973, when the Supreme Court guaranteed the right to abortion, in its landmark decision of *Roe v. Wade*.[2] The Court did, however, return some power to regulate abortion to the states when it decided *Planned Parenthood of Southeastern Pennsylvania v. Casey*.[3]

Recently, however, the United State Supreme Court decided *Stenberg v. Carhart*,[4] which took away some of the power returned to the states by *Casey*.[5] The Court did so by striking down a Nebraska Statute banning partial-birth abortions.[6] As the dissenting opinions of *Stenberg* point out, the majority opinion is not just,[7] and it does not accurately follow *Casey*,

---

1. *See infra* note 14.

2. 410 U.S. 113 (1973) (holding that abortion is a right protected by the Constitution). *See also infra* note 14.

3. 505 U.S. 833, 875-76 (1992) (holding that a state may regulate pre-viability abortions, if the regulation will further the State's interest in fetal life).

4. 530 U.S. 914 (2000).

5. *See id.*

6. *See* NEB. REV. STAT. §§ 28-326, 28-328 (Supp. 2000). Partial birth abortion "is a late term abortion involving partial delivery of the baby before it is aborted by the attending physician." Allison D. Gough, Note, *Banning Partial-Birth Abortion: Drafting a Constitutionally Acceptable Statute*, 24 U. DAYTON L. REV. 187, 188 (1998); *see also infra* section III.A for a more detailed discussion on partial birth abortions.

7. It is not a just opinion because thirty states have statutes banning partial birth abortions. *See infra* note 57. This shows that a majority of people in this country and the legislatures of those states, find the partial birth abortion method unacceptable. *See* Karen E. Walther, *Partial-Birth Abortion: Should Moral Judgement Prevail Over Medical Judgement?*, 31 LOY. U. CHI. L.J. 693, 707 (2000); *see also infra* note 57 and accompanying text; *see also Stenberg*, 530 U.S. at 953 (Scalia, J., dissenting).

as it claims to.[8] The statute at issue should have been upheld as it promotes the state's interest in fetal life, and does not create an undue burden on the right to abortion.[9] Further, although the statute does not include an exception to protect the health of the mother, there is no evidence that a partial-birth abortion will ever be necessary to protect the health of the mother.[10]

This comment first discusses the law on abortion before *Stenberg* focusing on the *Roe v. Wade* and *Planned Parenthood of Southeastern Pennsylvania v. Casey* decisions. The next section describes the partial-birth abortion procedure and the dilation and evacuation abortion procedure[11] as well as state statutes banning partial-birth abortions.[12] The following section investigates two Federal Circuit holdings which examine specific statutes banning partial-birth abortions.[13] Finally, this comment analyzes the Supreme Court opinions in *Stenberg* proposing that the dissenting opinions, and not the majority, are more just in that they follow the will of the "people," and more accurately follow the precedent set by *Casey*.

II. THE STATE OF THE LAW ON ABORTION BEFORE *STENBERG V. CARHART*

A. *Roe v. Wade*

Roe v. Wade was the first decision by the United States Supreme Court

---

Further, it is unjust for five *appointed* Justices to make a decision that strikes down a statute that a majority of the people support, especially when the statute is not unconstitutional. *See id.* at 955 (Scalia, J., dissenting).

8. *See Stenberg*, 530 U.S. at 982-83 (Thomas, J., dissenting). Under *Casey*, the statute should be upheld as long as it furthers a legitimate state interest, provides an exception to protect the life and health of the mother and does not create an undue burden on the mother. *See* Planned Parenthood of S.E. Pa. v. Casey, 505 U.S. 878-79 (2000); *see also infra* section II.B.1.

9. *See Stenberg*, 530 U.S. at 982 (Thomas, J., dissenting).

10. *See id.* at 966 (Kennedy, J., dissenting); *see also* Maria Parker, *Partial-Birth Abortions Are Often Elective*, BOSTON GLOBE, Jan. 16, 2001, at A10 (stating "partial birth abortion is never necessary to preserve a woman's life, health, or future fertility, even in cases of fetal abnormalities.").

11. The dilation and evacuation method is contrasted with the partial birth method to show their differences. The Court struck down the Nebraska statute in *Stenberg* in part because they found the statute's description of "partial birth abortion" was so vague as to include the dilation and evacuation method. *See Stenberg*, 530 U.S. at 938-39; *see also infra* sections III.A. and V.A & B.

12. For a list of statutes banning partial birth abortions, *see infra* note 57.

13. *See* Women's Med. Prof. Corp. v. Voinovich, 130 F.3d 187 (6th Cir. 1997); Hope Clinic v. Ryan, 195 F.3d 857 (7th Cir. 1999); *see also infra* sections IV.A & B.

guaranteeing a fundamental and constitutional right to abortion.[14] The Court based this right on the "Fourteenth Amendment's concept of personal liberty and restrictions upon state action."[15] According to the Court, denying this right could be detrimental to women's physical and mental health, as well as to their life in general.[16] However, the right to an abortion is not absolute, it must be balanced with "important state interests in regulation."[17] To balance these interests the court developed a trimester plan.[18]

The *Roe* Court held that during the first trimester of pregnancy,[19] the state has no interest in regulating abortion; the decision is left entirely to the woman and her doctor.[20] The state may regulate abortions in the second trimester of pregnancy if the regulations are "reasonably related to maternal health."[21] During the third trimester of pregnancy, the state has a compelling interest in the "potentiality of human life," and may regulate or ban abortions, unless an abortion is necessary to protect "the life or health of the mother."[22] This trimester framework remained the law until 1992, when the Court decided *Planned Parenthood of Southeastern Pennsylvania v. Casey*.[23]

---

14. A majority of states enacted anti-abortions statutes in the mid-nineteenth century, which over time became more strict and severe in penalties. *See Roe*, 410 U.S. at 139. By the 1950's abortion was almost completely outlawed throughout the country. *See id.* However, by the 1960's "a trend toward liberalization of abortion statutes" developed, and one-third of the states changed their laws, making the penalties less severe. *Id.* at 140. The liberalization of the statutes in the 1960's created "vigorous opposing views," which pushed the Court to pass judgement on the issue. *Id.* at 116.

15. *Id.* at 153.

16. *See id.*

17. *Id.* at 154 (holding that states have "important interests in safeguarding health, in maintaining medical standards, and in protecting potential life"). The Court uses the term "potential life" because an unborn fetus is not considered a "'person' as used in the Fourteenth Amendment." *Id.* at 158.

18. *See Roe*, 410 U.S. at 164-65.

19. The first trimester is the first three months of pregnancy. *See id.* at 162-64.

20. *See Roe*, 410 U.S. at 163-64.

21. *Id.* at 164.

22. *Id.* at 164-65. The state may ban third trimester abortions because after this time the fetus becomes viable and can survive outside the womb, therefore the state has a compelling interest in protecting the potential life of the child. *See id.* A viable fetus is one that can survive outside the womb. *See* Planned Parenthood of S.E. Pa. v. Casey, 505 U.S. 833, 870 (1992). Viability is estimated to occur at twenty gestational weeks. *See* Walther, *supra* note 7, at 696 n.30 (citing STEDMAN'S MEDICAL DICTIONARY 1936 (26th ed. 1995)).

23. 505 U.S. 833 (1992).

B. Planned Parenthood of Southeastern Pennsylvania v. Casey

1. The Majority Opinion

After nineteen years of following their decision in *Roe v. Wade*,[24] the Court modified its holding when it decided *Planned Parenthood of Southeastern Pennsylvania v. Casey*.[25] *Casey* concerned the Pennsylvania Abortion Control Act ("the Act"),[26] which required certain prerequisites before a woman could obtain an abortion.[27] As a prerequisite, the Act required "voluntary and informed" consent before a woman could receive an abortion.[28] For a woman to voluntarily give informed consent, her doctor must have informed her, at least twenty-four hours before the abortion procedure, of any "risks and alternatives" to the abortion, of the "gestational age" of the fetus, and of any "risks associated with carrying her child to term."[29] The Act also provided that for a minor to get an abortion, she must get consent of one of her parents or a judicial bypass.[30] Under the Act, a married woman had to provide a signed statement that she informed her husband of her intent to get an abortion.[31] The Act, however, provided an exception to these requirements in the case of a medical emergency.[32]

The majority of the Court in *Casey* upheld all the provisions of the Act, except the section requiring a married woman to notify her husband of the abortion.[33] In doing this, the Court modified its holding in *Roe* regarding the conflicting interests of a woman's right to abortion, and the state's interest in regulating abortion.[34] Specifically, the Court in *Casey* abandoned

---

24. *See* 410 U.S. 113 (1973). After the Court decided *Roe v. Wade*, it affirmed its decision in the following cases: Carey v. Population Servs. Int'l, 431 U.S. 678, 684 (1977); Moore v. E. Cleveland, 431 U.S. 494, 500-01 (1977); Akron v. Akron Ctr. for Reprod. Health, Inc., 462 U.S. 416, 420 (1983); Thornburgh v. Am. Coll. of Obstetricians and Gynecologists, 476 U.S. 747, 759 (1986); Webster v. Reprod. Health Servs., 492 U.S. 490, 521 (1989); Cruzan v. Dir., Mo. Dept. of Health, 497 U.S. 261 (1990).

25. 505 U.S. 833 (1992).

26. Pennsylvania Abortion Control Act of 1982, 18 PA. CONS. STAT. §§ 3203 (2000).

27. *See id.*

28. 18 PA. CONS. STAT. § 3205(a) (1990).

29. *Id.* § 3205(a)(1).

30. *See id.* § 3206.

31. *See id.* § 3209.

32. *See id.* §§ 3202, 3205(a), 3206(a), 3209(c).

33. *See* Planned Parenthood of S.E. Pa. v. Casey, 505 U.S. 833 (1992). The majority of the Court in *Casey*, consisted of J.J. O'Connor, Kennedy, Souter, Blackmun, and Stevens. *Id.*

34. *See Casey*, 505 U.S. at 876.

the trimester approach and replaced it with the "undue burden" standard.[35]

The Court chose the undue burden standard because it no longer thought the trimester approach adequately represented the state's "interest in protecting fetal life or potential life."[36] Specifically, the trimester approach in *Roe* did not allow any state regulations prior to viability unless they were to protect the health of the mother.[37] The undue burden standard of *Casey*, however, allows a state to make regulations before viability that are "designed to further the State's interest in fetal life."[38] Following this standard, the Court found that the informed consent requirement, as well as the minor's parental consent or judicial waiver requirement, did not pose an undue burden on the woman's right to choose an abortion.[39]

The majority found the portion of the Act that required a married woman to notify her husband of the abortion, was an undue burden.[40] They found the spousal notification requirement was an undue burden because some women might not be able to obtain an abortion at all.[41] This requirement does more than "make abortions a little more difficult or expensive to obtain," for many women it makes abortion impossible, and that "will impose a substantial obstacle," and therefore and undue burden.[42] Accordingly, the Court struck down the spousal notification portion of the

---

35. *See id.* A regulation is an undue burden if it does more than "create a structural mechanism by which the State . . . may express profound respect for the life of the unborn" and it creates "a substantial obstacle to the woman's exercise of the right to choose." *Id.* at 877.

36. *Id.* at 876.

37. *See supra* notes 18-22 and accompanying text. When the Court decided *Roe*, there was a concern for women's health, because most abortions could only be obtained illegally, and those methods were "unsafe and unsanitary," which often resulted in infections and even death. Walther, *supra* note 7, at 701, (citing ACLU, *The Right to Choose at 25: Looking Back and Ahead*, *at* http://www.aclu.org/issues/reproduct/rrujan98.html (last visited Feb. 7, 2002)).

38. *Casey*, 505 U.S. at 877.

39. *Id.* at 882, 885, 889 (finding that "a 24-hour delay does not create any appreciable health risk. In theory, at least, the waiting period is a reasonable measure to implement the State's interest in protecting the life of the unborn, a measure that does not amount to an undue burden").

40. *See id.* at 887-98. The Court found that women who are in healthy marriages would consult their husbands when making this type of a decision. *See id.* at 892-93. However, those women who are in both physically and mentally abusive marriages would have a good reason for not wanting to tell their husbands of their decision. *See id.* at 893. They might be afraid for their physical and mental safety, and the Act does not provide an exception for these circumstances. *See id.* at 833, 893. *See also* 18 PA. CONS. STAT. § 3209(b).

41. *See supra* note 40.

42. *See Casey*, 505 U.S. at 893-94.

Act.[43] The dissenting opinions, however, propose a different view; they would have upheld all portions of the Act, and overruled *Roe*.[44]

2. The Dissenting Opinions in *Casey*

Chief Justice Rehnquist wrote that *Roe v. Wade* was decided incorrectly, and should be overruled.[45] According to Rehnquist, *Roe* went too far in declaring abortion a fundamental right, and the "undue burden" standard of *Casey*[46] "will not, . . . result in the sort of 'simple limitation,' easily applied."[47] Rehnquist further states that the "undue burden" standard is more subjective than the *Roe* trimester approach, and allows judges to make opinion-based decisions.[48] Rehnquist would sooner apply the rational basis standard.[49] Using this standard, he would uphold only the portions of the Act that are "rationally related to a legitimate state interest."[50]

Justice Scalia's dissent would also uphold all portions of the Act, but for different reasons than those offered by Chief Justice Rehnquist.[51] Scalia feels there is no constitutionally guaranteed right to abortion, and that states can prohibit abortion, if their citizens vote against it.[52] Scalia goes on to write that the state does not violate the Constitution when it makes a "choice between two positions on which reasonable people can disagree [even if] it intrudes upon a liberty."[53] Under Scalia's reasoning, even if abortion is a liberty guaranteed by the Constitution, the state may make a law that intrudes upon it because reasonable people disagree about it.[54] However, abortion, according to Scalia, is not guaranteed by the Constitution for two reasons: "(1) the Constitution says absolutely nothing about it, and (2) the longstanding traditions of American society have permitted it

---

43. *See id.* at 898.

44. *See id.* at 944-1002. There are two dissenting opinions in *Casey*. *See id.* The first dissent is written by C.J. Rehnquist, and joined by J.J. White, Scalia, and Thomas. *See id.* at 944. The second dissent is written by J. Scalia, and joined by C.J. Rehnquist, J.J. White, and Thomas. *See id.* at 979.

45. *See Casey*, 505 U.S. at 944 (stating "[w]e believe that *Roe* was wrongly decided, and that it can and should be overruled consistently with our traditional approach to stare decisis in constitutional cases.")

46. *See id.* at 964-65.

47. *Id.*

48. *See id.* at 965.

49. A state regulation will pass the rational basis test if it is "rationally related to a legitimate state interest." *Id.* at 966, *citing* Williamson v. Lee Optical of Oklahoma, Inc., 348 U.S. 483, 491 (1955).

50. *Id.*

51. *See Casey*, 505 U.S. at 979-1002.

52. *See id.* at 979.

53. *Id.* at 980.

54. *See id.*

to be legally proscribed."[55] Following this analysis, Justice Scalia, would uphold all the provisions of the Act.

III. PARTIAL-BIRTH ABORTIONS AND STATUTES BANNING THEM

A. An Explanation of Partial-Birth Abortions

Not long after the *Casey* decision, some states enacted anti-partial-birth abortion statutes.[56] There are currently thirty states, which have statutes on the books banning partial-birth abortions.[57] In 1995, 1997 and 1999, the Federal Government tried to pass legislation banning partial-birth abortions, but President Clinton vetoed the first two bills, and *Stenberg v. Carhart* rendered the third invalid.[58] However, before the statutes are addressed, a thorough description will be provided of the most commonly used abortion method, dilation and evacuation (D&E), as well as a thorough description of "partial-birth" abortion, clinically known as dilation and extraction (D&X), as they both are related to the anti-partial-birth abortion statutes.[59]

---

55. *Id.*

56. Partial birth abortion is the common name for dilation and extraction (D&X), which is defined as "deliberate dilation of the cervix . . . instrumental conversion of the fetus to a footling breech; breech extraction of the body excepting the head; and partial evacuation of the intracranial contents of a living fetus to effect vaginal delivery of a dead but otherwise intact fetus." Carolyn Bower, *Validity, Construction, and Application of Statutory Restrictions on Partial Birth Abortions*, 76 A.L.R. 5th 637, 646 n.1 (2000) (citing Carhart v. Stenberg, 11 F. Supp. 2d 1099 (D. Neb. 1998), *judgment aff'd*, 192 F.3d 1142 (8th Cir. 1999), *cert. granted in part*, 2000 WL 21145 (U.S. 2000)).

57. *See* ALA. CODE § 26-23-3 (1997); ALASKA STAT. § 18.16.050 (Michie 2000); ARIZ. REV. STAT. § 13-3603.01 (West Supp. 2001); ARK. CODE ANN. § 5-61-202, 203 (Michie 1997); FLA. STAT. ANN. § 390.011 (West 2001); GA. CODE ANN. § 16-12-144 (1999); 720 ILL. COMP. STAT. ANN. § 513/10 (West 1999); IND. CODE ANN. § 16-34-2-1 (Michie Supp. 2001); IOWA CODE ANN. § 707.8A (West Supp. 2001); KAN. STAT. ANN. § 65-6721 (Supp. 1999); KY. REV. STAT. ANN. § 311.720 (Banks-Baldwin 2001); LA. REV. STAT. ANN. § 14:32.9 (West Supp. 2001); MICH. COMP. LAWS ANN. §§ 333.17016, 333.17516 (West Supp. 1999); MISS. CODE ANN. § 41-41-73 (West 1999); MO. REV. STAT. § 565.300 (2001); MONT. CODE ANN. § 50-20-109 (1999); NEB. REV. STAT. §§ 28-326, 28-328 (1998); N.J. STAT. ANN. § 2A-65A-6 (West 2000); OHIO REV. CODE ANN. § 2919.151 (West Supp. 2001); R.I. GEN. LAWS § 23-4.12 (2001); S.C. CODE ANN. § 44-41-85 (2001); S.D. CODIFIED LAWS § 34-23A-27 (Michie Supp. 2001); TENN. CODE ANN. § 39-15-209 (1997); UTAH CODE ANN. § 76-7-310.5 (1999); VT. STAT. ANN. TIT. § 18.2-74.2 (1998); VA. CODE ANN. § 18.2-74.2 (Michie 2001); WIS. STAT. § 940.16 (1997).

58. *See* Walther, *supra* note 7, at 706-07; *see also infra* notes 92-96 and accompanying text for a discussion of the three federal bills.

59. Many of the partial birth abortion statutes give a description of the proce-

1. Dilation and Evacuation Abortion Method

Dilation and evacuation (D&E) is the most commonly used abortion procedure after the thirteenth week of gestation.[60] There are several variations to the technique, which are used at different stages of fetal development during the second trimester.[61] During the thirteenth to fifteenth weeks of gestation the technique used is much like vacuum aspiration,[62] however, "the cervix must be dilated more widely because surgical instruments are used to remove larger pieces of tissue."[63] Once the cervix is dilated, the fetus and placenta are removed.[64] "Because fetal tissue is friable and easily broken, the fetus may not be removed intact. The walls of the uterus are scraped with a curette to ensure that no tissue remains."[65]

The technique used after fifteen weeks of gestation is somewhat different because the fetus (especially the head) is bigger, and the "bones are more rigid."[66] Because of this higher degree of development, "dismemberment or other destructive procedures are more likely to be required than at earlier gestational ages to remove fetal and placental tissue."[67] After

---

dure, which some courts have found to be vague. *See* Stenberg v. Carhart, 530 U.S. 914, 938-39 (2000); *see also* Evans v. Kelly, 977 F. Supp. 1283, 1311 (1997). These courts have struck down the "vague" statutes as the D&X procedure cannot be distinguished from the D&E procedure. *See id.*; *see also* Jennifer Landrum Elliott, Comment, *Will Charlie Brown Finally Kick the Football?: Missouri Enacts the Next Generation of Partial Birth Abortion Restrictions*, 44 ST. LOUIS U. L.J. 1083, 1097-98 (2000).

60. *See Stenberg*, 530 U.S. at 924, (citing CENTERS FOR DISEASE CONTROL AND PREVENTION, ABORTION SURVEILLANCE – UNITED STATES, 1996, 41 (July 30, 1999)). D&E "'together with a modified form of vacuum aspiration used in the early second trimester' accounts for about 95% of all abortions performed from 12 to 20 weeks of gestational age." *Id.*

61. *See Stenberg*, 530 U.S. at 925, (citing AMERICAN MEDICAL ASSOCIATION, REPORT OF BOARD OF TRUSTEES ON LATE-TERM ABORTION, APP. 490).

62. Vacuum aspiration is defined as "insertion of a vacuum tube (cannula) into the uterus to evacuate the contents." *Stenberg*, 530 U.S. at 923, (citing OBSTETRICS: NORMAL & PROBLEM PREGNANCIES 1253-54 (S. Gabbe, J. Niebyl, & J. Simpson eds. 3d ed. 1996)).

63. *Stenberg*, 530 U.S. at 924, (citing AMERICAN MEDICAL ASSOCIATION, REPORT OF BOARD OF TRUSTEES ON LATE-TERM ABORTION, APP. 490-91).

64. *See id.*

65. *Id.*

66. *Stenberg*, 530 U.S. at 925, *citing* AMERICAN MEDICAL ASSOCIATION, REPORT OF BOARD OF TRUSTEES ON LATE-TERM ABORTION, APP. 491.

67. *Id.* Dr. Carhart in his trial testimony described what is meant by dismemberment in D&E abortions: part of the fetus "such as a foot or a hand" is grasped by an instrument and pulled towards the vagina. *Stenberg*, 530 U.S. at 925 (citing 11 F.

twenty weeks of gestation, "[s]ome physicians use intrafetal potassium chloride or digoxin to induce fetal demise prior to a late D&E . . . to facilitate evacuation."[68] Although there are some variations, this is the primary technique for later term abortions.[69]

2. Dilation and Extraction Abortion Method

Dilation and Extraction (D&X), sometimes called intact D&E by the medical community, is commonly referred to as partial-birth abortion.[70] It is also the abortion method that states and Congress have attempted to ban.[71] D&X is performed after sixteen weeks of gestation,[72] as vacuum aspiration is no longer effective.[73] This form of abortion can "proceed[] in one of two ways, depending on the presentation of the fetus."[74] The physician can either perform the abortion as the fetus presents itself, or the physician can convert the presentation of the fetus to a footling breech position[75] (if the fetus is not already in that presentation).[76]

Supp. 2d at 1104 (testimony of Dr. Carhart)). Traction is created between the vagina and the uterus, which pulls the fetus apart into pieces. *See id.*

68. *Stenberg*, 530 U.S. at 925, *citing* AMERICAN MEDICAL ASSOCIATION, REPORT OF BOARD OF TRUSTEES ON LATE-TERM ABORTION, APP. 491-92.

69. *See Stenberg*, 530 U.S. at 925, *comparing* AMERICAN MEDICAL ASSOCIATION, REPORT OF BOARD OF TRUSTEES ON LATE-TERM ABORTION, APP. 490-92, *with* W. HERN, ABORTION PRACTICE 146-56 (1984), *and* MEDICAL AND SURGICAL ABORTION 133-35.

70. *See Stenberg*, 530 U.S. at 927, *citing* Carhart v. Stenberg, 11 F. Supp. 2d at 1111-12. D&X is commonly known as partial birth abortion because the fetus is partially delivered, intact, before it is killed and aborted. *See generally* MARTIN HASKELL, M.D., DILATION AND EXTRACTION FOR LATE SECOND TRIMESTER ABORTION, *reprinted in* 139 CONG. REC. E1092, 1093 (daily ed. Apr. 29, 1993) (describing D&X procedure). Delivery is defined as: "[t]o extract from an enclosed place, as the fetus from the womb." STEDMAN'S MEDICAL DICTIONARY 453 (26th ed. 1995).

71. *See supra* notes 57-58 and accompanying text; *see also infra* notes 92-96 and accompanying text.

72. *See* Stenberg v. Carhart 530 U.S. 914, 926 (2000), *citing* Carhart v. Stenberg, 11 F. Supp. 2d at 1105, *compare with* HASKELL, *supra* note 70 at E1092 (stating that D&X is performed on patients twenty to twenty-four weeks after their last menstrual period, and for some patients may be performed twenty-five to twenty-six weeks after their last menstrual period).

73. Vacuum aspiration is no longer effective because the size of the fetus's skull has grown too large. *See* HASKELL, *supra* note 70, at E1092.

74. *Stenberg*, 530 U.S. at 927 (quoting AMERICAN MEDICAL ASSOCIATION, REPORT OF BOARD OF TRUSTEES ON LATE-TERM ABORTION, APP. 490).

75. The fetus is in a footling breech position when the feet occupy the lower part of the uterus. *See* MICHAEL BRUDENELL ET AL, THE COMPLETE BOOK OF BABY CARE 45 (1978).

76. *See Stenberg*, 530 U.S. at 927 (quoting AMERICAN MEDICAL ASSOCIATION,

A physician may face one of two possible scenarios when following the first method. First, the fetus may be presented head first, "a vertex presentation."[77] Second the fetus may be presented feet first, "a breech presentation."[78] If the fetus is in a vertex presentation "the doctor collapses the skull; and the doctor then extracts the entire fetus through the cervix."[79] However, if the fetus presents itself in a breech position the doctor partially delivers the fetus into the cervix, with forceps, and then "collapses the skull, and extracts the fetus through the cervix."[80]

A doctor who follows the second possible method of D&X[81] manipulates the presentation of the fetus to a footling breech position.[82] This is the procedure the American College of Obstetricians and Gynecologists uses to describe D&X abortion.[83]

B. Statutes Banning Partial-Birth Abortions

Many people find the above described dilation and extraction method of performing abortion inhumane.[84] In response to D&X, thirty states have

---

REPORT OF BOARD OF TRUSTEES ON LATE-TERM ABORTION, APP. 490).

77. *Id.* The fetus is in a vertex presentation when "[t]he head . . . occupies . . . the lower part of the uterus." BRUDENELL, *supra* note 75, at 45.

78. *Stenberg*, 530 U.S. at 927 (quoting AMERICAN MEDICAL ASSOCIATION, REPORT OF BOARD OF TRUSTEES ON LATE-TERM ABORTION, APP. 490).

79. *Id.*, *citing* GYNECOLOGIC, OBSTETRIC, AND RELATED SURGERY 1043 (D. Nichols & D. Clarke-Pearson eds. 2d ed. 2000).

80. *Id.* (quoting AMERICAN MEDICAL ASSOCIATION, REPORT OF BOARD OF TRUSTEES ON LATE-TERM ABORTION, APP. 490); *see also* HASKELL, *supra* note 70, at E1093 (stating that the physician delivers the fetus, except the head, into the cervix, and then collapses the head by placing a suction catheter into the base of the skull and removing the fetus's brain so the head may then be fully delivered).

81. *See supra* notes 77-78 and accompanying text.

82. *See Stenberg*, 530 U.S. at 927-28 (citing GYNECOLOGIC, OBSTETRIC, AND RELATED SURGERY 1043 (D. Nichols & D. Clarke-Pearson eds. 2d ed. 2000)).

83. *See Stenberg*, 530 U.S. at 928, *citing* AMERICAN COLLEGE OF OBSTETRICIANS AND GYNECOLOGISTS EXECUTIVE BOARD, STATEMENT ON INTACT DILATION AND EXTRACTION APP. 599-560 (Jan. 12, 1997) which defines the second D&X method as:

> 1. deliberate dilation of the cervix, usually over a sequence of days;
> 2. instrumental conversion of the fetus to a footling breech;
> 3. breech extraction of the body excepting the head; and
> 4. partial evacuation of the intracranial contents of a living fetus to effect vaginal delivery of a dead but otherwise intact fetus.

Id.

84. *See Stenberg*, 530 U.S. at 920-23.

enacted statutes banning this method of abortion.[85] These statutes are similar, primarily because the legislatures were banning the same procedure.[86] In addition, many were modeled after the 1997 federal bill banning partial-birth abortions.[87]

The Iowa statute banning partial-birth abortions is a good example of these statutes.[88] The statute describes a partial-birth abortion as one "in which a person partially vaginally delivers a living fetus before killing the fetus and completing the delivery."[89] "Vaginally delivers a living fetus before killing the fetus" is defined as "deliberately and intentionally delivering into the vagina a living fetus or a substantial portion of a living fetus for the purpose of performing a procedure the person knows will kill the fetus, and then killing the fetus."[90] Significantly, the statute also includes an exception allowing for a partial-birth abortion if it is "necessary to save the life of the mother."[91]

Congress has tried to ban partial-birth abortions three times.[92] The first two federal bills were similar to the Iowa statute, but President Clinton vetoed both of them, and they never became law.[93] Congress tried again to ban partial-birth abortions with the 2000 bill.[94] However, because the holding of *Stenberg v. Carhart*[95] rendered the bill unconstitutional this was a futile effort.[96]

---

85. *See supra* note 57.

86. *See* Walther, *supra* note 7, at 707-08.

87. *See id.*

88. *See* IOWA CODE § 707.8A (2001).

89. IOWA CODE § 707.8A(1)(c) (2001).

90. IOWA CODE § 707.8A(1)(d) (2001).

91. IOWA CODE § 707.8A(2) (2001).

92. *See* Partial Birth Abortion Ban Act of 1995, H.R. REP. NO. 104-267, 104th Cong. (1995); Partial Birth Abortion Ban Act of 1997, H.R. REP. NO. 105-24 (1997) 105th Cong.; Partial Birth Abortion Ban Act of 1999, S. 1692, 106th Cong. (2000).

93. President Clinton vetoed the bill, because he felt the exception was not enough to protect the health of the mother. *See* Walther, *supra* note 7, at 707 n.116.

94. *See* S. 1692, 106th Cong. (2000).

95. 530 U.S. 914 (2000).

96. Under *Stenberg*, the statute would be an "undue burden" on a woman trying to obtain an abortion. *Id.* For a full discussion on this matter see *infra* notes 202-12 and accompanying text.

IV. FEDERAL CIRCUIT COURT DECISIONS ON THE CONSTITUTIONALITY OF STATUTES BANNING PARTIAL-BIRTH ABORTIONS

A. Circuit finding the Statutes Unconstitutional: The Sixth Circuit: *Women's Medical Professional Corp. v. Voinovich*[97]

In the *Voinovich* case an Ohio statute banning partial-birth abortions was challenged by the Women's Medical Professional Corporation, an abortion group that used the banned method.[98] The Ohio statute described partial-birth abortion as "the termination of a human pregnancy by purposefully inserting a suction device into the skull of a fetus to remove the brain."[99] A doctor charged criminally or civilly for performing this procedure had an affirmative defense under the statute if he could prove "that all other available abortion procedures would have posed a greater risk to the health of the pregnant woman."[100]

The Sixth Circuit found the Ohio statute unconstitutional for three main reasons: (1) It was vague; (2) it created an undue burden for women seeking abortions; and (3) it did not contain an adequate health exception.[101] The court found the statute unconstitutionally vague because the definition of partial-birth abortion, could also encompass the D&E procedure.[102] This would make it difficult for doctors to know whether or not they were violating the law, and "'[i]t is a basic principle of due process that an enactment is void for vagueness if its prohibitions are not clearly defined.'"[103]

The majority's vagueness argument is not convincing, as the dissent in *Voinovich* points out, the D&E procedure "does not terminate the pregnancy by purposely inserting [a] suction device into the fetal skull to excavate the contents of the skull."[104] In fact, the definition in the statute is the medical definition for the D&X procedure, and does not resemble the

---

97. 130 F.3d 187 (6th Cir. 1997).

98. *See id*; *see also* OHIO REV. CODE ANN. § 2919.151(C) (West 1996).

99. Women's Med. Prof'l Corp. v. Voinovich, 130 F.3d 187,198 (6th Cir. 1997) (quoting OHIO REV. CODE ANN. § 2919.151(A) (West 1996)).

100. OHIO REV. CODE ANN. § 2919.151(C) (West 1996).

101. *See* Gough, *supra* note 6, at 195-98, (citing *Voinovich*, 130 F.3d at 197-98, 201); *see also* Elizabeth A. Cooke, *South Carolina Bans Partial-Birth Abortions*, 49 S.C. L. REV. 1041, 1049-50 (1998) (citing *Voinovich*, 130 F.3d at 196, 201, 203).

102. *See Voinovich*, 130 F.3d at 198.

103. Gough, *supra* note 6, at 195 (quoting Grayned v. City of Rockford, 408 U.S. 104, 108 (1972)).

104. *Voinovich*, 130 F.3d at 215 (Boggs, J., dissenting); *see also supra* Part III.A1 & 2.

medical definition for the D&E procedure.[105] This makes it difficult to believe that a physician would not be able to tell whether or not he was in compliance with the statute. Therefore, the court's finding that the statute was void for vagueness is unconvincing.

The court also found the statute violated *Casey's* undue burden standard and created a substantial obstacle for a woman wanting a second trimester abortion.[106] The court found an undue burden because, as previously discussed, they felt the definition of partial-birth abortion could include the most common abortion procedure (D&E).[107] Since D&E would be prohibited as well, women would be deterred from receiving an abortion, as other methods may not be as safe as the D&E procedure - this would create a substantial obstacle.[108]

This analysis by the court is also questionable because it is based on the belief that the D&E procedure is prohibited by the statute.[109] As discussed above, the statute is clear in its definition of partial-birth abortion in that it only prohibits the D&X procedure.[110] The dissent in *Voinovich* also pointed out that prohibiting only D&X can not be an undue burden since there was no finding that D&X was a safer procedure or that there were no other safe procedures available for a woman seeking to have an abortion.[111] This analysis undermines the court's decision, because D&X is not the most common abortion method used and it is not a safer method, therefore banning it cannot place a substantial obstacle in front of a woman seeking a second trimester abortion.

The court's final reason for striking down the statute as unconstitutional was based on the health exception.[112] The court found the health exception inadequate because it spoke only of protecting the physical health of the mother and not her mental health.[113] The Sixth Circuit determined a mental health exception must be included, because the Supreme Court in *Doe v. Bolton*[114] defined health as including mental health.[115] Therefore, for the

---

105. *See* Part III.A.1 & 2, for the medical explanations of the D&E and D&X procedures.

106. *See* Gough, *supra* note 6, at 197, (citing *Voinovich*, 130 F.3d at 201 (finding "'that a statute which bans a common abortion procedure would constitute an undue burden'")).

107. *See* Gough, *supra* note 6, at 196-98.

108. *See id.*

109. *See id.*

110. *See supra* notes 102-05 and accompanying text.

111. *See* Walther, *supra* note 7, at 710, (citing *Voinovich*, 130 F.3d at 213 (Boggs, J., dissenting)).

112. *See Voinovich*, 130 F.3d at 206-10.

113. *See* Elliot, *supra* note 59, at 1102, (citing *Voinovich*, 130 F.3d at 206-10).

114. 410 U.S. 179 (1973).

115. *See id.* at 192.

health exception to be sufficient it must make provisions for a woman's mental health as well as her physical health, and because it did not, the statute was found unconstitutional.[116]

It is difficult to imagine a scenario where only the D&X procedure could be used to protect a woman's mental health.[117] As mentioned earlier, the dissent noted that it has not been found whether the D&X procedure is safer than other methods. Therefore, how can it ever be necessary to protect the health of a woman?[118] Further, by allowing a physician to decide when the D&X procedure is the only possible procedure that could be used to protect a woman's mental health would dilute the states' rights in regulating abortion.[119] It would give the power to regulate abortion to the physician, because he could use whatever method he found was needed regardless of any state laws on the matter.[120] Therefore, the Sixth Circuit's holding on the health exception seems to extend a bit too far.

B. Circuit Finding the Statutes Constitutional: The Seventh Circuit: *Hope Clinic v. Ryan*[121]

Two cases within the Seventh Circuit were decided differently concerning the constitutionality of partial-birth abortion statutes,[122] so that when both cases were appealed up to the Seventh Circuit, the court decided them together *en banc*.[123] The first case was *Hope Clinic v. Ryan*[124] from Illinois

---

116. *See* Elliot, *supra* note 59, at 1102, (citing *Voinovich*, 130 F.3d at 209 (holding that a woman must be able to obtain an abortion "if carrying a fetus to term would cause severe non-temporary mental and emotional harm")).

117. *See* Jeremy Alder, *Nebraska Decision Condones Infanticide*, U-WIRE (Austin, Texas), July 14, 2000, *available at* 2000 WL 21061150 (stating that the "'American Medical Association has declared the partial birth abortion procedure to be unnecessary'").

118. *See supra* note 111 and accompanying text.

119. *See* Dorinda C. Bordlee, *Partial Birth: What Next?*, CHRISTIANITY TODAY (Louisiana), Aug. 7, 2000, *available at* 2000 WL 11062785; *see also* Stenberg v. Carhart, 530 U.S. 914, 964 (2000) (Kennedy, J., dissenting) (stating that to "defer" to a doctor's judgment "is no different than forbidding [a state] from enacting a ban at all").

120. *See id.*; *see also* Hope Clinic v. Ryan, 195 F.3d 857, 873 (7th Cir. 1999) (holding that to allow a physician to decide when a procedure is necessary to protect the health of a woman would "amount to a rule that no state may regulate any abortion procedure").

121. 195 F.3d 857 (7th Cir. 1999).

122. The statutes were from Wisconsin and Illinois, and were virtually identical defining "partial birth abortion" as "partially vaginally delivers a living fetus" then kills it and follows by "completing delivery." 720 ILL. COMP. STAT. 513/10 (West 1999); *see also* WIS. STAT. § 940.16 (1997).

123. *See* Elliot, *supra* note 59, at 1104. The Seventh Circuit decided the two cases under the same name: *Hope Clinic v. Ryan*. *See* 195 F.3d 857 (7th Cir. 1999).

and the second case was *Planned Parenthood of Wisconsin v. Doyle*[125] from Wisconsin. The district court in *Ryan* struck down the statute as unconstitutional, whereas the district court in *Doyle* upheld the statute as constitutional.[126] The Seventh Circuit agreed with the *Doyle* court and upheld the statutes, as they could be interpreted as constitutional.[127]

The statutes in Wisconsin and Illinois were similar therefore, the court was able to decide them together.[128] The main issues before the Seventh Circuit were whether the definition of partial-birth abortion was vague as to include the D&E procedure and thereby creating an undue burden for a woman seeking a second trimester abortion; and whether the statutes' lack of a health exception left them unconstitutional.[129] The court found that the statutes' were not so vague as to render them unconstitutional and that a health exception was not necessary.[130]

The court first discussed the issue of vagueness.[131] In doing so the court noted that the reason the legislature did not use the more precise medical definition was to avoid creating a loophole where "any small variation (such as a change in the method of reducing the head size, or snipping off a toe to defeat the "otherwise intact" specification) would take the abortion outside the prohibition . . . ."[132] The court further found that although the definitions of partial-birth abortion contained in the statutes were not exact to the medical definition, abortion doctors know that the term refers only to the D&X method.[133] As proof of this the court looked to Indiana (which had a similar partial-birth abortion statute) and found that the number of D&E abortions performed each year did not decline in that state.[134] However, if any ambiguities should appear in the future, the court stated that the state courts could interpret the statutes so they only apply to the D&X method, as that was the legislature's intent.[135] Therefore, the court found

---

124. 995 F.Supp. 847 (N.D. Ill. 1998).

125. 44 F.Supp.2d 975 (W.D. Wis. 1999).

126. *See* Hope Clinic v. Ryan, 995 F.Supp. 847 (N.D. Ill. 1998); *see also* Planned Parenthood of Wis. v. Doyle, 44 F.Supp.2d 975 (W.D. Wis. 1999).

127. *See* Hope Clinic v. Ryan, 195 F.3d 857, 865 (7th Cir. 1999).

128. *See id.* at 862-63; *see also supra* note 123.

129. *See Hope Clinic*, 195 F.3d at 863-65.

130. *See id.* at 874-75.

131. *See id.* at 863-64; *see also* Kristin O'Connell, *Constitutional Law: State Partial Birth Abortion Statutes May Be Constitutional*, 27 J.L. MED. & ETHICS 384, 384 (1999).

132. *Hope Clinic*, 195 F.3d at 863.

133. *See Hope Clinic*, 195 F.3d at 864; *see also* Walther, *supra* note 7 at 714.

134. *See* 195 F.3d at 870-71. This proved to the court that physicians could tell the difference between the prohibited partial birth abortion method (D&X) and the legal D&E procedure. *See id.*; *see also* Walther, *supra* note 7, at 716.

135. *See Hope Clinic*, 195 F.3d at 867.

the statutes were not unconstitutionally vague.

"'Partial-birth abortion' has always been a legal term . . . to describe one specific abortion procedure," the D&X abortion.[136] Although physicians do not use the term "partial-birth abortion" to refer to the D&X procedure, as the Seventh Circuit noted, they are aware of its common usage.[137] Even if a physician were unsure whether the term "partial-birth abortion" referred only to the D&X procedure and not to the D&E procedure, the statutory definition, although not medically precise, does give enough information for a physician to conclude that the statute is only meant to ban the D&X procedure.[138] Since a physician will be able to determine that only the D&X abortion method is banned by the statute, it cannot be unconstitutionally vague.

The next issue the court addressed, whether the statute posed an undue burden on the woman's right to choose, is tied up with the vagueness issue,

---

136. Parker, *supra* note 10, at A10.

137. *See id.*; *see also Hope Clinic*, 195 F.3d at 864. As noted above, Indiana physicians have not stopped using the D&E abortion method, which is strong evidence that physicians do in fact know that the term "partial birth abortion" refers to only the D&X abortion method. *See supra* note 134 and accompanying text.

138. *See* James Bopp, Jr., & Curtis R. Cook, *Partial-Birth Abortion: The Final Frontier of Abortion Jurisprudence*, 14 ISSUES L. & MED. 3, 15-16 (1998) (discussing the Partial -Birth Abortion Ban Act of 1977, H.R. 1122, 105th Cong. § 1531(a) & (b) (Oct. 9, 1997) which is substantially similar to the Wisconsin and Illinois statutes). The statute is clear in what abortion method it is banning because:

> *Delivery* is a term of art: "in obstetrics the birth of a child; parturition." Thus, to partially deliver a fetus and violate the ban, the physician would deliberately and intentionally manipulate the body of the intact, living child, or a substantial portion thereof, into the vagina for the purpose of performing a procedure the physician knows will kill the fetus, and then kills the fetus.
>
> Therefore, the term deliver may not be construed to cover the following situation: A physician has dilated the cervix in preparation for performing an in-utero dismemberment procedure usually referred to as dilation and evacuation ("D&E"), but before he inserts his instruments into the uterus, he notices the baby's hand or foot poking into the vagina. Since the physician has done nothing to manipulate the body of the living baby into the vagina, he cannot be said to have partially delivered the child and the statute is inapplicable.
>
> Nor does the statute apply in the following situation: a woman presents with an incomplete spontaneous breech delivery ("spontaneous abortion") of a sixteen week old fetus with hydrocephalus (enlarged head), with the head lodged at the cervix and the baby still alive. Again, the physician has done nothing to manipulate the body of the living baby into the vagina, so the statute would not apply.

Id. at 16 (footnotes omitted).

in that, if the statute were vague and could cover multiple abortion methods, it would place a substantial obstacle in front of a woman looking to have a late term abortion.[139] Since the court determined that the statutes only banned the D&X abortion method, there could be no undue burden on the woman's right to choose.[140] The court also discussed whether the absence of a health exception from the statute created an undue burden.[141] There were two reasons why the Seventh Circuit found it did not.[142] The first was that it is never necessary to perform a D&X procedure to protect the health of the mother; and the second was that forcing a health exception would make the statute ineffective, because a physician could perform the procedure at any time and claim it was necessary to protect the health of the mother.[143]

Although the Wisconsin and Illinois statutes banned the D&X abortion procedure, the statutes do not create an undue burden, because as the Seventh Circuit stated, there were other abortion methods available to women, including the most commonly used D&E procedure.[144] If only one abortion method for late term abortions is taken away as an option, leaving the most common and safest method, then it is difficult to see how this could place a substantial obstacle in front of a woman. There is no substantial obstacle, because there are still methods by which she can have an abortion if she so desires.[145]

Further, there is no undue burden on the mother because of the statutes' lack of a health exception.[146] There is no need for a health exception because other abortion methods are available.[147] More importantly, not only is there no evidence that the D&X procedure would ever be necessary, but also there is evidence that it is more dangerous to the mother's physical health than the D&E procedure, which would still be available under the statute.[148] C. Everett Koop, a former Surgeon General, stated in 1996 that

---

139. *See Hope Clinic*, 195 F.3d at 871.

140. *See id.* at 874 (finding that since there were other abortion methods available which were not included in the ban, there was no undue burden on the woman's right to choose abortion); *see also* Elliot, *supra* note 59, at 1107.

141. *See Hope Clinic*, 195 F.3d at 871-72.

142. *See id.*

143. *See id.* at 872-74.

144. *See id.* at 874; *see also* Elliot, *supra* note 59, at 1107.

145. *See Hope Clinic*, 195 F.3d at 174; *see also* O'Connell, *supra* note 131 at 385 (stating that "because D&X abortions are not the only way to perform an abortion -- even in late[-]stage, second trimester abortions -- proscribing D&X abortions does not unduly burden a woman's right to an abortion").

146. *See Hope Clinic*, 195 F.3d at 875.

147. *See supra* notes 144-45 and accompanying text.

148. *See supra* note 10 and accompanying text. *See also* Colleen Reilly, *Partial-Birth Risks*, PATRIOT-NEWS, (Penn.), July 17, 2000, at A07 (quoting "Dr. Curtis

"'partial-birth abortion is never medically necessary to protect a mother's health or her future fertility. This procedure [D&X] can pose a significant threat to both.'"[149] Since the D&X procedure cannot ever be seen as necessary to protect the health of the mother, the Seventh Circuit's decision on that issue is a valid one.[150]

The Seventh Circuit's final point on whether the lack of a health exception created an undue burden was a practical one.[151] If the statute contained a health exception it would create two legal alternatives, both of which are unacceptable.[152] The first is that a judge or jury would decide whether the doctor acted correctly in deciding that the D&X procedure was necessary to protect the health of the mother.[153] This, as the court correctly points out, would be problematic, as it would lower the community's faith in the capability and judgement of its doctors.[154] The second alternative is that the judgement of whether the D&X procedure is necessary should be entirely left up to physicians.[155] This is problematic because the doctor could perform the D&X procedure, regardless of whether the medical community would have followed that course of action in the same situation.[156] In effect, this method would "amount to a rule that anything goes with regard to late-term abortions."[157] Therefore, the Seventh Circuit was correct to find that the lack of a health exception does not create an undue burden on a woman seeking a late term abortion.

V. STENBERG V. CARHART

A. Facts of the Case

In *Stenberg v. Carhart*,[158] the United States Supreme Court addressed a

---

Cook, a member of the American College of Obstetrics and Gynecology" saying "that the technique [the D&X Procedure] is largely abandoned in modern obstetrics because of unacceptable risk. The woman is placed at greater risk for both bleeding and delayed complications of infection."); Roger Myers, *Panel Ponders Bill to Ban Partial-Birth Abortions*, TOPEKA CAP. J., Mar. 9, 2000, at C4 (stating that "partial birth abortions shouldn't be performed because they are needlessly risky").

149. Reilly, *supra* note 148, at A07.

150. *See Hope Clinic*, 195 F.3d at 872-73; *see also supra* notes 146-49, and accompanying text.

151. *See Hope Clinic*, 195 F.3d at 873-74.

152. *See id.*

153. *See id.*; *see also* Walther, *supra* note 7, at 717 n.188.

154. *See Hope Clinic*, 195 F.3d at 873; *see also* Walther, *supra* note 7, at 717 n.188.

155. *See Hope Clinic*, 195 F.3d at 874.

156. *See id.*

157. Myers, *supra* note 148, at C4 (quoting Cynthia Patton, Esq.).

158. 530 U.S. 914 (2000).

Nebraska state statute banning partial-birth abortions.[159] The Nebraska statute is similar to the Iowa statute discussed above;[160] it defines partial-birth abortions as a "procedure in which the person performing the abortion partially delivers vaginally a living unborn child before killing the unborn child and completing the delivery."[161] "[P]artially delivers vaginally a living unborn child before killing the unborn child" is further defined as "deliberately and intentionally delivering into the vagina a living unborn child, or a substantial portion thereof, for the purpose of performing a procedure that the person performing such procedure knows will kill the unborn child and does kill the unborn child."[162] The Nebraska statute, like the Iowa statute, also has an exception to the partial-birth abortion ban if the abortion method is "necessary to save the life of the mother."[163]

The case began when Dr. Leroy Carhart, who performs partial-birth abortions in Nebraska, challenged the statute.[164] Dr. Carhart sued in federal district court for an injunction forbidding the enforcement of the statute, claiming it violated a woman's constitutional right to obtain an abortion.[165] The district court found in favor of Dr. Carhart, stating the statute violated the Constitution by placing an "undue burden" on a woman's right to obtain an abortion.[166] The Eighth Circuit affirmed the district court's holding, stating that the statute was an "undue burden" because its language was broad, and therefore, it would also ban not only D&X abortions, but also D&E abortions, which are the most commonly used.[167] The United States Supreme Court granted Nebraska's petition for certiorari, to consider whether the statute did in fact pose an undue burden on the mother's right to choose, and whether the statute violated the Constitution by not including a health exception.[168]

B. The Majority Opinion

The majority opinion in *Stenberg* was written by Justice Breyer and was joined by Justices Stevens, O'Connor, Souter and Ginsburg.[169] The Court affirmed the Eighth Circuit holding that the Nebraska statute was uncon-

---

159. *See* NEB. REV. STAT. §§ 28-326, 28-328 (1998).
160. *See supra* notes 88-96 and accompanying text.
161. NEB. REV. STAT. § 28-326(9) (1998).
162. *Id.*
163. NEB. REV. STAT. § 28-328(1).
164. *See Stenberg*, 530 U.S. at 922.
165. *See id.*
166. *See* Bower, *supra* note 56 at 657.
167. *See* Bower, *supra* note 56 at 657-58 (citing Carhart v. Stenberg, 192 F.3d 1142 (8th Cir. 1999) *cert. granted in part* 2000 WL 21145 (U.S. 2000)).
168. *See Stenberg*, 530 U.S. at 922.
169. *See id.* at 919.

stitutional because it places an "undue burden" on a woman's right to choose abortion.[170] The majority purported to base its opinion on the holding of *Casey*,[171] and also because the statute "lacks any exception 'for the preservation of the . . . health of the mother.'"[172]

The majority acknowledges that *Casey* allows a state to regulate and even ban post-viability abortions as long as there is an exception to save the life or health of the mother.[173] Following this line of reasoning, the Court found that since a life or health exception is required for regulations on post-viability abortions, it is also required for regulations of pre-viability abortions.[174] The majority further found that the health exception requires that a "[s]tate cannot subject women's health to significant risks both in that context, [where the pregnancy creates the health risk] *and also* where state regulations force women to use riskier methods of abortion."[175] The Court believed that in some instances the Nebraska statute forces women to use riskier abortion methods, and since the statute does not supply a *health* exception, it is unconstitutional.[176]

The majority's reasoning here is less than convincing. It is true that *Casey*'s *plurality* opinion requires a state regulation on abortion to make an exception when it is necessary to protect the health or life of the mother.[177] This however, is not the holding of the case and is not binding on the Court. Notwithstanding this, the Court seems to have ignored the fact that the D&X abortion method is much riskier for the mother, and banning this method could in fact protect the physical health of the mother.[178] There-

---

170. *See id.* at 922; *cf. supra* notes 144-45 and accompanying text.

171. *See also supra* notes 35-36 and accompanying text.

172. *Stenberg*, 530 U.S. at 930 (citing *Casey*, 505 U.S. at 879) (joint opinion of O'Connor, Kennedy, and Souter, JJ.).

173. *See Stenberg*, 530 U.S. at 930 (citing 505 U.S. at 879) (quoting *Roe*, 410 U.S. at 164-65). The Court in *Stenberg* refers to the joint opinion in *Casey*, not to the majority opinion. *See id.*

174. *See Stenberg*, 530 U.S. at 930.

175. *Id.* The Court states that this rule comes from *Casey*, but does not give a page on which it can be found. *See id.*

176. *See id.* at 934-35. The Nebraska statute provides an exception if the mother's *life* is in danger, but does not state an exception for when her *health* is in danger. *See supra* note 163 and accompanying text.

177. *See Casey*, 505 U.S. at 879 (joint opinion of O'Connor, Kennedy, and Souter, J.J.).

178. *See* Rick Santorum, *Santorum Remains Hopeful that Partial Birth Abortions Ban will be Enacted into Law*, GOV'T PRESS RELEASES, Oct. 20, 1999, *available in* 1999 WL 28846082 (citing the American Medical Association: "the partial birth abortion procedure is never medically necessary to save the mother's life. In fact, this procedure can actually threaten a mother's health and her ability to carry future children to term."); *see also supra* notes 10, 146-47 and accompanying text; and *infra* notes 186-87 and accompanying text.

fore, banning D&X would not force women to use riskier abortion methods, but safer ones, including the D&E abortion method.[179]

The Court also mentions that a state cannot ban a type of abortion if that abortion method is necessary to prevent a health risk which the pregnancy has created.[180] When the Court talks about a health risk, it is referring to "'all factors – physical, emotional, psychological, familial, and the woman's age – relevant to the well-being of the patient.'"[181] This is a broad definition and when looking at it, it becomes clear that almost anything can fall under this category of health risk, including emotional distress.[182]

Although one might think that doctors would not perform this disturbing procedure unless it was absolutely necessary, the opposite appears to be true. In a study of partial-birth abortions, twenty percent were done because of health reasons, most of which were mental in nature, and some of those were because the mother was agoraphobic (fear of leaving the home).[183] In another study a mere nine percent of the partial-birth abortions were for health reasons, and of those a majority were because the mother was suffering from depression, caused by her pregnancy.[184] These facts show that a vast majority of partial-birth abortions are not completed to protect the health of the mother.[185] Therefore, it is clear that in the minority of cases when the D&X procedure is chosen to protect the health of the mother it is not to protect the physical health of the mother, but to protect her mental health. However, it is hard to believe that the D&X procedure could ever actually protect a mother's mental health especially con-

---

179. *See supra* notes 10, 117-18, 148-49, 178 and accompanying text; *see also infra* notes 185, 219-20 and accompanying text.

180. *See supra* note 175 and accompanying text.

181. Bordlee, *supra* note 119, at 68 (citing, Doe v. Bolton, 410 U.S. 179, 192 (1973)).

182. *See id.* (stating that under this standard of health risk "a doctor can perform a partial birth abortion on a healthy baby of a healthy mother who may find the late-term pregnancy just too emotionally stressful").

183. *See* Bopp & Cook, *supra* note 138, at 9; *see also* Myers, *supra* note 148, at C4 (stating that a D&X could be used to protect a woman's mental health when she cannot handle the "emotional consequences of putting a child up for adoption").

184. *See* Bopp & Cook, *supra* note 138, at 11; *see also* Myers, *supra* note 148, at C4.

185. *See* Santorum, *supra* note 178 (finding that Ron Fitzsimmons, the Executive Director of the National Coalition of Abortion Providers, misled the public on this issue, the fact being "that partial birth abortion is a widespread, *elective* procedure used in the fifth and sixth month of gestation, usually on healthy babies and on healthy mothers.") (emphasis added); *see also* Alder, *supra* note 117 (stating that of the 1,500 partial birth abortions performed by one New Jersey abortion clinic most are performed "for elective, not medical reasons").

sidering the brutality of the D&X procedure.[186]

The following is an account by a nurse sitting in on the D&X procedure, it is dramatic, but accurately shows how mentally straining the D&X procedure can be on a mother:

> The mother was 6 months pregnant, 26 ½ weeks. A doctor told her that her baby had Down Syndrome, and she decided to have an abortion. She came in the first 2 days to have the laminaria inserted and changed, and *she cried the whole time*. On the third day she came in to receive the partial-birth abortion.
>
> Dr. Haskell brought the ultrasound in and hooked it up, so that he could see the baby. I could clearly see the baby's heartbeat on the ultrasound screen . . . .
>
> Dr. Haskell went in with forceps and grabbed the baby's legs and pulled them down into the birth canal. Then he delivered the baby's body and arms – everything but the head. The doctor kept the baby's head just inside the uterus.
>
> The baby's little fingers were clasping and unclasping, and his feet were kicking. Then the doctor stuck the scissors through the back of his head, and the baby's arms jerked out in a flinch, a startle reaction, like a baby does when you throw him up in the air and he thinks he might fall. The doctor opened up the scissors, stuck a high-powered suction tube into the opening, and sucked the baby's brains out. Now the baby went completely limp.
>
> I was really completely unprepared for what I was seeing. *I almost threw up* as I watched the doctor do these things . . . .
>
> Dr. Haskell [next] delivered the baby's head. He cut the umbilical cord and delivered the placenta. He threw that baby in a pan, along with the placenta and the instruments he used. I saw the baby move in the pan, and I asked another nurse and she said it was just "reflexes." . . . .
>
> The woman wanted to see her baby, so they cleaned the baby up and put it in a blanket and handed the baby to her. *She cried the whole time*. Looking down into the baby's face, she said, "*Please forgive me. I am so sorry. I love you.*" *I was crying too. I couldn't take it.*[187]

After reading this account it is impossible to believe that the D&X abortion method can ever protect a woman's mental health, therefore the Court

---

186. For a complete comparison between the D&X and D&E procedures see *supra* Part III.A.1 & 2.

187. *Effects of Anesthesia During a Partial-Birth Abortion: Thursday, March 21, 1996 House of Representatives, Subcommittee on the Constitution, Committee on the Judiciary*, Washington DC, 14 ISSUES L. & MED. 71, 99-100 (1998) (statement of Brenda Pratt Shafer, Registered Nurse) [hereinafter *Effects of Anesthesia*] (emphasis added).

mistakenly found that a health exception was necessary for the statute to be constitutional.

The Court next addressed the issue of state interests in regulating abortion.[188] For a state to regulate abortion before viability, it must further its interests in protecting the health of the mother[189] or the potentiality of human life.[190] The majority found that Nebraska could not further any interests in protecting the health of the mother because the statute created health risks for the mother.[191] The Court also found the statute did not "directly further an interest 'in protecting the potentiality of human life.'"[192] Nebraska argued the regulation furthered an interest in potential human life because it "'show[s] concern for the life of the unborn,'[and] 'prevent[s] cruelty to partially born children.'"[193] This statement did not move the majority because they found that to protect the potential for human life the statute would have to save the life of the fetus.[194]

The Court's analysis on this issue is shaky. As discussed above, the statute does not create a health risk for the mother if anything, it protects the mother from having a procedure that could cause her both physical and mental harm.[195] Therefore, the statute absolutely furthers the state's interests in protecting the health of the mother. It forces her to use safer abortion methods.[196]

The Court also found that the statute did not further the state's interest in protecting fetal life because it did not save the life of the fetus, however this is not a compelling statement in light of the facts.[197] Because this method of abortion is considered to be quite painful for the fetus, preventing it would spare the fetus needless pain.[198] Also, Nebraska's argument

---

188. *See* Stenberg v. Carhart, 530 U.S. 914, 930 (2000).

189. *See id.*; *see also supra* notes 21, 37 and accompanying text.

190. *See Stenberg*, 530 U.S. at 930; *see also supra* notes 36-39 and accompanying text.

191. *See Stenberg*, 530 U.S. at 931; *see also supra* notes 117-18, 148-49, 178-79, 185-187 and accompanying text.

192. *Stenberg*, 530 U.S. at 930.

193. *Id.* at 930-31 (citing Brief for Petitioners at 48).

194. *See id.* (citing Planned Parenthood of S.E. Pa. v. Casey, 505 U.S. 833, 885 (1992), where the Court found that a 24-hour waiting period furthered the state's interests in fetal life because a woman *might* change her mind during that period, and not go through with the abortion).

195. *See supra* notes 10, 117-18, 148-49, 178-87 and accompanying text.

196. *See Reilly*, *supra* 148, at A7.

197. *See Stenberg*, 530 U.S. at 930; *cf.* Planned Parenthood of S.E. Pa. v. Casey, 505 U.S. 833, 876 (1992). The *Casey* opinion never states that the fetus' life must be saved by a statute for the statute to promote the state's "interests in protecting fetal life or potential life." *Id.*

198. Many medical professionals agree that the D&X abortion method creates much pain for the fetus, during the whole procedure. *See Effects of Anesthesia*, *supra*

that the ban prevents cruelty to the fetus supports this, because to prevent pain is to prevent cruelty, and this will promote a respect for potential human life.[199] Considering this, a statute banning a form of painful cruelty necessarily promotes a state's interest in the fetal life.

The Nebraska statute was also struck down on the grounds that it posed an undue burden on the woman's right to choose abortion, as defined in *Casey*.[200] The Court found the statute "'plac[es][] a substantial obstacle in the path of a woman seeking an abortion of a nonviable fetus.'"[201] The statute creates a substantial obstacle for women because, as the court found, its language encompasses not only dilation and extraction (commonly known as partial-birth abortions), but also dilation and evacuation, which is the most commonly used second trimester abortion method.[202]

The language of the statute banning partial-birth abortions describes them as "delivering into the vagina . . . a substantial portion [of the living fetus]" for the purpose of killing the fetus.[203] The Court found this language to be broad and covering both D&E and D&X abortion procedures since both procedures call for delivery of at least a portion of the fetus into the vagina while it is still living.[204] The Court further found that, although the statute could be construed to apply to only D&X abortions, it had not been construed as such in the lower federal courts, and the Supreme Court "normally follows lower federal court interpretations of state law."[205] For that reason, the Court found the Nebraska Statute placed an "undue burden" on a woman's right to choose abortion.

The Court is overly broad in its interpretation of the statute because, as the Seventh Circuit found in *Hope Clinic v. Ryan*, the statute clearly prohibits only the D&X abortion procedure.[206] "Also, large numbers of phy-

---

note 187 at 93 and accompanying text (stating that the anesthesia the mother receives does not dull the pain for the fetus).

199. *Brief of the Attorney General of the State of Nebraska in Stenberg v. Carhart*, 16 ISSUES L. & MED. 3, 32 (2000) [hereinafter *Brief of the Attorney General*] (stating that the statute "serves the State's interest in preventing cruelty to partially-born children and 'unacceptable disrespect for potential human life'").

200. *See Stenberg*, 530 U.S. at 938; *see also* notes 35-38 and accompanying text.

201. *Stenberg*, 530 U.S. at 938 (citing *Casey*, 505 U.S. at 877).

202. *See id.*

203. NEB. REV. STAT. § 28-326(9) (1998). *See also supra* notes 136-38 and accompanying text.

204. *See Stenberg*, 530 U.S. at 938.

205. *Stenberg*, 530 U.S. at 940 (citing McMillian v. Monroe County, 520 U.S. 781 (1997) and Brocket v. Spokane Arcades, Inc., 472 U.S. 491 (1985)).

206. *See id.* at 938; *compare with* Hope Clinic v. Ryan, 195 F.3d 857, 864 (7th Cir. 1999) (finding that doctors know that "partial birth abortion" refers only to the D&X procedure). See also *supra* note 136-38 and accompanying text for an explana-

sicians are quite comfortable with the term partial-birth abortion. On May 19, 1997, the American Medical Association endorsed the federal Partial-Birth Abortion Ban Act after minor revisions were made, none of which involved tampering with the term partial-birth abortion."[207]

When examining the procedures for D&X and D&E abortions and comparing them with the definition for partial-birth abortion in the statute, it becomes obvious that the statute only prohibits the D&X abortion procedure.[208] During the D&E procedure, one does not purposely bring (deliver) a "substantial portion" of the fetus "into the vagina."[209] In the D&E procedure, if a "substantial portion" of the fetus (i.e. a foot or arm) were to enter the vagina, it would not have happened because the physician *purposely brought* (delivered) it into the vagina, and therefore the physician could not be guilty of violating the statute.[210] The Court further states that the D&E procedure requires delivering a portion of the fetus into the vagina, but that simply is not true.[211] In fact it calls for pulling a portion of the fetus *between* the uterus and the vagina for traction, therefore no parts of the fetus are purposely brought *into* the vagina.[212] Because the statute is clear in banning only the D&X abortion procedure and not the D&E abortion procedure, it cannot be an undue burden as the majority suggests.

C. The Concurring Opinions[213]

1. Justice Stevens

Justice Ginsburg joined Justice Stevens in his concurring opinion.[214] Justice Stevens concurred in the Court's decision to strike down the Nebraska statute, banning partial-birth abortions for two reasons.[215] The primary reason was that the statute violated the Fourteenth Amendment's

---

tion of why a partial birth abortion statute cannot be seen as including the D&E abortion procedure.

207. Bopp & Cook, *supra* note 138, at 21.

208. See *supra* sections III.A.1 & 2 for a complete description of the D&E and D&X abortion methods.

209. NEB. REV. STAT. § 28-326(9) (1998). *See also supra* notes 136-38 and accompanying text.

210. NEB. REV. STAT. § 28-326(9) (1998). *See also supra* notes 70, 136-38 and accompanying text.

211. *See Stenberg*, 530 U.S. at 939.

212. *See supra* notes 67 & 138.

213. There are three concurring opinions in *Stenberg v. Carhart*. *See* Stenberg v. Carhart, 530 U.S. 914 (2000). Justice Stevens writes the first, Justice O'Connor writes the second, and Justice Ginsburg writes the third. *See id.* at 946, 950-51.

214. *See Stenberg*, 530 U.S. at 946.

215. *See id.*

guarantee of liberty.[216] Justice Stevens further stated, that the liberty guarantee encompasses a woman's right to choose, and the states do not have "any legitimate interest in requiring a doctor to follow any procedure other than the one that he or she reasonably believes will best protect the woman in her exercise of this constitutional liberty."[217]

Justice Stevens appeared to overlook the fact that the D&X procedure is never considered to be necessary to protect the woman, and that, in fact, it presents a danger to her.[218] The state has a legitimate interest in protecting the health of the mother,[219] and if a vast majority of medical professionals agree that the procedure is never necessary and in fact is dangerous, the state may ban the procedure.[220] Furthermore, if the state were to allow physicians to always determine whether it is medically necessary for the D&X procedure to be performed, it would take away the state's power to regulate abortion.[221] It could also lower the esteem and trust in the medical community if physicians were to perform a procedure which is frowned upon by the majority of the profession.[222] Following that reasoning, the state does have a legitimate interest in preventing a doctor from using a specific abortion procedure.

The second reason why Justice Stevens believed the statute should be struck down was because both D&E and D&X abortion procedures were so similar that they could not be distinguished.[223] He found they were "equally gruesome," and the D&X procedure was not closer to infanticide than the D&E procedure, therefore it was "irrational" to ban only one.[224]

Although, one can agree with Justice Stevens that both the D&X and D&E are gruesome, it is not necessarily "irrational" to ban only one.[225] First, the D&X procedure poses a genuine health risk, both physical and

---

216. *See Stenberg*, 530 U.S. at 946 (Stevens, J., concurring).

217. *Id.*

218. *See supra* notes 10, 117-18, 148-49, 178-79, 185-87 and accompanying text.

219. *See* Planned Parenthood of S.E. Pa. v. Casey, 505 U.S. 833, 875-76 (1992).

220. *See* Alder, *supra* note 117 (quoting the American Medical Association as "declar[ing] the partial birth abortion procedure to be unnecessary and . . . urg[ing] Congress to ban the procedure"); *see also supra* notes 10, 117-18, 148-49, 178-79, 185-87 and accompanying text.

221. *See* Hope Clinic v. Ryan, 195 F.3d 857, 873-74 (7th Cir. 1999); *see also supra* notes 151-57 and accompanying text.

222. *See supra* notes 219-20. *See also Brief of the Attorney General*, *supra* note 199, at 32 (stating that the state has a legitimate interest in "preserving the integrity of the medical profession").

223. *See Stenberg*, 530 U.S. at 946-47.

224. *Id.*

225. *See id.*

mental, on the mother.[226] Second, the D&X procedure is more like infanticide because the fetus is almost completely delivered *alive* before it is killed.[227] The procedure is dangerous, it is not the most commonly used, and it is traumatic for both the mother and the fetus, therefore it can be distinguished from the D&E procedure, and it *is* rational to ban only the D&X procedure.

2. Justice O'Connor[228]

Justice O'Connor's concurring opinion has two main focus points: (1) the lack of a health exception in the statute,[229] and (2) that the statute poses an undue burden on the woman's right to choose.[230] Justice O'Connor recognizes that a state may completely ban abortions after the fetus is viable as long as there is an exception when it is necessary to save the "'life or health of the mother.'"[231] Building on this principle, she infers that a ban on a specific type of pre-viability abortion must also contain an exception when it is necessary to save the "life or health of the mother."[232] Therefore, since the Nebraska Statute does not contain an exception when it is necessary to save the health of the mother, she finds the statute is unconstitutional.[233]

Justice O'Connor bases her opinion on *Casey's* plurality opinion, which required a health exception for the mother in abortion legislation.[234] *Casey*, however, does not reach a scenario where a health exception would not be necessary, because there are still other, safer options available.[235]

---

226. *See supra* notes 10, 117-18, 148-49, 178-79, 185-87, 219-20 and accompanying text.

227. *See Brief of the Attorney General*, *supra* note 199, at 32; *see also* Jill R. Radloff, Note, *Partial-Birth Infanticide: An Alternate Legal and Medical Route to Banning Partial-Birth Procedures*, 83 MINN. L. REV. 1555, 1555-56 (1999). See *supra* section III.A.2, for a medical discussion of the D&X abortion procedure.

228. No other Justices joined Justice O'Connor in her concurring opinion. *See Stenberg*, 530 U.S. at 947 (O'Connor, J., concurring).

229. *See Stenberg*, 530 U.S. at 947 (O'Connor, J., concurring). *See supra* note 163 and accompanying text.

230. *See Stenberg*, 530 U.S. at 948 (O'Connor, J., concurring).

231. *Id.* at 947 (quoting Planned Parenthood of S.E. Pa. v. Casey, 505 U.S. 833, 879 (1992)).

232. *Id.* at 948.

233. *Id. See also* NEB. REV. STAT. § 28-328 (1998). The Nebraska statute does in fact have an exception to the ban of partial birth abortions where it is necessary to save the life of the mother. *See id.* § 28-328(1); *see also supra* note 163 and accompanying text.

234. *See Casey*, 505 U.S. at 879.

235. *See id.* at 880; see also *supra* notes 10, 117-18, 148-49, 178-79, 185-87, 219-20 and accompanying text for discussions on the lack of necessity of a health

Casey's requirement for a health exception, was intended for legislation which left no other safer alternatives for the woman.[236] This is not the case in the Nebraska statute.[237] The Nebraska statute bans only one type of abortion, leaving other options available, one of which happens to be the most commonly used late term abortion technique (D&E).[238] Therefore, Justice O'Connor is not following the precedent set forth in *Casey*, as she purports to, but is extending it to include legislation banning only one form of abortion.

Justice O'Connor goes on to find that the statute poses an undue burden on a woman's right to choose because it includes both D&E and D&X abortion procedures.[239] She further states that this is an undue burden because the D&E abortion procedure is "the most commonly used method for [performing] previability second trimester abortions."[240] Since D&E is the most common late term abortion procedure, banning it would create "'a substantial obstacle to a woman seeking an abortion'" before her fetus is viable.[241]

Justice O'Connor interprets *Casey's* "undue burden" standard to hold that a state cannot ban the commonly used abortion procedure because this would place a substantial obstacle in front of the woman.[242] This is an accurate interpretation of *Casey* since it states that a substantial obstacle is one that prevents a woman from obtaining an abortion, and banning the most commonly used abortion method could in some instances have that effect.[243] O'Connor, like the majority, however, is mistaken in finding that

---

exception in partial birth abortion legislation.

236. *See Casey*, 505 U.S. at 879-80 (ruling on a statute which placed substantial obstacles in front of women looking to obtain any type of abortion; if the women could not or would not fulfill the requirements of the statute, they were prevented from receiving any abortion).

237. *See generally* NEB. REV. STAT. §§ 28-326, 28-328 (1998); see also *supra* notes 159-63 and accompanying text for a discussion of the statute.

238. *See* NEB. REV. STAT. § 28-328 (2002); *see also Stenberg*, 530 U.S. at 949 (O'Connor, J., concurring) (quoting plurality); *and* Rebecca L. Andrews, Note, *The Unconstitutionality of State Legislation Banning "Partial-Birth" Abortion*, 8 B.U. PUB. INT. L.J. 521, 527 (1998).

239. *See* Stenberg v. Carhart, 530 U.S. 914, 948 (2000) (O'Connor, J., concurring) (stating "it is not possible to interpret the statute's language as applying only to the D&X procedure").

240. *See id. See also* Andrews, *supra* note 238, at 527 (stating that 85% of all abortions performed after twelve weeks are by the D&E method).

241. *Stenberg*, 530 U.S. at 949 (O'Connor, J., concurring) (quoting Planned Parenthood of S.E. Pa. v. Casey, 505 U.S. 833, 884 (1992)).

242. *See id.*

243. *See Casey*, 505 U.S. at 887-98 (holding that a spousal notification requirement would *prevent* some women from being able to obtain an abortion); see also *supra* section II.B. 1 for a complete explanation of the holding in *Casey*.

the Nebraska statute bans the D&E procedure.[244] As discussed above, the statute is clear, in that it only bans the D&X abortion method and not the D&E abortion method.[245] Therefore if the statute bans one method, but leaves available the most common one, it is difficult to conclude that the statute places a substantial obstacle in front of a woman wanting an abortion.

Although Justice O'Connor agrees with the majority that the Nebraska statute is unconstitutional, she does suggest that other state statutes banning partial-birth abortion are constitutional.[246] She finds that these statutes are more specific in their definitions, so they only ban the D&X procedure.[247] Following this theory, they are not an undue burden.[248] These statutes also provide a health exception, so that the banned D&X procedure may be used if it is necessary to save the health or life of the mother.[249] Therefore, had the Nebraska statute more specifically banned only the D&X abortion procedure, and included an exception to protect the health of the mother, Justice O'Connor would have upheld the statute.[250]

3. Justice Ginsburg

Justice Stevens joined Justice Ginsburg in her concurring opinion.[251] Justice Ginsburg wrote a separate concurring opinion to stress her belief that the only purpose for the statute was to "chip away at the private choice shielded by *Roe v. Wade*, even as modified by *Casey*."[252] She finds that this "chipping away" is the only interest the state had in banning partial-birth abortions because the ban does not protect the life or health of the mother, and no fetal life is saved.[253] Further, like Justice Stevens, Justice Ginsburg finds no distinction between D&E and D&X abortions.[254] There-

---

244. See *supra* notes 200-12 and accompanying text for a discussion and commentary on the majority opinion on this issue; see also *supra* notes 102-11, 131-38 and accompanying text for a discussion and commentary on circuit court opinions on the possible inclusion of D&E in the partial birth abortion statutes.

245. *See supra* notes 200-12 and accompanying text.

246. Justice O'Connor cites to Kansas, Utah and Montana statutes. *See Stenberg*, 530 U.S. at 950-51 (O'Connor, J., concurring) (citing KAN. STAT. ANN. § 65-6721(b)(2) (Supp. 1998)); MONT. CODE ANN. § 50-20-401(3)(c)(ii) (Supp. 1999); UTAH CODE ANN. § 76-7-310.5(1)(a) (1999).

247. *See id.*

248. *See Stenberg*, 530 U.S. at 950.

249. *See supra* note 246.

250. *See Stenberg*, 530 U.S. at 951.

251. *See id.*

252. *Id.* (Ginsburg, J., concurring) (citing Hope Clinic v. Ryan, 195 F.3d 857, 880-82 (7th Cir. 1999) (Posner, C.J., dissenting)).

253. *See* Stenberg v. Carhart, 530 U.S. at 952 (Ginsburg, J., concurring).

254. *See id.* (stating that D&E "is no less distressing or susceptible to gruesome

fore, Ginsburg finds that because the Nebraska statute interferes with a woman's constitutional right to a pre-viability abortion, and the state has no legitimate interest in the interference, the statute is unconstitutional.[255]

Contrary to Justice Ginsburg's belief, Nebraska has several legitimate interests in banning partial-birth abortions.[256] The state has an interest in "preventing cruelty to partially born children and 'unacceptable disrespect for potential human life' . . . [during] the D&X procedure, [where] the child is more outside the womb than inside . . . . Furthermore, the State has a compelling interest in erecting a barrier to infanticide."[257] The state also has an interest in protecting the health and safety of the mother, and "in preserving the integrity of the medical profession."[258]

These state interests *are* protected by the Nebraska statute.[259] It prevents a woman from undergoing a surgery, which could be dangerous to her health, and possibly leave her unable to ever carry a child to term.[260] It promotes respect for fetal life by stopping a cruel and barbaric procedure which causes the partially-born child pain.[261] Further, it protects the integrity of the medical profession, because so many people are horrified with this procedure, and have lost respect for medical professionals who perform it.[262]

---

description").

255. *See id.* at 952.

256. *Brief of the Attorney General*, *supra* note 199, at 32.

257. *Id. See also* Alder, *supra* note 117 (stating that "under the law, the only difference between a partial birth abortion and infanticide is about six inches, the length of the child's head and neck, still inside the mother's womb, upon execution").

258. *See Brief of the Attorney General*, *supra* note 199, at 32; *see also* Alder, *supra* note 117; Planned Parenthood of S.E. Pa. v. Casey, 505 U.S. 833, 878 (1992).

259. See *supra* notes 180-201 and accompanying text for a full discussion on how the statute furthers the state's interest in protecting the well-being of the mother, respect for fetal life, and the medical profession.

260. *See* Reilly, *supra* note 148 at A07 (quoting Dr. Cook as saying that "[h]e and other doctors have seen women who have had trouble maintaining pregnancies after undergoing a partial birth abortion").

261. *See* Santorum, *supra* note 178; *see also* Bordlee, *supra* note 119, at 68 (stating that "partial birth [abortion] bans serve the legitimate state interest of constructing a firewall between abortion and infanticide").

262. *See* Hope Clinic v. Ryan, 195 F.3d 857, 873-74 (7th Cir. 1999); *see also* Walther, *supra* note 7, at 717 n.188.

D. The Dissenting Opinions[263]

1. Chief Justice Rehnquist

The Chief Justice dissents in a short paragraph to state his belief that *Casey*[264] was decided incorrectly.[265] Regardless of this belief, however, he states that the majority in *Stenberg* did not correctly apply the holding from *Casey*, and therefore, he joins Justices Kennedy and Thomas in their dissents.[266]

2. Justice Scalia

Justice Scalia dissents for several reasons, but the first and primary reason for his dissent is that he believes that partial-birth abortions are inhuman and barbaric.[267] He feels that, by striking down a law that prohibits partial-birth abortions, an inhumane act, the Court has placed this opinion in that same category as the *Korematsu*[268] and *Dred Scott*[269] opinions.[270] This comparison expresses Scalia's feeling that the *Stenberg* opinion is a gross deviation from what the United States and its Constitution stands for – "'to establish Justice, insure domestic Tranquility, . . . and secure the Blessings of Liberty to ourselves and our Posterity.'"[271] Justice Scalia

---

263. There are four dissenting opinions in *Stenberg v. Carhart*. *See generally* Stenberg v. Carhart, 530 U.S. 914 (2000). Chief Justice Rehnquist writes the first, the second is written by Justice Scalia, Justice Kennedy writes the third, and Justice Thomas writes the fourth. *See id.* at 952, 953, 956, 980.

264. 505 U.S. 833 (1992).

265. *See Stenberg*, 530 U.S. at 952 (Rehnquist, C.J., dissenting).

266. *See id.* at 952-53 (stating "[d]espite my disagreement with the opinion, . . . the *Casey* joint opinion represents the holding of the Court in that case. I believe Justice Kennedy and Justice Thomas have correctly applied *Casey's* principles and join their dissenting opinions"). For a discussion on the majority holding from *Casey*, *see also supra* notes 24-44 and accompanying text. Chief Justice Rehnquist concurs with Justices Kennedy and Thomas in their application of the holding from *Casey*. *See Stenberg*, 530 U.S. at 952. For a discussion on Justices Kennedy's and Stevens' dissents, *see infra* sections V.D 3-4.

267. *See Stenberg*, 530 U.S. at 953, 954 (Scalia, J., dissenting). Scalia specifically refers to the Nebraska law banning partial birth abortions as a "humane (not to say anti-barbarian) law." *Id.*

268. 323 U.S. 214 (1944) (holding it constitutional to place a Japanese American in a "camp" during the Second World War).

269. 60 U.S. 393 (19 How. 1857) (holding that a black slave who escaped to a "free state" was not a "citizen" as written in the Constitution).

270. *See Stenberg*, 530 U.S at 953 (Scalia, J., dissenting) (stating "I am optimistic enough to believe that, one day, *Stenberg v. Carhart* will be assigned its rightful place in the History of this Court's jurisprudence beside *Korematsu* and *Dred Scott*.").

271. *Id.* (Scalia, J., dissenting) (quoting U.S. CONST. preamble). Following

further believes that this cannot be achieved if the Court does not allow the states to protect "our half-born posterity."[272]

These are valid points made by Justice Scalia. The majority's opinion is an inhumane act that allows the continued disrespect for human life.[273] Considered in this way, one cannot help but agree that this case does in fact bear a close resemblance with the decisions in *Korematsu*, and *Dred Scott*.[274] The majority's opinion is also unjust in that it contravenes the right of the states to regulate abortion, as guaranteed in *Casey*.[275] Further, the majority in this case does not insure domestic tranquility by allowing a procedure that horrifies a vast majority of people in this country.[276] Therefore, the majority opinion cannot uphold the values that the Constitution represents.

Justice Scalia further believes that *Casey* should be overruled.[277] In Scalia's opinion, *Casey* has no basis in the Constitution, or in history, and this makes it an unconstitutional decision.[278] When reading the Constitution, there is no explicit right for a woman to choose abortion, therefore, according to Scalia, this means the framers of the Constitution intended to leave the decision to the states to decide.[279] In addition, Scalia notes that this is a controversial issue, and "a 5-to-4 vote on a policy matter by

---

Scalia's logic, if the Constitution is meant to "secure the Blessings of Liberty to . . . our posterity", then it would be contradictory to also say that the Constitution prohibits the states from banning partial birth abortions, which destroy "our posterity." *Id.*

272. *See id.*

273. The majority has validated a procedure which shows a complete lack of respect for human life as it is painful for the child. It differs from infanticide only in the fact that the baby's head has not yet been delivered from the womb before the baby is killed. *See* Alder, *supra* note 117 (stating that when the D&X procedure is performed "all of the baby's organs are functioning, she can cry, suck her thumb, grasp with her hand and even has hair growing on top of her head").; *See also* Bordlee, *supra* note 119, at 68; Santorum, *supra* note 178; *and supra* note 257 and accompanying text.

274. *See supra* notes 268-69 and accompanying text.

275. *See* Planned Parenthood of S.E. Pa. v. Casey, 505 U.S. 833, 876 (1992); *see also supra* section II.B.1. for an in-depth discussion of the majority opinion in *Casey*.

276. This holding will create the opposite of tranquility, because so many are opposed to partial birth abortions. *See supra* notes 7 & 57 and accompanying text.

277. *See* Stenberg v. Carhart, 530 U.S. 914, 956 (2000) (Scalia, J., dissenting).

278. *See id.* Justice Scalia is a strict constructionist, which explains his belief that for *Casey* to be upheld, the right of a woman to choose abortion must be written in the Constitution, or it must be deeply rooted in our nation's history. *See, e.g.*, Harmelin v. Michigan, 501 U.S. 957 (1991) (Scalia, J., concurring) (finding no proportionality guarantee because it is not specifically stated in the Constitution or deeply rooted in the nation's history).

279. *See Stenberg*, 530 U.S. at 956 (Scalia, J., dissenting); *see also* U.S. CONST. amend. X (giving states any "powers not delegated to the United States by the Constitution, nor prohibited by it to the States").

unelected lawyers should not overcome the judgement of 30 state legislatures."[280] He further states that "the Court should return this matter to the people—where the Constitution, by its silence on the subject, left it."[281]

Justice Scalia, being a strict constructionist, goes a little too far in this argument. Although the right to an abortion is not actually written in the Constitution, and is not deeply rooted in our nation's history, it is part of a woman's liberty interest, and to ignore this right could be harmful to her health.[282] However, he is correct in stating that the states should have the power to decide this issue.[283] Under *Casey* the states have the right to regulate abortion as long as it does not place a substantial obstacle in front of a woman seeking an abortion.[284] Because this statute still permits the most common late term abortion procedure, it cannot be seen as placing a substantial obstacle in front of the woman, and therefore, the issue, as Justice Scalia believes, should be left to the states to decide.[285]

Scalia also states that the right to abortion is not firmly rooted in United States' history.[286] He mentions this because in the past, the Court has upheld rights that were not specifically mentioned in the Constitution if they were firmly rooted in our Nation's history.[287] Following this logic, Scalia states that "the legal traditions of the American people would have sustained such a limitation upon abortion."[288] This argument has validity; given the history of abortion in this country it is clear that this legislation would have been upheld.[289]

Scalia further believes *Casey* should be overruled because its undue burden standard is "'hopelessly unworkable in practice.'"[290] Scalia finds

---

280. *Stenberg*, 530 U.S. at 955 (Scalia, J., dissenting).

281. *Id.* at 956; *see also supra* note 14.

282. *See* Roe v. Wade, 410 U.S. 113, 153 (1973); *see also supra* Part II.A, for a complete discussion on the holding of *Roe* and why abortion is guaranteed by the Constitution.

283. *See* Planned Parenthood of S.E. Pa. v. Casey, 505 U.S. 833, 877 (1992) (holding that a state may regulate abortion as long as it does not create an undue burden on the woman's right to choose to have an abortion).

284. *See id.*

285. *See Stenberg*, 530 U.S. at 954-55 (Scalia, J., dissenting); *see also supra* notes 278-81 and accompanying text.

286. *See Stenberg*, 530 U.S. at 956 (Scalia, J., dissenting).

287. *See id.* at 954-56. The Court has traditionally protected certain rights such as marriage, even though it is not written in the Constitution, because it is considered such a basic part of the United States' history. *See, e.g.*, Loving v. Virginia, 388 U.S. 1 (1967) (holding marriage is deeply rooted in our country's history, and therefore, the framers did not feel the need to specifically state it in the Constitution).

288. *Stenberg*, 530 U.S. at 955 (Scalia, J., dissenting).

289. The states have, throughout history, tried to ban abortion. *See supra* note 14 for a history of how abortion has been treated in this country.

290. *Stenberg*, 530 U.S. at 954 (quoting Planned Parenthood of S.E. Pa. v. Ca-

this standard entirely too vague and therefore, it cannot be properly applied.[291] He supports this statement by saying that what he considers to be an undue burden "is different from what the majority considers to be an 'undue burden.'"[292] Scalia states that the undue burden standard is a value standard, and that he values the life of the partially-born child more than the freedom of the mother to abort it.[293] By placing more value on the life of the partially-born child, Scalia does not find that banning partial-birth abortions places a substantial obstacle in front of the woman who wants an abortion.[294] Scalia believes that since the Justices cannot agree on how this standard should be applied then the standard is too vague, because any opinion based on this standard is really only based on the individual Justices' values.[295]

The "undue burden" standard from *Casey* is not as unworkable as Justice Scalia states. Without a context, it is a vague standard, but when examining the *Casey* decision and its application of the standard, the standard becomes much less vague. Before viability, if the regulation would ultimately prevent a woman from having an abortion, the regulation is deemed an undue burden.[296] If however, the regulation would not prevent abortion, but would make it a little more difficult to obtain, the mother's right to have an abortion must be balanced with the state's right to protect its interests.[297] If the state's interests outweigh the mother's right to abortion then the regulation is not an undue burden.[298] This standard is not too vague, and if followed properly, it is easily applied.

Finally, Scalia feels there is no need for a health exception in the Nebraska statute banning partial-birth abortions.[299] He finds that a health exception would contravene the purposes of having the statute, because it would give doctors the power to decide whether the procedure should be used.[300] Scalia also points out that there is no evidence that the D&X procedure would ever be necessary to protect the health of the mother.[301]

---

sey, 505 U.S. 833, 986 (1992) (Scalia, J., dissenting)).

291. *See id.* at 954-55.

292. *Id.* at 954.

293. *See id.* at 954-55.

294. *See id.* at 955.

295. *See id.*

296. *See Casey*, 505 U.S. at 877.

297. *See id.* at 876, 882, 885.

298. *See id.*

299. *See Stenberg*, 530 U.S. at 953 (Scalia J., dissenting).

300. *See id.* Scalia finds that the health exception would make the statute completely ineffective because it will be the doctor who decides which procedure to use with or without the statute in operation. *See id.*

301. *See Stenberg*, 530 U.S. at 954 (Scalia, J., dissenting); *see also id.* at 963 (Kennedy, J., dissenting) (citing American Medical Association, Board of Trustees

Justice Scalia correctly finds that there is no need for a health exception, because as discussed above, the D&X procedure is never necessary to protect the health of the woman, and in fact can be more dangerous than the D&E procedure.[302] Also as the Seventh Circuit decided in *Hope Clinic v. Ryan*, allowing a health exception would take away the power of the statute.[303] Since the definition of *health* is so broad, physicians could perform the procedure whenever they wanted to and the statute would in effect be meaningless.[304]

3. Justice Kennedy[305]

Justice Kennedy dissents from the majority opinion because it does not accurately apply *Casey*.[306] The majority did not follow *Casey*, in that they did not give sufficient importance to the states' interests, they ignored the fact that there were still abortion procedures available and they found a necessity for a health exception.[307] Kennedy stated that *Casey* broadened the importance of the states' interests in regulating abortion, and that *Casey* is "premised on the States having an important constitutional role in defining their interests in the abortion debate."[308] For example, the state has an interest in preventing a procedure which would "cause the medical profession or society as a whole to become insensitive, even disdainful, to life, including life in the human fetus."[309] Kennedy further stated that these interests are legitimate interests to regulate an abortion method and that the majority's view of *Casey* was too narrow, limiting state interests to

---

Factsheet on HR 1122 (June 1997)); M. HASKELL, DILATION AND EXTRACTION FOR LATE SECOND TRIMESTER ABORTION, *reprinted in* 139 CONG. REC. E1092 (daily ed. Apr. 29, 1993); *and Stenberg* 530 U.S. at 965-66 (Kennedy, J., dissenting) (citing AMERICAN COLLEGE OF OBSTETRICIANS AND GYNECOLOGISTS EXECUTIVE BOARD, *Statement on Intact Dilation and Extraction* (Jan. 12, 1997)). For a more in depth discussion on the non-necessity of a health exception for a ban on D&X abortions see *supra* notes 10, 117-18, 148-49, 151-57, 178-79, 185-87 & 219-20 and accompanying text.

302. *See supra* notes 10, 117-18, 148-49, 178-79, 185-87 & 219-20 and accompanying text.

303. 195 F.3d 857, 874 (7th Cir. 1999); *see also supra* notes 151-57 and accompanying text.

304. *See supra* notes 180-87 & 300 and accompanying text.

305. Chief Justice Rehnquist joins Justice Kennedy in his dissent. *See Stenberg*, 530 U.S. at 956 (Kennedy, J., dissenting).

306. *See id.* at 956-57 (Kennedy, J., dissenting).

307. *See id.* at 957, 965-66, 977-78.

308. *Id.* at 961 (finding that the majority did not place a "proper weight" on Nebraska's interests in banning the D&X abortion procedure).

309. *Id.*

protecting the health of the mother and the potentiality of human life.[310]

Justice Kennedy correctly interprets *Casey* to allow the states greater breadth in determining what interests are important for them to protect when creating abortion legislation.[311] This analysis makes it more difficult to find state legislation banning abortion unconstitutional, because it places a greater weight on the state's interests, and makes it easier to outweigh the woman's right to choose.[312] However, even using the narrower holding the majority employed, the statute should be found constitutional as it does promote the state's interest in protecting the health of the mother, and also protects the state's interests in fetal life.[313]

Casey's undue burden standard, according to Kennedy, requires only that a regulation does not impose a significant threat to the life or health of the woman, and "'which in no real sense depriv[e] the women of the ultimate decision.'"[314] Further, Kennedy disagreed with the majority by finding that the statute did not pose a significant threat to the life or health of the woman, and it did not deprive her of obtaining an abortion because the statute only prohibited the D&X abortion procedure, leaving other abortion procedures available.[315] Since the statute only prohibited the D&X procedure, there were still other options available for a woman wanting an abortion, such as the D&E procedure, therefore, the statute could not pose a threat to her health or eliminate the abortion option.[316]

---

310. *See id.* at 962; *see also supra* notes 190-91. For a full discussion on Nebraska's interests see notes 256-62 and accompanying text.

311. *See* Planned Parenthood of Southeastern Pa. v. Casey, 505 U.S. 833, 875 (1992) (holding that a woman does not have the "right to decide whether to have an abortion 'without interference from the State'").

312. *See id.* at 876, 882, 885; *see also supra* notes 299-300 and accompanying text.

313. *See Brief of the Attorney General, supra* note 199, at 32; *see also supra* notes 189-99 and accompanying text for a discussion on how the statute furthers the state's interests in protecting the health of the mother and potential life.

314. *Stenberg*, 530 U.S. at 965 (Kennedy, J., dissenting) (quoting *Casey*, 505 U.S. at 875) (alteration in original).

315. *See id.* at 974. Kennedy found that the statute only included the D&X procedure because, the term partial birth abortion is commonly known to refer to the D&X abortion procedure. *See id.* He quoted the American Medical Association as stating "The 'partial birth abortion' legislation is by its very name aimed exclusively [at the D&X.] There is no other abortion procedure which could be confused with that description." *Id.* (alteration in original). Kennedy further found that the definition of the procedure in the statute could only refer to the D&X procedure when the medical descriptions of the procedures were considered, therefore, a physician is provided with "the fair warning required by the law." *Id.*; *see also* NEB. REV. STAT. §§ 28-326, 28-328 (1998); *cf. supra* Parts III.A.1 & 2, which give a complete medical description of the D&E and D&X abortion procedures.

316. *See Stenberg*, 530 U.S. at 965-66.

Kennedy is a little too generous to the states when giving the undue burden standard from *Casey*. In fact, as discussed above, the undue burden standard firstly requires an examination as to whether or not the regulation would ultimately prevent a woman from obtaining an abortion.[317] If it would not, then there is a balancing between the state's rights in regulating the abortion, and the woman's rights in obtaining an abortion.[318] Following this standard, the statute would be upheld because it does not prevent a woman from receiving an abortion.[319] The state's interests in protecting the health of the mother, respecting human life and the medical profession, and erecting a barrier between abortion and infanticide, outweigh the mother's interests in receiving this type of abortion.[320] So, although the statute would be upheld under Kennedy's standard, which is more generous to the states, the statute could still be upheld under the actual standard.

Kennedy further disagreed with the majority because he did not find that a health exception was a good idea, let alone necessary.[321] It would not be a good idea because it would take away the state's power to regulate abortion and make the statute ineffectual.[322] "Requiring Nebraska to defer to Dr. Carhart's judgement is no different than forbidding Nebraska from enacting a ban at all; for it is now Dr. Leroy Carhart who sets abortion policy for the State of Nebraska, not the legislature or the people."[323] These policy reasons aside, Justice Kennedy also does not think that a health exception is necessary because "[n]o studies support the contention that the D&X abortion method is safer than other abortion methods."[324] If there is no proof that D&X could ever be the only possible procedure to protect the health of the mother, then Kennedy, cannot find that a health

317. *See Casey*, 505 U.S. at 876, 882, 885.

318. *See id.*

319. As Justice Kennedy concluded, the statute only bans the D&X procedure and therefore, leaves other safe abortion methods available, such as the D&E procedure. *See supra* note 315; *see also* Parker, *supra* note 10, at A10 (stating that partial birth abortions are a known term to describe D&X); Bopp & Cook, *supra* note 138, at 15-16 (stating that the definitions to partial birth abortion in statutes can only refer to the D&X method).

320. *See Brief of the Attorney General*, *supra* note 199, at 32 (noting that the state has an interest in preventing disrespect towards fetal life and the medical profession); *see also* Santorum, *supra* note 170 (stating that banning the D&X procedure brings humanity to the community).

321. *See Stenberg*, 530 U.S. at 964 (stating that a health exception is a bad idea because "[c]asting aside the views of distinguished physicians and the statements of leading medical organizations, the Court awards each physician a veto power over the State's judgement that the procedure should not be performed").

322. *See id.* at 964-65.

323. *Id.* at 965.

324. *Id.* at 966.

exception is necessary for the statute to be constitutional.[325]

Like the court in *Hope Clinic v. Ryan*, Justice Kennedy takes a practical look at the effects of having a health exception requirement in the partial-birth abortion statute.[326] This is a sensible analysis because requiring a health exception would repudiate the states' power to regulate abortion, as granted by *Casey*.[327] If a doctor could decide without regard to the statute whether he felt it was necessary to perform the D&X procedure, it would make the statute meaningless.[328] If a physician can decide when *he* finds it necessary to use the procedure, then the state has lost its power to regulate abortion, and by requiring a health exception the majority has in effect overruled *Casey*.[329] Further, as discussed in depth above, a health exception is not necessary to protect the mother.[330]

4. Justice Thomas[331]

Justice Thomas, like Chief Justice Rehnquist and Justice Scalia, believed that *Roe v. Wade* was wrongly decided.[332] He states that the right to abortion is not within the text of the Constitution and "[a]lthough a State *may* permit abortion, nothing in the Constitution dictates that a State *must* do so."[333] Following this, the Nebraska statute should be upheld, because Nebraska is not violating any Constitutional right by banning an abortion technique.[334]

If one were to interpret the Constitution as narrowly as Justice Thomas would like, the citizens of the United States would lose more than just their

---

325. *See id.* at 966-67.

326. 195 F.3d 857, 874 (7th Cir. 1999) (finding that a health exception would allow a doctor to perform the banned procedure, even if the medical profession as a whole would not have followed that course of action). *See also supra* notes 153-57 and accompanying text.

327. *See* Planned Parenthood of S.E. Pa. v. Casey, 505 U.S. 833, 876 (1992) (holding that the woman's right to an abortion is not an absolute right, and it may be interfered with by the state, to support legitimate state interests).

328. *See* Myers, *supra* note 148, at C4.

329. The majority opinion purports to follow *Casey*, however, by requiring a health exception they are in fact overruling *Casey*, because they are taking away the broad power of the state to regulate abortion, and placing it in the hands of the physician. *Compare Casey*, 505 U.S. at 876 *with* Stenberg v. Carhart, 530 U.S. 930 (2000); *see also supra* notes 180-86 & 188-99 and accompanying text.

330. *See supra* notes 10, 117-18, 148-49, 178-79, 185-87 & 219-20 and accompanying text.

331. Chief Justice Rehnquist and Justice Scalia joined Justice Thomas in his dissent. *See Stenberg*, 530 U.S. at 980 (Thomas, J., dissenting).

332. *See id.*

333. *Id.*

334. *See id.* at 982.

right to an abortion.[335] As discussed above, the right to an abortion may not be specifically stated in the Constitution, but it is a right that is implicit in the Constitution.[336] It is not an absolute right, however. It may be infringed upon by the states, as long as that infringement would not pose an undue burden in the exercise of that right.[337]

Justice Thomas believes that the majority opinion is wrong because it recognizes an individual right to abortion. He goes on to state that even if there were such a right, and *Casey* was the controlling precedent, the majority did not correctly follow that precedent as they claim to, and, in effect, overruled it.[338] First, Thomas states that the majority did not give enough credence to Nebraska's state interests.[339] Second, the majority ignored the doctrine of Constitutional Avoidance.[340] Third, the statute did not pose an undue burden on the woman's right to choose an abortion.[341] And fourth, the health exception is too expansive, and is not necessary.[342]

Under *Casey*, "seven Members of th[e] Court, . . . acknowledged that States have a legitimate role in regulating abortion and recognized the States' interest in respecting fetal life at all stages of development."[343] Thomas states that the majority violates *Casey*, because it does not give credence to the state's interests, which are to "express . . . profound respect [for the partially-born] by prohibiting a procedure that approaches infanticide, and thereby dehumanizes the fetus and trivializes human life."[344] By not recognizing Nebraska's interests in regulating abortion, the majority takes away the broad interests the state had in regulating abortion under

---

335. For example, the right to privacy, marriage, and even expression (i.e. what clothes you wear) would no longer be protected as they are not specifically articulated in the Constitution. *See generally* U.S. CONST.

336. *See supra* notes 14-17; *see also* Roe v. Wade, 410 U.S. 113, 153-58 (1973).

337. *See id.*; *see also Casey*, 505 U.S. at 876-77.

338. *See Stenberg*, 530 U.S. at 982.

339. *See id.*

340. *See id.* at 996. The doctrine of constitutional avoidance is employed when it is possible for the Court to narrow the interpretation of the statute so as to "avoid the constitutional question." *Id.*

341. *See id.* at 1014.

342. *See id.* at 1011-12.

343. Stenberg v. Carhart, 530 U.S. 914, 981 (2000) (Thomas, J., dissenting) *citing* Planned Parenthood of S.E. Pa. v. Casey, 505 U.S. 833, 877 (O'Connor, Kennedy, and Souter, J.J., joint opinion), 944 (Rehnquist, C.J., White, Scalia, and Thomas, J.J., concurring in judgement in part and dissenting in judgement in part).

344. *Id.* at 1006. The majority opinion stated that a state only has an interest in protecting the health of the mother, and not the life of the fetus. *See Stenberg*, 530 U.S. at 930. The majority also found that a state's interest in the life of the fetus could only be furthered when that life was saved. *See id.*; *see also supra* notes 188-94 and accompanying text.

Casey.[345]

The majority in this case did take away the state's broad power to determine its interests in regulating abortion when they refused to recognize Nebraska's interest in placing a barrier between abortion and infanticide.[346] The issue is not whether the majority thinks that banning partial-birth abortion will further this interest, but whether Nebraska thinks it will.[347] Further, Nebraska has ample reason to believe that this statute will form a barrier between partial-birth abortion and infanticide.[348] Therefore, Nebraska has a valid interest in erecting a barrier between infanticide and abortion.[349]

Justice Thomas's second disagreement with the majority concerns their failure to follow the doctrine of Constitutional Avoidance.[350] According to Justice Thomas, the Nebraska statute could have been interpreted to be constitutional, and therefore it should have been.[351] Justice Thomas outlined three reasons for his assertion. First, the term partial-birth abortion is commonly known to mean D&X abortion.[352] Second, the definition of

---

345. *See Casey*, 505 U.S. at 876; *see also supra* notes 36 & 39 and accompanying text.

346. *See Stenberg*, 530 U.S. at 930.

347. *See Casey*, 505 U.S. at 874 (stating that a state's interests are valid as long as they do not have the *purpose* or effect of preventing a woman from obtaining an abortion). The Nebraska statute does not have the purpose or effect of preventing a woman from receiving an abortion, therefore Nebraska's interests are valid under *Casey*. *See id.*

348. *See* Alder, *supra* note 117 (stating that partial birth abortion would be identical to infanticide, "under the law" if the head of the child were also delivered, before the child was killed). There are also cases where the only reason the child is not fully delivered before it is killed is because the doctor is holding its head inside the womb. *See Stenberg*, 530 U.S. at 988 (Thomas, J., dissenting) (quoting a statement by Dr. Pamela Smith).

349. Justice Thomas focuses only on Nebraska's interest in promoting respect for fetal life by placing a barrier between infanticide and abortion. However, as discussed above, Nebraska had several other valid interests upon which the majority did not place sufficient weight. *See Stenberg*, 530 U.S. at 930; *see also supra* notes 188-99, 222, 256-58 & 320 and accompanying text, for discussions on Nebraska's interests in banning partial birth abortion.

350. *See Stenberg*, 530 U.S. at 996. Constitutional avoidance is a doctrine whereby the Court, whenever possible, interprets statutes narrowly to refrain from reaching constitutional difficulties. *See id.* (Thomas, J., dissenting) (citing Frisby v. Schultz, 487 U.S. 474, 482 (1988); United States v. Morton, 467 U.S. 822, 828 (1984); Erznoznik v. Jacksonville, 422 U.S. 205, 216 (1975)).

351. *See supra* note 350.

352. *See Stenberg*, 530 U.S. at 994-95 (Thomas, J., dissenting) (citing the use of the term partial birth abortion to refer to D&X by "a majority of state legislatures, the United States Congress, medical journals, physicians, reporters," the American Medical Association, and judges).

partial-birth abortion in the statute, when read in context, can only be read as banning D&X abortion.[353] Since, as Justice Thomas contends, the statute can be read as referring to D&X abortion only, it should be read as such and the statute should be upheld as constitutional.[354]

Because Justice Thomas has found that the Nebraska statute only bans the D&X abortion method and not the D&E abortion method, the statute cannot pose an undue burden on a woman exercising her right to abortion.[355] Thomas states that under *Casey*, as long as the statute does not prevent a woman from obtaining an abortion, it is not an undue burden.[356] Since the statute only bans the D&X abortion procedure, and the more common D&E abortion procedure is still available, the statute cannot pose an undue burden on a woman seeking an abortion.[357]

Justice Thomas misstates the undue burden standard of *Casey* when he says that a regulation is not an undue burden if it does not prevent the ability of a woman to receive an abortion.[358] If the regulation would pose such a difficulty as to prevent some women from obtaining an abortion it would be an undue burden.[359] However, Justice Thomas is correct in concluding that this statute would not pose an undue burden, as it leaves the most common abortion procedure available.[360]

Justice Thomas states that the majority expands the *Casey* health exception to the point that "[t]he exception entirely swallows the rule."[361] Under *Casey* the health exception was required "if '*continuing her pregnancy* would constitute a threat' to the woman."[362] In his dissent, Justice Thomas states that the majority broadens the health exception, because they find a

---

353. *See id.* at 990 (Thomas, J., dissenting).

354. *See id.* at 990-91. Justice Thomas' argument is even more compelling when considered with the previous discussions, which suggest that partial birth abortion statutes only ban D&X abortion. *See supra* notes 101-11, 131-38, 200-12, 315 & 319 and accompanying text.

355. *See Stenberg*, 530 U.S. at 996-97.

356. *See id.* at 1013-14 (Thomas, J., dissenting) (stating that even if the regulation makes it marginally more difficult to obtain an abortion it is not an undue burden).

357. *See id.* at 1015.

358. *See supra* note 356 and accompanying text.

359. *See Casey*, 505 U.S. at 887-98; *see also supra* section II.B.1 for a complete discussion of *Casey*.

360. *See* Bopp & Cook, *supra* note 138, at 20-23 (stating that partial birth abortion statutes are clear enough that they ban only the D&X abortion procedure, leaving several other abortion methods available).

361. Stenberg v. Carhart, 530 U.S. 914, 1012 (2000) (Thomas, J., dissenting); *see also* Bordlee, *supra* note 119, at 68 (stating that a health exception as the majority states it would create a "hollow partial birth abortion ban").

362. *Stenberg*, 530 U.S. at 1019 (Thomas, J., dissenting) (quoting *Casey*, 505 U.S. at 880).

health exception is necessary if the procedure "has any comparative health benefits."[363] "In effect, no regulation of abortion procedures is permitted because there will always be some support for a procedure and there will always be some doctors who conclude that the procedure is preferable."[364]

Justice Thomas accurately shows how the majority has expanded the *Casey* standard.[365] If the majority had followed *Casey* the statute would have been upheld.[366] This statute, under *Casey*, does not need a health exception because it does not ban all types of abortion and, in fact, it leaves available the most common and safest method.[367] Therefore, if "continuing her pregnancy would constitute a threat to the woman" she could still terminate her pregnancy by using any abortion method other than D&X.[368] The majority has broadened the standard, making it impossible to regulate abortion. As Justice Thomas explains, "[a]ny doctor who wishes to perform such a procedure . . . will be able to do so with impunity."[369]

VI. CONCLUSION

The *Stenberg* opinion takes a step backwards in abortion jurisprudence.[370] Although the majority purports to follow the precedent set forth in *Casey*, it does not.[371] Instead, it in effect overrules *Casey*.[372] It does so by narrowing the states' interests in regulating abortion; by interpreting the statute to include the most common form of abortion; and by expanding the requirement of a health exception.[373] By doing this, the

---

363. *Id.* at 1012 (citing *Stenberg*, 530 U.S. at 937 (majority opinion)).

364. *Id.*; *see also* Bordlee, *supra* note 119, at 68; Myers, *supra* note 148, at C4 (stating that such a broad health exception renders the statute meaningless).

365. *See supra* notes 361-64 and accompanying text.

366. *See Stenberg*, 530 U.S. at 1017-18 (Thomas, J., dissenting); *see also supra* notes 339-42 and accompanying text.

367. *See* Radloff, *supra* note 227, at 1570; *see also* Reilly, *supra* note 148, at A7; *supra* notes 10, 117-18, 147-49, 178-79, 185-87 & 219-20 and accompanying text.

368. *Stenberg*, 530 U.S. at 1009-10 (Thomas, J., dissenting) (quoting *Casey*, 505 U.S. at 880); *see also* Andrews, *supra* note 238, at 526-29.

369. *Stenberg*, 530 U.S. at 1013 (Thomas, J., dissenting); *see also supra* notes 151-57.

370. *See Stenberg*, 530 U.S. at 982 (Thomas, J., dissenting); *cf. Stenberg*, 530 U.S. at 920-21 (majority opinion).

371. *Compare id.* at 930-31; *with Casey*, 505 U.S. at 833. See also *supra* notes 169-212 for a complete discussion of the majority holding of *Stenberg*, and its noncompliance with the *Casey* opinion.

372. *Compare Casey*, 505 U.S. at 878; *with Stenberg*, 530 U.S. at 914; *see also supra* notes 180-86, 189-99, 329, 361 & 364 and accompanying text.

373. *See Stenberg*, 530 U.S. 930-31, 937-38. See also *supra* notes 189-200 and

Court has made it virtually impossible for a state to regulate abortion,[374] and thus has put forth an unjust opinion.[375]

Under *Casey*, the states could further any interests in regulating abortion as long as those interests did not have the purpose or effect of preventing a woman from receiving an abortion.[376] *Casey* also specifically stated that a state has an interest in promoting fetal life at all stages of development.[377] Therefore, if the Court had followed *Casey*, then the Nebraska statute would have been upheld as it does not have the purpose or effect of preventing abortion, and it furthers several state interests, one of which is respect for fetal life.[378] However, even under the Court's narrower interpretation of what are valid state interests, the Nebraska statute should still be upheld.[379] The statute could still be upheld, because it protects the mother's health.[380]

The Court makes a further blunder when it reads the Nebraska statute to include the D&E abortion method, thereby finding the statute to pose an undue burden on the woman's right to an abortion.[381] If the statute did ban the D&E abortion method it would be an undue burden as it could prevent a woman from having an abortion.[382] The statute, however, bans "partial-birth abortion," on the premise that this term is only known to mean D&X

---

accompanying text, for a complete discussion and analysis of the majority's opinion concerning state interest; *supra* notes 200-12 and accompanying text, for a complete discussion and analysis of the majority's opinion concerning the statute's creation of an undue burden; and *supra* notes 173-86 and accompanying text for a complete discussion and analysis of the majority opinion concerning a health exception.

374. *See supra* note 367 and accompanying text.

375. With this opinion the Court has ignored the voices of thirty states, attempting to further legitimate interests which were in their means to protect under *Casey*. *See* Walther, *supra* note 7, at 706-07; *see also supra* note 57; *Casey*, 505 U.S. at 878.

376. *See Casey*, 505 U.S. at 878.

377. *Id.*

378. *See Brief of the Attorney General, supra* note 199, at 32. See also *supra* notes 194-99, 222, 227, 256-58 & 320 and accompanying text for discussions on Nebraska's interests in banning partial birth abortion.

379. The majority narrows valid interests to include only an interest in protecting the health of the mother, and in saving fetal life. *See Stenberg*, 530 U.S. at 930-31.

380. The D&X procedure is considered by many to be dangerous, and less safe than the more common D&E procedure, therefore, banning the procedure would protect the mother's health. *See* Reilly, *supra* note 148, at A7; Alder, *supra* note 117; Parker, *supra* note 10, at A10; Myers, *supra* note 148, at C4.

381. *See Stenberg*, 530 U.S. at 938-39.

382. *See Casey*, 505 U.S. at 887-98; see also *supra* section II.B.1. for a complete explanation of the holding in *Casey*.

abortion, and therefore cannot be read to include D&E abortion.[383] Even if the term "partial-birth abortion," were not commonly known to refer only to D&X abortion, the definition of it in the statute clearly only refers to D&X by its use of the word "deliver."[384] Therefore, since the statute leaves other abortion methods available, it is not an undue burden.

Finally, the Court expands the *Casey* requirement of a health exception to such a point that it overrules *Casey*.[385] The *Casey* health exception only required that a woman be able to have an abortion at any time during her pregnancy if her health was in danger.[386] Following this rule, the Nebraska statute would be upheld because it only bans one method of abortion; therefore, a woman may still receive an abortion if it is necessary.[387]

The majority overrules *Casey*, by holding that a health exception is necessary when a procedure has any form of benefit.[388] Under this rule, the Nebraska statute is meaningless, as a doctor can perform the procedure whenever he deems it would be better, regardless of what the medical profession as a whole would do in that situation.[389] Therefore, under the holding of this case, the states have lost their power to regulate abortion and further their interests.

Aimée M. Gauthier

---

383. *See* Hope Clinic v. Ryan, 195 F.3d 857, 867 (7th Cir. 1999) (stating that physicians are aware that "partial birth abortion" only refers to D&X abortion); *see also supra* notes 138, 206-08 and accompanying text.

384. *See* NEB. REV. STAT. §§ 28-326, 28-328 (1998); *see also supra* notes 134-38, 206-12 and accompanying text.

385. *Compare Casey*, 505 U.S. at 877; *with Stenberg*, 530 U.S. 930-31; *see also supra* notes 180-86, 189-99, 329-31 & 361-69 and accompanying text.

386. *See Casey*, 505 U.S. at 880.

387. *See* NEB. REV. STAT. §§ 28-326, 28-328 (1998).

388. *See Stenberg*, 530 U.S. at 938. This overrules *Casey*, because it takes the power to regulate abortion away from the states and gives it to the physician. *See* Bordlee, *supra* note 119, at 68; *see also* Myers, *supra* note 148, at C4.

389. *See* Bordlee, *supra* note 119, at 68; *see also* Myers, *supra* note 148, at C4. See *supra* notes 10, 117-18, 148-49, 178-79, 185-87 & 219-20 and accompanying text, for discussions on how the D&X abortion method is not beneficial to a woman seeking an abortion.

NOTE:

Reasoned Compulsory Licensing: Applying U.S. Antitrust's "Rule of Reason" to TRIP's Compulsory Licensing Provision

"Pharmaceutical companies would rather treat a bald American than a dying African."

-Richard Holbrooke
U.S. Ambassador to the United Nations[1]

"We recognize that AIDS is a major problem, but weakening intellectual property rights is not the solution."

-Alan F. Holmer
President of the Pharmaceutical Research and Manufacturers of America[2]

I. INTRODUCTION

Compulsory licensing is not a new concept. Nor is compulsory licensing found only in special fields of intellectual property. Many examples of compulsory licensing exist in the United States alone, touching the areas of high technology,[3] pharmaceuticals,[4] music[5] and more.[6]

---

1. Steve Sternberg, *AIDS Activists Discount Big Drugmakers' Gifts: Generic Producers Not Allowed to Fight Epidemic for Less*, USA TODAY, July 11, 2000 at 9D (convening the first National Security Council session to address a health issue, AIDS).

2. *Clinton, Drug Makers Move to Improve Access to Medicines*, AIDS POLICY AND LAW, May 26, 2000 (LRP Publications) at 8.

3. *See, e.g.*, *In re* Digital Equipment Corp., 2000 FTC LEXIS 28 (2000) (microprocessors); *In re* Intel Corp., 1999 FTC LEXIS 145 (1999) (advance technical infor-

Compulsory licensing allows a government to force the licensing of a patent to the non-patent-holders and, thus (by itself or through third parties), distribute the production of the patent to its citizens.[7] Although some payment may be made to the patent owner, compulsory licensing generally represents a modest return on invested capital and rarely includes the monopoly premium that a patent holder otherwise can receive by controlling the supply.[8]

Each sector has distinct reasons for using compulsory licensing. In the high tech sector, compulsory licensing can be used to prevent anti-competitive behavior in companies and thereby promote the economy.[9] In the pharmaceutical sector, compulsory licensing can be used similarly to promote the economy but foremost to protect people from deadly diseases in situations of national emergency.[10] Just as different circumstances may justify compulsory licensing, the mode of compulsory licensing, the arguments for and against it, and the results created are all very distinct.

This Note will analyze the viability of applying the "rule of reason" of U.S. antitrust cases to the more rare national emergency compulsory licensing of the World Trade Organization's Trade Related Aspects of Intellectual Property Agreement [hereinafter "TRIPS Agreement"].[11] As a focus, this Note will explore the impact compulsory licensing could have

---

mation); Garabet v. Autonomous Techs. Corp. and Summit Tech., Inc., 116 F. Supp. 2d 1159 (C.D. Cal. 2000) (laser vision correction).

4. *See, e.g.*, United States v. Bristol-Myers Co., 1982 U.S. Dist. LEXIS 11315 (D.C. Cir. 1982) (semisynthetic penicillin); United States v. Glaxo Group Ltd., 410 U.S. 52 (1973) (bulk-form fungicide griseofulvin).

5. *See, e.g.*, Recording Indus. Ass'n. of Am. v. Librarian of Congress, 176 F.3d 528 (D.C. Cir. 1999) (digital audio music); ABKCO, Inc. v. Stellar Records, Inc., 96 F.3d 60 (2nd Cir. 1996) (lyrics on screen).

6. *See, e.g.*, Pioneer Hi-Bred Int'l, Inc. v. J.E.M. Agric. Supply, Inc., 49 U.S.P.Q. 2d 1813 (N.D. Iowa 1998) (hybrid seed corn).

7. *See* Kevin Kennedy, *The Protection and Enforcement of Intellectual Property Rights*, *in* WORLD TRADE LAW 1081, 1113 (Raj Bhala & Kevin Kennedy eds., 1998); Alan M. Fisch, *Compulsory Licensing of Pharmaceutical Patents: An Unreasonable Solution to an Unfortunate Problem*, 34 JURIMETRICS J. 295, 300 (1994).

8. *See* A. Samuel Oddi, *The International Patent System and Third World Development: Reality or Myth?*, 1987 DUKE L.J. 831, 835-36 (1987); MICHAEL P. RYAN, KNOWLEDGE DIPLOMACY 26-31 (1998).

9. *See* Richard H. Marschall, *Patents, Antitrust, and the WTO/GATT: Using TRIPS as a Vehicle for Antitrust Harmonization*, 28 LAW & POL'Y INT'L BUS. 1165, 1170-71 (1997) (stating that "[u]nder the rule of reason, a particular practice is lawful only if its procompetitive effects outweigh its anticompetitive effects").

10. *See* NATIONAL RESEARCH COUNCIL, GLOBAL DIMENSIONS OF INTELLECTUAL PROPERTY RIGHTS IN SCIENCE AND TECHNOLOGY 164-65 (Mitchel B. Wallerstein et al. eds., 1993); BRIAN MCDONALD, THE WORLD TRADING SYSTEM 170 (1998).

11. *See infra* notes 279-313 and accompanying text.

on Acquired Immuno-Deficiency Syndrome [hereinafter "AIDS"] and Human Immunodeficiency Virus [hereinafter "HIV"].[12] AIDS/HIV is rapidly infecting millions of people and devastating populations.[13] Generally, AIDS drugs are not available to the poorest and sickest people because of the costs associated with the drugs.[14] Even when the expensive drugs are available, they cost more money per month than most people who need them will earn in one year.[15] To address the issue, some advocates insist that AIDS drugs be made available, through compulsory licensing, to those countries most affected by AIDS.[16]

A discussion of compulsory licensing is not complete without mentioning its many drawbacks, including the creation of a serious set of problems for the patent holders. One problem is that compulsory licensing may result in a decreased incentive to invent drugs prone to compulsory licensing, thereby decreasing the Research and Development budget for AIDS/HIV in certain companies.[17] There is also the possibility that more AIDS resistant viruses might develop, thereby making current AIDS drugs ineffective.[18] Another problem is that these drugs could overwhelm the infrastructure of developing countries and render the drugs useless.[19] There is also a fear that cheap drugs will be leaked into the high price markets.[20]

---

12. AIDS is caused by an infection of HIV. HIV destroys the immune system which protects humans from disease. *See* RICHARD SLOANE, THE SLOANE-DORLAND ANNOTATED MEDICAL-LEGAL DICTIONARY 497, 564 (Supp. 1992).

13. *See* UNAIDS, *HIV/AIDS Information and Data, available at* http://www.unaids.org/epidemic_update/report/Epi_report_chap_glo_estim.htm (last visited July 12, 2001).

14. *See, e.g.*, AIDS Coalition to Unleash Power, *Candidate Gore Zaps, at* www.actupny.org/actions/gorezaps.html (last visited July 12, 2001) (ACT UP is an activist group rallying for direct action to end the AIDS crisis).

15. *See id.*

16. For an interesting discussion of this proposition, see Mary T. Griffin, *AIDS Drugs & The Pharmaceutical Industry: A Need for Reform*, 17 AM. J.L. & MED. 363, 404 (1991) (suggesting direct price controls instead of compulsory licensing). *See also South Africa's PMA Resumes Litigation Against Health Minister*, MARKETLETTER (Marketletter Publications Ltd., London, U.K.), Aug. 21, 2000 (reporting that South African Health Minister is pursuing the use of compulsory licensing).

17. *Compare* Amy E. Carroll, Comment, *Not Always the Best Medicine: Biotechnology and the Global Impact of U.S. Patent Law*, 44 AM. U.L. REV. 2433, 2469 (1995), *with* Griffin, *supra* note 16, at 382-83.

18. *See* Ellie Tesher, *Cheap AIDS Drugs for Africa Ensnarled*, TORONTO STAR, Aug. 24, 2000, at A21.

19. *See* Duane Nash, *South Africa's Medicines and Related Substances Control Amendment Act of 1997*, 15 BERKELEY TECH. L.J. 485, 498-99 (2000) (citing Daniel Boden et al., *HIV-1 Drug Resistance in Newly Infected Individuals*, 282 J. AM. MED. ASS'N 1135 (1999); Susan J. Little et al., *Reduced Antiretroviral Drug Susceptibility Among Patients with Primary HIV Infection*, 282 J. AM. MED. ASS'N 1142 (1999)).

20. *See* Nash, *supra* note 19, at 501-02.

This Note will begin by presenting historical background materials, including definitions, international treaties and agreements, to set a framework for analysis of the issues presented by compulsory licensing.[21] Section II will also explore U.S. legislation and U.S. cases addressing compulsory licensing.[22] U.S. antitrust case law will be used as a model of anticompetitive compulsory licensing. Then, the South African legislation on compulsory licensing, Canada's history of compulsory licensing, and India's compulsory licensing laws will be examined.[23]

Section III will examine the AIDS crisis in those developing countries which have the highest poverty rates and currently have the highest AIDS rates.[24] These countries also typically have the weakest intellectual property rights laws.[25] High poverty and low intellectual property rights is a threatening combination for many pharmaceutical manufacturers. Section III will also analyze the arguments for and against compulsory licensing of pharmaceutical patents for anti-AIDS drugs, and the common ground between the arguments.[26] Finally, there will be a comparison of compulsory licensing as a remedy for AIDS and for antitrust.[27]

The Note's conclusion will suggest that the standard for antitrust compulsory licensing – the rule of reason – appropriately transfers to national emergency compulsory licensing.[28] Furthermore, the rule of reason in this situation would allow TRIPS members to examine whether compulsory licensing would alleviate or advance a national emergency.[29] If compulsory licensing would advance a national emergency, or in other words, create more problems than it would solve, then it should not be done. If, on the other hand, compulsory licensing helps to alleviate the national emergency, then it should be done. This Note does not conclude whether

---

21. *See infra* notes 30-74 and accompanying text.
22. *See infra* notes 75-144 and accompanying text.
23. *See infra* notes 145-80 and accompanying text.
24. *See infra* notes 181-202 and accompanying text.
25. *See* Robert M. Sherwood, *Why a Uniform Intellectual Property System Makes Sense for the World, in* NATIONAL RESEARCH COUNCIL, *supra* note 10, at 68-88 (discussing how a uniform system of intellectual property rights would be beneficial and rebutting arguments why developing countries refuse to raise intellectual property rights); Alan S. Gutterman, *The North-South Debate Regarding the Protection of Intellectual Property Rights*, 28 WAKE FOREST L. REV. 89, 92 (1993) (giving reasons developing countries are reluctant to provide strong intellectual property rights because providing these rights would work as a disincentive for local innovators, and because these rights would give foreign companies undue control over the protected items).
26. *See infra* notes 203-78 and accompanying text.
27. *See infra* notes 279-311 and accompanying text.
28. *See infra* notes 312-13 and accompanying text.
29. *See id.*

any individual country should be allowed compulsory licensing, but merely offers an alternative method to examining the factors and weighing them carefully.

II. BACKGROUND

1. Compulsory Licensing of Patents Under the Agreement on Trade-Related Aspects of Intellectual Property Rights: How the World Got to Where It Is Now.

Patents[30] are an important part of intellectual property. Dating back to the time of the Venetians, before the eighteenth century, development and invention have been encouraged and rewarded by giving creators limited monopoly control over profits to be made.[31] The aim of U.S. intellectual property policy since George Washington has been to promote the public interest to which "private property rights have been the means toward that end."[32]

Early in the development of patent rights, distinctions between different nations' patent laws were the impetus for international agreements concerning the protection of intellectual property.[33] Prior to international agreements, inventors had to disclose their inventions to other countries upon patenting in their own country, and then had to follow difficult procedural rules, pay taxes, and protect their use in those foreign countries.[34] Patentees were forced to maintain patent agents in each country to protect their interests.[35]

Frequently, patentees were artisans who were drawn to a foreign nation to teach their trade to foreign nationals who were, for the patent period, prevented from competing with their teacher.[36] These patentees were some of the first and most outspoken proponents of international agreements in patent protection and thus, international conventions on patent protection arose.[37]

---

30. "A grant of right to exclude others from making, using or selling one's invention and includes right to license others to make, use or sell it." BLACK'S LAW DICTIONARY 1125 (6th ed. 1990).

31. *See* RYAN, *supra* note 8, at 6; NATIONAL RESEARCH COUNCIL, *supra* note 10, at 46.

32. RYAN, *supra* note 8, at 6.

33. *See* Edith Tilton Penrose, *The Economics of the International Patent System*, *in* THE INTERNATIONAL INTELLECTUAL PROPERTY SYSTEM: COMMENTARY AND MATERIALS PART ONE 635, 635-36 (Frederick Abbott et al., 1999).

34. *See id.*

35. *See id.* at 636.

36. *See* NAT'L RESEARCH COUNCIL, *supra* note 10, at 45-46 (discussing the granting of early English patents for 14 years, or two generations of apprentices).

37. *See* Penrose, *supra* note 33, at 636.

The first international patent convention was the Vienna Congress of 1873.[38] The resolutions from the Vienna Congress were general but began the heated debate over patent rights and "clearly endorsed the principle of patent protection."[39] The issue of compulsory licensing was broached at this conference, and compulsory licensing was adopted "in cases in which the public interest should require it;" however, it was not a popularly accepted resolution.[40] The Convention's protection of intellectual property rights is now generally regarded as a feeble attempt at protection because the Convention only prescribed a minimum set of standards.[41]

Five years later, the Paris Conference convened to solidify and establish the basis of an international agreement on patent protection.[42] Although the patentee's point of view dominated at this conference, there were lively debates over the ability of governments to compel licenses.[43] The debate focused on whether a proposal to subject "natural property" of inventors to expropriation was necessary.[44] Some believed that subjecting natural property to the government's expropriation was not necessary because, if the government needed to take property, it certainly had the power to override private interests.[45] The proposal, however, was adopted for the sake of making the principle explicit.[46] The Convention was signed on March 20, 1883,[47] and in 1884 the International Union for the Protection of Industrial

---

38. *See id.* at 637.

39. *Id.*

40. *See id.*

41. *See* Robert J. Pechman, *Seeking Multilateral Protection for Intellectual Property: The United States "TRIPs" over Special 301*, 7 MINN. J. GLOBAL TRADE 179, 181 (1998).

42. *See* Penrose, *supra* note 33, at 638.

43. *See id.* at 639-40.

44. *See id.* at 640 (defining "natural property" as the "right of inventors and of industrial creators in their work" and the "most sacred of all property"). Decades earlier, Thomas Jefferson wrote:

> If nature has made any one thing less susceptible than all others of exclusive property, it is the action of the thinking power called an idea, which an individual may exclusively possess as long as he keeps it to himself; but the moment it is divulged, it forces itself into the possession of everyone . . . . Inventions then cannot, in nature, be a subject of property.

NAT'L RESEARCH COUNCIL, *supra* note 10, at 26.

45. *See* Penrose, *supra* note 33, at 640.

46. *See id.*

47. *See* Laurinda L. Hicks & James R. Holbein, *Convergence of National Intellectual Property Norms in International Trading Agreements*, 12 AM. U.J. INT'L L. & POL'Y 769, 778 (1997).

Property came into existence.[48]

The General Agreement on Trade and Tariffs [hereinafter "GATT"] was created in the 1940s to centralize international trade issues.[49] The rounds, or trade negotiations, held under GATT in Uruguay between 1986 and 1994 expanded the idea for GATT into a separate and viable organization called the World Trade Organization [hereinafter "WTO"].[50] While GATT deals exclusively with trade in goods, the WTO deals with trade in services and trade-related intellectual property and investment issues.[51] On December 15, 1994, the WTO created the Agreement on Trade-Related Aspects of Intellectual Property Rights,[52] which came into effect on January 1, 1995.[53]

Section 5 of the TRIPS Agreement governs patent licenses pursuant to international law.[54] Section 5, article 27 of the TRIPS Agreement states that "patents [shall be] available for any inventions, whether products or processes, in all fields of technology, without discrimination, subject to the normal tests of novelty, inventiveness and industrial applicability."[55] The TRIPS Agreement has not created a single, universal patent system, rather it sets forth ground rules for member countries.[56]

At the outset, the TRIPS Agreement states that its purpose is to "reduce distortions and impediments to international trade, and taking into account the need to promote effective and adequate protection of intellectual property rights, and to ensure that measure and procedures to enforce intellectual property rights do not themselves become barriers to legitimate

---

48. *See* Paris Convention for the Protection of Industrial Property, Mar. 20, 1883, as last revised July 14, 1967, 21 U.S.T. 1583, 828 U.N.T.S. 305 [hereinafter "Paris Convention"]; *see also* Penrose, *supra* note 33, at 644-45 (discussing the addition of twenty-nine other countries to the Union after the original eleven signers).

49. *See* MCDONALD, *supra* note 10, at 41.

50. *See id.* (effective January 1, 1995); *see also* WTO, *The Multilateral Trading System – Past, Present and Future*, *available at* http://www.wto.org/english/thewto_e/whatis_e/inbrief_e/inbr01_e.htm (last visited Aug. 5, 2001). The website also provides a brief history of the WTO. *See id.*

51. *See* MCDONALD, *supra* note 10, at 41.

52. *See Agreement on Trade-Related Aspects of Intellectual Property Rights, Including Trade in Counterfeit Goods*, 33 I.L.M. 81 (1994) [hereinafter "TRIPS Agreement"].

53. *See* WTO *Overview: The TRIPS Agreement*, at http://www.wto.org/english/tratop_e/trips_e/intel2_e.htm (last visited July 12, 2001).

54. *See id.*

55. *Id.*

56. *See The Right to Good Ideas: How Patents Help the Poor: Intellectual – Property Rights Are Not Just for the Rich World. Carefully Constructed, They Can Help the Poorest Too*, THE ECONOMIST (2001) June 23, 2001, at 21 [hereinafter "*The Right to Good Ideas*"].

trade."[57] Furthermore, Article 7 defines the objective of the TRIPS Agreement: as "contribut[ing] to the promotion of technological innovation . . . in a manner conducive to social and economic welfare, and [] a balanc[ing] of rights and obligations."[58] The overall goal is clearly the protection of intellectual property rights.[59] However, the principles in Article 8 define situations in which members of the TRIPS Agreement may "adopt measures necessary to protect . . . and promote the public interest," as long as the measures are consistent with the Agreement.[60]

Patent licenses are specifically subject to compulsory licensing by the TRIPS Agreement even though the Agreement does not use the words "compulsory license."[61] Article 28 of the TRIPS Agreement describes the rights conferred by a patent, and Articles 30 and 31 describe the exceptions to those rights conferred by a patent.[62] Article 30, entitled "*Exceptions to Rights Inferred*," is a general list of exceptions to patent holders' rights.[63] Article 31 allows for other exceptions, and is entitled "*Other Use Without Authorization of the Right Holder*," and is known outside of the Agreement as the provision that allows compulsory licensing.[64] The words of Article 31 unmistakably permit compulsory licensing: "[A]llows for other use . . . without the authorization of the right holder, including use by the government or third parties authorized by the government."[65] Throughout this Note, "*Other Use Without Authorization of the Right Holder*," will be referred to as "compulsory licensing."

Because the patent can be used without the permission of the patent holder, compulsory licensing promotes a public interest that is distinct from the protection of property rights.[66] Compulsory licensing in TRIPS is meant to promote the public interest in protecting health and nutrition, and

---

57. TRIPS Agreement, *supra* note 52, at 84; *see also* Carlos M. Correa, *Patent Rights*, *in* INTELLECTUAL PROPERTY AND INTERNATIONAL TRADE: THE TRIPS AGREEMENT 189, 209 (Carlos M. Correa & Abdulqawi A. Yusef eds., 1998).

58. TRIPS Agreement, *supra* note 52, at 86-87.

59. *See id.* at 86.

60. *See id.* at 87 (addressing the public interest in paragraph one, and the prevention of abuse of intellectual property rights which restrains trade in paragraph two). There are 141 member-countries of the TRIPS Agreement. *See The Right to Good Ideas*, *supra* note 56.

61. *See infra* notes 62-65 and accompanying text; *see also* Correa, *supra* note 57, at 208.

62. *See* TRIPS Agreement, *supra* note 52, at 94-95.

63. *See id.* at 95.

64. *See id.* at 95-96.

65. *See id.*; *see also* DANIEL GERVAIS, THE TRIPS AGREEMENT: DRAFTING HISTORY AND ANALYSIS 165 (1998) (suggesting some alternative terms: compulsory licensing, non-voluntary licensing or licensing as of right).

66. *See* TRIPS Agreement, *supra* note 52.

promoting the public interest in areas of vital importance to national socio-economic and technological development.[67]

Article 31 provides five grounds for granting compulsory licenses: (1) refusal to deal; (2) national emergency or extreme urgency; (3) anti-competitive practices; (4) non-commercial use; and (5) dependent patents.[68] This Note is centered around and compares compulsory licensing based on a national emergency or extreme urgency from the TRIPS Agreement, and anti-competitive practices, as discussed through the United States' antitrust laws. Subsections (b) and (k) of Article 31 deal respectively with these two exceptions.[69]

Subsection (b) provides that one who attempts to use a patent without authorization of the patent holder must first make efforts to obtain authorization, but "[t]his requirement may be waived by a Member in the case of a national emergency or other circumstances of extreme urgency . . . . [T]he right holder shall, nevertheless, be notified as soon as reasonably practicable."[70] Subsection (k), however, provides that efforts to obtain authorization are not required "where [compulsory licensing] is permitted to remedy a practice determined after judicial or administrative process to be anti-competitive."[71] In other words, compulsory licensing "enables the government granting the patent to force the patentee to license the invention if the government does not approve of the patent's use[, and c]onsequently, another individual or company is allowed to make and sell the invention."[72]

Many conditions must be fulfilled before compulsory licenses will be granted, such as: (1) review on individual merits of compulsory license; (2) efforts previously made to obtain authorization; (3) license limited to purpose for which it was authorized; (4) non-exclusive; (5) non-assignable; (6) predominant authorization for domestic market; (7) revocation; (8) remuneration; and (9) judicial review of remuneration.[73]

The fourth Ministerial Conference for the WTO in Doha, Qatar, in November 2001, issued a declaration in order to elucidate the TRIPS Agreement. The declaration stated: "[T]he TRIPS Agreement does not and

---

67. *See id.* art. 8.

68. *See* Correa, *supra* note 57, at 210 (summarizing TRIPS Agreement provisions of article 31).

69. *See* TRIPS Agreement, *supra* note 52.

70. *Id.*

71. *Id.*

72. Cole M. Fauver, *Compulsory Patent Licensing in the United States: An Idea Whose Time Has Come*, 8 NW. J. INT'L L. & BUS. 666, 667 (1988).

73. *See id.* at 668; *see also* Correa, *supra* note 57, at 210 (indicating in table form the conditions to which compulsory licenses should be subjected under TRIPS Agreement).

should not prevent Members from taking measures to protect public health . . . . [T]he Agreement can and should be interpreted and implemented in a manner supportive of WTO Members' right to protect public health and, in particular, to promote access to medicines for all."[74]

2. Various Approaches to Compulsory Licensing

a. United States Legislation and Policy Regarding Compulsory Licensing

In May, 2000, President Clinton signed the African Growth and Opportunity Act.[75] The Act's aim was to foster the mutual interest of the United States and the countries of sub-Saharan Africa to promote stable and sustainable economic growth and development in sub-Saharan Africa.[76] The law states that sub-Saharan African nations are eligible for trade benefits if the President determines:

> that the country—(1) has established, or is making continual progress toward establishing—(A) a market-based economy that protects private property rights . . . (C) the elimination of barriers to United States trade and investment including by—(i) the provision of national treatment and measures to create an environment conducive to domestic and foreign investment; (ii) the protection of intellectual property . . . .[77]

---

74. WTO, *Declaration on the TRIPS Agreement and Public Health* at http://www.wto.org/english/thewto_e/minist_e/min01_e/mindecl_trips_e.htm (last visited Dec. 6, 2001). The declaration, although hailed as new hope for access to AIDS drugs for poor countries, has been received with mixed emotions, and dashed hopes. *See* Geoff Dyer, *Activists Point to Flaws in Declaration on Drug Patents*, FINANCIAL TIMES, Nov. 16, 2001, *available at* http:// globalarchive.ft.com/globalarchive/article.html?id=011116000410&query=Geoff+Dyer (last visited Apr. 10, 2002); Jayanta Ghosh, *No Gains From Doha, Say Pharma Firms*, TIMES OF INDIA, Nov. 27, 2001, *available at* http:// timesofindia.indiatimes.com/articleshow.asp?art_id=569873751&sType=1 (last visited Apr. 10, 2002).

75. *See* Mark Felsenthal, *Legislation: Clinton Signs Africa-Caribbean Trade Bill, Sees Momentum Building for China Measure*, INTERNATIONAL TRADE DAILY (The Bureau of National Affairs, Inc.), May 19, 2000.

76. *See* African Growth and Opportunity Act of 2000, Pub. L. No. 106-200, 114 Stat. 251 (May 18, 2000) (Section 104-Eligibility Requirements).

77. *Id.*; *see also UN and Big Pharma Join to Improve HIV/AIDS Care in Developing Countries*, MARKETLETTER, May 22, 2000, *available in* 2000 WL 7541832 (U.N. Secretary-General Kofi Annan stating that this "is a positive and really strong indication to the patients they are not being abandoned.").

The available trade benefits, with regard to AIDS and HIV, are that the law encourages businesses and legislators to meet the needs of the AIDS crisis in sub-Saharan Africa because that is the focus of U.S. foreign policy in that region.[78] Although this executive order may be in jeopardy under the new administration of President Bush,[79] Secretary of State, Colin L. Powell referred to the Act during his visit to South Africa in May 2001 as beneficial, encouraging free trade between Africa and America.[80]

There were also bills which preceded this Act that were never made into law. These bills discussed the compulsory licensing of patented inventions relating to health and illustrate divergent views on this topic.[81] A Democrat-sponsored bill provided for compulsory licensing in much the same way as Article 31 does in the TRIPS Agreement.[82] However, a Republican-sponsored bill stated:

> Whereas . . . [the] pharmaceutical industry estimates its losses due to piracy at $5,000,000,000 per year . . . [and] deficiencies in the TRIPS Agreement will delay realization, by United States industries that are dependent on the protection of intellectual property rights . . . [be it r]esolved . . . that the United States should pursue further strengthened levels of intellectual property protection, even after the TRIPS Agreement is implemented.[83]

Needless to say, United States policy toward the protection of patents, and the availability of patents in urgent situations, is not uniformly supported.[84]

The United States government can, however, use a patent without negotiating for its use through eminent domain. The patentee's sole remedy is

---

78. *See* Pub. L. No. 106-200, §§ 128, 129(b)(3) (2000).

79. *See* Salih Booker, *Bush Medicine for Africa: George W. Bush's Foreign Policy Toward Africa*, THE NATION, February 26, 2001.

80. *See* John Murphy, *Powell Tells S. African Students U.S. Will Be a Partner for Peace*, BALT. SUN, May 26, 2001.

81. *See* H.R. 2927, 106th Cong. (1999) *and* H. Con. Res. 240, 103rd Cong. (1994).

82. *See* H.R. 2927, 106th Cong. (1999) (using the words "compulsory licensing" and allowing the Secretary of Health and Human Services to "establish other use of the subject matter of the patent without authorization of the right holder if the Secretary determines . . . [it] necessary to alleviate health or safety needs" not adequately satisfied by the right holder).

83. H. Con. Res. 240, 103rd Cong. (1994) (expressing the "sense of Congress" in regards to intellectual property protection).

84. *Compare* H. Con. Res. 240, 103rd Cong. (1994) *with* H.R. 2927, 106th Cong. (1999). For information regarding compulsory licensing after the September 11, 2001 terrorist attack, see *infra* note 226 and accompanying text.

recovery under title 28, section 1498 of the United States Code.[85] This particular use of a patent is considered a taking under the Fifth Amendment of the U.S. Constitution.[86] Section 1498 also provides that if the government uses a patent without negotiating for its use, then it must make "reasonable and entire compensation for such use."[87] The United States also allows compulsory licensing under the Atomic Energy Act,[88] the Clean Air Act,[89] and the Plant Variety Protection Act;[90] all of these acts are compliant with the TRIPS provisions.[91]

While TRIPS was being discussed and debated in the Uruguay Rounds, several countries proposed certain objectives and requirements for TRIPS.[92] The United States' objectives included limiting impediments to trade in goods and services, and as a means to this end, the U.S. proposal included standards also found in its domestic laws.[93] The United States discouraged compulsory licensing in its proposal and encouraged longer patent terms, especially from developing countries.[94]

In the U.S. antitrust sector, the Sherman Act is extremely important.[95] The Sherman Act prohibits "[e]very contract, combination in the form of trust or otherwise, or conspiracy, in restraint of trade or commerce among

---

85. *See* 28 U.S.C. § 1498(a) (1994).

86. *See* Hughes Aircraft Co. v. United States, 86 F.3d 1566, 1571 (Fed. Cir. 1996).

87. 28 U.S.C. § 1498(a) (1994).

88. 42 U.S.C. §§ 2011-2297 (1994). Atomic energy is "subject at all times to the paramount objective of making the maximum contribution to the common defense and security." *Id.* § 2011. *See also* RAJ BHALA & KEVIN KENNEDY, WORLD TRADE LAW: THE GATT-WTO SYSTEM, REGIONAL ARRANGEMENTS, AND U.S. LAW 1113 (1998).

89. 42 U.S.C. §§ 7401-7671 (1994). Air pollution prevention is "the reduction or elimination, through any measures, of the amount of pollutants produced or created at the source." *Id.* § 7401; *see also* BHALA & KENNEDY, *supra* note 88, at 1113.

90. 7 U.S.C. §§ 2321-2583 (2000) "The Secretary may declare a protected variety open to use on a basis of equitable remuneration to the owner, not less than a reasonable royalty, when the Secretary determines that such declaration is necessary in order to insure an adequate supply of fiber, food, or feed in this country . . ."; *Id.* § 2404; *see also* Fisch, *supra* note 7, at 301 n.40.

91. *See* BHALA & KENNEDY, *supra* note 88, at 1113.

92. *See* Michael L. Doane, *TRIPS and International Intellectual Property Protection in an Age of Advancing Technology*, 9 AM. U.J. INT'L L. & POL'Y 465, 474 (1994) (naming Japan, the European Community and the United States).

93. *See id.*

94. *See id.* Currently, the U.S. patent term for pharmaceutical patents is twenty years which is the same as under TRIPS. *See* 35 U.S.C. § 154 (1998) *and* TRIPS Agreement, art. 33. Canada has recently changed their patent term to twenty years from a seventeen year term in 1989. *See* Mota, *infra* note 162, at 547. India has a patent term of ten years. *See Compulsory License Likely in Patents Amendment Bill*, CHEMICAL BUSINESS NEWSBASE, Jan. 22, 2001.

95. *See* 15 U.S.C. § 1 (1994 & Supp. 1999).

the several States, or with foreign nations."[96] The Sherman Act has a long history in the United States, originating on July 2, 1890.[97] The Sherman Act also declares the violation of the Act a felony, and prescribes the punishment of fines, or imprisonment, or both.[98] The Sherman Act has been applied in antitrust cases ranging from the bulk sale of drugs[99] to laser eye surgery.[100] The Sherman Act applies not because a company is making too much money, but because it is abusing its monopoly at the expense of the consumer and at the expense of further innovation.[101]

b. United States Cases on Compulsory Licensing

The U.S. Supreme Court decided, in *United States v. Glaxo Group Ltd.*,[102] that despite the defendant's defense of non-use of the patent, the government could challenge the validity of the patent because the public's interest in free competition prevailed.[103] The Court also stated that since mandatory sales granted as relief by the court below could be overcome by the plaintiff, the plaintiff should also be compelled to license the patent.[104] The Court referred to *Besser Mfg. Co. v. United States*,[105] *International Salt Co. v. United States*,[106] and *Hartford-Empire Co. v. United States*,[107] in its conclusion stating that "compulsory patent licensing at reasonable charges are recognized antitrust remedies."[108]

Other Supreme Court cases, some of which preceded the above cases by one hundred years, probably set the stage for the approval of compulsory

---

96. *Id.*

97. *See id.*

98. *See id.*

99. *See* United States v. Glaxo Group Ltd., 410 U.S. 52 (1973).

100. *See* Garabet v. Autonomous Technologies Corp., 116 F. Supp. 2d 1159 (C.D. Cal. 2000).

101. *See* B. Zorina Khan, *Symposium on Antitrust and Intellectual Property: Federal Antitrust Agencies and Public Policy Toward Antitrust and Intellectual Property*, 9 CORNELL J.L. & PUB. POL'Y 133, 152 (1999) (according to the 1995 Department of Justice and the Federal Trade Commission joint issue of Guidelines for the Licensing of Intellectual Property).

102. 410 U.S. 52 (1973).

103. *See id.* at 57-58.

104. *See id.* at 63-64.

105. 343 U.S. 444, 449 (1952) (recognizing compulsory licensing and sale of patented devices as antitrust remedies, particularly appropriate in this concrete block-making case, where the Court found a "penchant for abuses of patent rights").

106. 332 U.S. 392 (1947) (directing defendant to sell, license or lease the salt machines under the patent to any buyer on non-discriminatory terms).

107. 323 U.S. 386 (1945) (restricting the lower court's decision to make appellant forfeit their patents, permitting compulsory licensing on reasonable grounds; for example, with reasonable royalty fees for only current or modified patents at issue).

108. *Glaxo Group Ltd.*, 410 U.S. at 64.

licensing.[109] Cases using the "rule of reason" standard to determine whether the Sherman Act had been violated also helped set the stage for compulsory licensing.[110] The rule of reason

> has been used to give the Act both flexibility and definition, and its central principle of antitrust analysis has remained constant. Contrary to its name, the Rule does not open the field of antitrust inquiry to any argument in favor of a challenged restraint that may fall within the realm of reason. Instead, it focuses directly on the challenged restraint's impact on competitive conditions.[111]

The Supreme Court first expounded this standard of analysis in *Standard Oil v. United States*[112] where it became accepted that "[o]nly unreasonable restraints of trade should be illegal."[113]

In *Professional Engineers*, the Court denied using the "rule of reason" to determine whether a professional organization's agreement not to take competitive bidding contracts was violative of the Sherman Act.[114] The Court stated that it "has adhered to the position that the inquiry mandated by the Rule of Reason is whether the challenged agreement is one that promotes competition or one that suppresses competition."[115] The Court found that although the professional organization wanted to apply the rule of reason, they could not because their agreement was on its face unlawful.[116] The Court reiterated that the rule of reason does not allow an evaluation of the facts if the agreement is per se illegal.[117]

---

109. *See* Khan, *supra* note 101, at 133 n.2 (1999) (citing Woodworth v. Edwards, 30 F. Cas. 567 (C.C.D. Me. 1847) (No. 18,014); and Kendall v. Winsor, 62 U.S. 322, 329 (21 How. 322, 329) (1859) ("Whilst the remuneration of genius and useful ingenuity is a duty incumbent upon the public, the rights and welfare of the community must be fairly dealt with and effectually guarded.")).

110. *See, e.g.*, Arizona v. Maricopa County Medical Society, 457 U.S. 332 (1982); Continental T.V., Inc. v. GTE Sylvania Inc., 433 U.S. 36 (1977); Goldfarb v. Virginia State Bar, 421 U.S. 773 (1975); United States v. Topco Assoc., Inc., 405 U.S. 596 (1972); Standard Oil Co. v. United States, 221 U.S. 1 (1911).

111. National Soc. of Prof'l Eng'rs v. United States, 435 U.S. 679, 688 (1978).

112. 221 U.S. 1, 60, 66 (1911).

113. Thomas A. Piraino, Jr., *Reconciling the Per Se and Rule of Reason Approaches to Antitrust Analysis*, 64 S. CAL. L. REV. 685, 689 (1991).

114. *See National Soc. of Prof'l Eng'rs*, 435 U.S. at 691-95.

115. *Id.* at 691. The Court explained the mechanics of the rule of reason: "Under this rule, the factfinder weighs all of the circumstances of a case in deciding whether a restrictive practice should be prohibited as imposing an unreasonable restraint on competition." *Id.* at 691 n.17 (quoting Continental T.V. v. GTE Sylvania, 433 U.S. 36, 49 (1977)).

116. *See id.* at 686, 699.

117. *See id.* at 689-90 (finding that precedence has foreclosed queries of whether

If the agreement in restraint of trade is *not* per se illegal, however, then the rule of reason is applicable.[118] Under the rule of reason, the Court should balance the many factors peculiar to the "business to which restraint is applied."[119] Those factors include the condition of the business before and after the agreement, the nature of the restraint, the effect of the restraint (probable or actual), the history of the restraint, and even the intent of the agreement's makers.[120] In *Continental T.V.*, the Court did not find anti-competitive behavior *per se*, but did find illegal behavior after using the rule of reason to analyze the whole situation.[121]

The TRIPS Agreement authorizes compulsory licenses to be granted in order to remedy anti-competitive practices.[122] It is clear that the U.S. cases that recognize compulsory licensing as an antitrust remedy[123] do not do so under the TRIPS Agreement, but rather under U.S. laws.[124] Nevertheless, the remedy is available for disagreements falling under the TRIPS Agreement.[125] Under the TRIPS Agreement, remedying anti-competitive practice with compulsory licensing is subject to special provisions for remuneration to the patent holder.[126] In the United States, compulsory licensing related to antitrust cases does, in fact, remedy anti-competitive practices by providing remuneration to the aggrieved non-patentee.[127]

---

certain monopolistic agreements promote trade and commerce better than competition).

118. *See* Continental T.V., Inc. v. GTE Sylvania Inc., 433 U.S. 36, 49 (1977).

119. *Id.* at 49 n.15 (quoting Justice Brandeis from his opinion in Chicago Bd. of Trade v. United States, 246 U.S. 231, 238 (1918)).

120. *See id.*

121. *See id.* at 59 (affirming the decision that the restriction of sales to franchisee's location was reasonable).

122. *See* TRIPS Agreement, *supra* note 52, sec. 5, art. 31(k).

123. See, for example, cases that contemplate or actually use compulsory licensing as a legitimate remedy to antitrust violations: *In re* Digital Equipment Corp., 2000 F.T.C. LEXIS 28 (2000) (compulsory licensing had the effect of preventing delay of production, impediment to innovations, and undermining of competitiveness); Garabet v. Autonomous Techs. Corp. and Summit Tech., Inc., 116 F. Supp. 2d 1159 (C.D. Cal. 2000); United States v. Glaxo Group Ltd., 410 U.S. 52 (1973).

124. The U.S. laws discussed in the cases *supra* note 123 are the Federal Trade Commission Act, 15 U.S.C. § 45 (2000); Clayton Act, 15 U.S.C. § 18 (2000); and Sherman Act, 15 U.S.C. § 1 (2000).

125. *See* TRIPS Agreement, *supra* note 52, sec. 5, art. 31(k). Some critics of the TRIPS Agreement's nebulousness observe that "competition safeguards" are available through the TRIPS Agreement, but do not *require* the use of anti-competition laws. *See* Marschall, *supra* note 9, at 1187-91.

126. *See* TRIPS Agreement, *supra* note 52, *and* CORREA, *supra* note 57, at 211 (assuming remuneration is used both as a remedy and a sanction against the abuse of conferred rights of patent holders).

127. *See, e.g.*, *In re* Intel Corp., No. 9288, 1999 FTC LEXIS 145 (Aug. 3, 1999); Image Tech. Serv. v. Eastman Kodak Co., No. C 87-1686 AWT 1996 U.S. Dist.

Many recent antitrust cases have further examined the issue of compulsory licensing.[128] In the 1999 case, *In re Intel*,[129] the Federal Trade Commission [hereinafter "FTC"] found that Intel was a monopolist, "engaged in 'conduct, other than competition on the merits or restraints reasonably "necessary" to competition on the merits, that reasonably appears capable of making a significant contribution to creating or maintaining monopoly power.'"[130] Despite one Commissioner's disagreement with the majority that the conduct at issue would *not* "appreciably affect—much less stem—the immense tide of invention and improvement that continuously drives this industry,"[131] the FTC placed strict orders on Intel Corporation.[132] The orders stated that Intel must "cease and desist" refusing access by any microprocessor customer to Advance Technology Information.[133] The Order allowed Intel to seek equitable remedies, but a dispute regarding remedies would not affect Intel's obligation to provide the Advance Technology Information.[134]

In *Image Technological Services*,[135] a California court based its decision regarding whether the plaintiff suffered anti-competitive injury on (among other case law) Supreme Court antitrust cases from the 1950s.[136] The court found injury in that case because there was "undeniabl[e] harm to the competitive market system, which the antitrust laws were designed to protect."[137] Furthermore, the court stated that "[c]onsistent with this conclusion, the Supreme Court has on more than one occasion recognized and approved compulsory licensing and sale of patented devices as remedies in antitrust cases."[138] This case falls under the third ground for granting com-

---

LEXIS 2386 (N.D.Cal. 1996); *and* Garabet v. Autonomous Techs. Corp. and Summit Tech., Inc., 116 F. Supp. 2d 1159, 1174-75 (C.D. Cal. 2000).

128. *See supra* note 127.

129. 1999 FTC LEXIS 145.

130. *Id.* (quoting Barry Wright Corp. v. ITT Grinnell Corp., 724 F.2d 227, 230 (1st Cir. 1983) (citation omitted)).

131. *In re Intel Corp.*, 1999 FTC LEXIS 145, at *6.

132. *See id.* at *10-16.

133. *See id.* at *10-11. "'Advance Technical Information' or 'AT Information' means confidential product information regarding a general purpose Intel microprocessor [needed to] enable a customer to [use and incorporate the information] in time for introduction into the [market] by the official release date for such microprocessors . . . ." *Id.* at *9.

134. *See In re Intel Corp.*, 1999 FTC LEXIS 145, at *12.

135. Image Tech. Serv. v. Eastman Kodak Co., 1996 U.S. Dist. LEXIS 2386, at *6-7 (N.D. Cal. 1996).

136. *See id.* (citing United States v. U.S. Gypsum Co., 340 U.S. 76, 94 (1950); Besser Mfg. Co. v. United States, 343 U.S. 444, 449 (1952)).

137. *Image Tech. Serv.*, 1996 U.S. Dist. LEXIS 2386, at *6 n.3.

138. *Id.* at *7. *But see* Tonya Trumm, Comment, *Expansion of the Compulsory Licensing Doctrine?* Image Technical Services, Inc. v. Eastman Kodak Co., 24 IOWA

pulsory licenses according to the TRIPS Agreement: anti-competitive practices.[139] It also meets some of the conditions laid out for the granting of compulsory licenses; for example, it has been reviewed on its individual merits for compulsory licensing, and remuneration was demanded.[140]

In *Garabet v. Autonomous Techs. Corp. and Summit Tech., Inc.*,[141] the plaintiffs challenged a merger of two vision correction corporations, which they asserted were restraints of trade and a monopolization in violation of the Sherman Act and the Clayton Act.[142] In the plaintiff's claim for a remedy of compulsory licensing, the court denied it and reiterated the general respect for compulsory licensing.[143] "Compulsory licensing is a very rare remedy, and is typically only ordered in cases of patent withholding, misuse, or use of a patent to monopolize a market."[144]

c. South Africa's Legislation on Compulsory Licensing: Focus on Pharmaceuticals.

Although the United States has proposed legislation that directly addresses the compulsory licensing issues, it has not been adopted. The United States, nevertheless, is not the only country whose legislative attempts at compulsory licensing have failed. In South Africa, a recently resubmitted amendment to its Intellectual Property Laws brought great attention to the compulsory licensing issue.[145] The amendment proposed to allow South Africa more leeway for compulsory licensing than under the TRIPS Agreement.[146] The amendment, called the Medicines and Related Substances Control Act,[147] faced strong opposition from the Pharma-

---

J. CORP. L. 157, 165-66 (1998) (protesting the result in *Image Technical Services* because it fails to consider the possible detrimental effects that compulsory licensing would have on the patent system).

139. *See* TRIPS Agreement, *supra* note 52, sec. 5.

140. *See also Image Tech. Serv.*, 1996 U.S. Dist. LEXIS 2386, at *15-16 (injunction forcing Kodak to sell its parts to Independent Service Organizations as it does to other customers under existing policies). *See generally* TRIPS Agreement, *supra* note 52.

141. 116 F. Supp. 2d 1159 (C.D. Cal. 2000).

142. *See id.* at 1160.

143. *See id.* at 1174.

144. *See id.* at 1174 n.16 (holding in the *Garabet* case that because there was no claim of patent misuse, the court is reluctant to give this antitrust remedy for simply the possession of patent).

145. *See South Africa's PMA Resumes Litigation Against Health Minister*, MARKETLETTER (Marketletter Publications Ltd., London, U.K.), Aug. 21, 2000.

146. *See* BSRSA 2000 Synopsis of Intellectual Property Laws Amendment Act 38 of 1997.

147. *See* Medicines and Related Substances Control Amendment Bill, [B72-97] at http://www.gov.za/bills/1997/b72a-97.pdf (last visited Aug. 5, 2001).

ceutical Manufacturers' Association [hereinafter "PharMA"], who filed a motion requesting that the legislation be declared unconstitutional.[148]

The forty-two pharmaceutical companies[149] listed on the motion requested specifically that the provision relating to cheaper and more available antiretroviral agents[150] be found unconstitutional.[151] This particular provision would allow the South African Minister of Health to engage in compulsory licensing.[152] Furthermore, PharMA contended that the Act was "discriminatory in respect of the enjoyment of patent rights in the pharmaceutical field which discrimination is in conflict with the provisions of Article 27 of the [TRIPS] Agreement."[153]

The PharMA and South Africa settled the suit at the final hour in April 2001.[154] Pressure from the United Nations and the European Union, as well as public relations pressure, was the impetus for the settlement.[155] The result was that PharMA cut the prices of its patented AIDS drugs for South Africa, and South Africa agreed to abide by TRIPS guidelines.[156]

d. Compulsory Licensing in Canada

In 1993, Canada passed an amendment to its Patent Act that eliminated the compulsory licensing scheme for medicines.[157] Prior to, and even af-

148. *See* Notice of Motion in the High Court of South Africa, Transvaal Provincial Division (No. 4183/98), Pharmaceutical Company Lawsuit Against the Government of South Africa, *available at* http://cptech.org/ip/health/sa/pharmasuit.html (last visited Mar. 27, 2002) [hereinafter "PharMa Motion"].

149. *See id.* (included within the forty-two movants were such names as Merck & Co., Inc., Eli Lilly and Co., Bristol-Myers Squibb Co., Boehringer-Ingelheim Int'l, and Glaxo Wellcome (South Africa)).

150. *See* Nash, *supra* note 19, at 491 (writing that antiretroviral therapy has reduced mortality rates of people living with HIV and improved quality of life, and that in order for antiretrovirals to be effective the health care system in which it is introduced must be able to manage it).

151. *See* PharMa Motion, *supra* note 148.

152. *See* Nash, *supra* note 19, at 492. Nash concludes that the Act may violate the TRIPS Agreement and, worse, it may set an example for widespread compulsory licensing, which may not serve the purpose intended. *See id.* at 501-2; *see also* David Benjamin Snyder, *South Africa's Medicines and Related Substances Control Amendment Act: A Spoonful of Sugar or a Bitter Pill to Swallow?*, 18 DICK. J. INT'L L. 175, 197-98 (1999).

153. PharMa Motion, *supra* note 148.

154. *See* Denise Gellene, *Settlement Does Little Harm to Drug Firms*, LOS ANGELES TIMES, Apr. 21, 2001 at C1.

155. *See id.*

156. *See id.*

157. *See* Sheldon Burshtein, *Sublicense or Supply Agreement? Supreme Court of Canada Interpretation Benefits Generic Pharmaceutical Industry*, 54 FOOD & DRUG L.J. 73, 75 (1999).

ter, the elimination of compulsory licensing in Canada there was a debate over whether compulsory licensing should be allowed for the public interest without proof of abuse.[158]

Compulsory licensing provisions for food and medicine have been part of Canada's Patent Act since 1923.[159] In the 1980s, a bill to substantially reduce the period in which compulsory licensing could be granted was passed.[160] This bill was later partially superseded by a bill passed in 1993 which abolished compulsory licensing for medicines.[161]

The European Community filed a complaint with the WTO alleging that Canada's Patent Act violated the TRIPS Agreement. The case was eventually settled under the Dispute Settlement Body [hereinafter "DSB"] of the WTO.[162] In April 2000, Canada informed the DSB that it needed a reasonable amount of time to comply with the settlement.[163]

e. Compulsory Licensing in India

India offers a unique approach to compulsory licensing. For over thirty years, Indian pharmaceutical companies have been operating under the India Patents Act of 1970.[164] India's Patents Act focuses on public interest rather than protecting the rights of patent owners.[165] Thus, the Act allows a license automatically for food, medicine, drugs, or chemical processes.[166] There is, however, a three-year period of exclusive rights for the original patent holder.[167] After those three years, any person can apply for a com-

---

158. *See* Michael Halewood, *Regulating Patent Holders: Local Working Requirements and Compulsory Licenses at International Law*, 35 OSGOODE HALL L.J. 243, 269-70 (1997). The Canadian domestic pharmaceutical industry has not loomed large which some argue is the reason Canada was able to allow compulsory licensing. *See also* Fisch, *supra* note 7, at 296-97 n.12.

159. *See* Edward Hore, *A Comparison of United States and Canadian Laws as They Affect Generic Pharmaceutical Market Entry*, 55 FOOD & DRUG L.J. 373, 381 (2000).

160. *See id.* at 382.

161. *See id.* at 383.

162. *See* Sue Ann Mota, *TRIPS-Five Years of Disputes at the WTO*, 17 ARIZ. J. INT'L & COMP. L. 533, 539 (2000); *see also infra* notes 237-50 and accompanying text.

163. *See* Mota, *supra* note 162, at 546.

164. *See* Donald G. McNeil Jr., *Selling Cheap 'Generic' Drugs, India's Copycats Irk Industry*, N.Y. TIMES, Dec. 1, 2000, at A1.

165. *See* Elizabeth Henderson, *TRIPs and the Third World: The Example of Pharmaceutical Patents in India*, *in* THE INTERNATIONAL INTELLECTUAL PROPERTY SYSTEM: COMMENTARY AND MATERIALS PART ONE 712, 716 (Frederick Abbott et al. eds., 1999).

166. *See id.* at 714.

167. *See id.*

pulsory license where "[t]he reasonable requirements of the public with respect to the patented invention have not been satisfied or [where] the patented invention is not available to the pubic at a reasonable price."[168] Practically speaking this means that Indians can pay one-twentieth to one-fiftieth of the price they would otherwise pay in the United States.[169] This still complies with the TRIPS Agreement because certain developing countries, India included, have a special time frame in which to become TRIPS compliant; India has until January 1, 2006.[170] So far, India's legislature has not brought its Act into compliance with the TRIPS Agreement.[171]

Despite the liberal compulsory licensing provisions, which many industry groups contest,[172] India has rarely used it.[173] However, it could be used as a negotiating tool with foreign nations.[174] Some of the problems for India, probably not atypical for many developing countries, is "that the majority of Indian patents are owned by foreigners and that a large proportion of these are not commercially worked . . . ."[175] Some countries require patents to be worked in the country that grants the patent, otherwise, in one to three years the unworked patent will be subject to compulsory license or lapse.[176] Compulsory licensing helps alleviate unworked patents in developing countries.[177]

India's goal, per the Patents Act, was to "encourage the founding of local industries to break the choke hold of foreign chemical companies."[178] In 1988, the Indian government established a group to find out if this intention had been realized; they reported that excluding some patents completely from patent law "served Indian interests well."[179] Serving the pub-

---

168. *Id.* (alterations in original).

169. *See* McNeil, *supra* note 164, at A1.

170. *See* TRIPS Agreement, *supra* note 52; GERVAIS, *supra* note 65, at 257 (complying with Articles 3, 4, and 5 is a current requirement, and these articles provide reciprocity between nations, fairness in treating foreign nations, and exemption of other multilateral agreements under World Intellectual Property Organization).

171. *See* McNeil, *supra* note 164, at A1.

172. *See* THE INTERNATIONAL INTELLECTUAL PROPERTY SYSTEM: COMMENTARY AND MATERIALS PART ONE 711 (Frederick Abbot et al. eds., 1999) [hereinafter Abbot]; *and* McNeil, *supra* note 164, at A1 (quoting PharMa spokeswoman Shannon Herzfeld as saying "[w]e object to their premise that intellectual property rights are a barrier to access to good medicine . . . . Anyone who says, 'We have to steal' is wrong. Stealing ideas is not how one provides good health care.").

173. *See* Abbott, *supra* note 172, at 711.

174. *See id.*

175. Henderson, *supra* note 165, at 712.

176. *See* NATIONAL RESEARCH COUNCIL, *supra* note 10, at 125.

177. *See* Henderson, *supra* note 165, at 715.

178. McNeil, *supra* note 164, at A1.

179. Ruth L. Gana, *Prospects for Developing Countries Under the TRIPs Agree-*

lic interest, or in other words accomplishing national welfare objectives, is a common theme for patents worldwide.[180]

III. ANALYSIS

1. AIDS and the Human Elements: AIDS Sufferers and Drug Creators

Now that a narrow survey of the history, laws, policies and cases concerning compulsory licensing, in the United States, South Africa, Canada and India, are in view, it is appropriate to look closer at the peculiarity of issues surrounding AIDS drugs and their compulsory licensing. Compulsory licensing in technical sectors may seem abstract to non-economists, but compulsory licensing of pharmaceutical drugs is a different matter, especially where the suffering of AIDS/HIV victims has a face. The virus infects thirty-four million people worldwide.[181] Furthermore, eighteen million people have already died from the disease.[182] Developing countries in general have ninety-five percent of the global total of HIV/AIDS infectees,[183] while sub-Saharan Africa claims seventy percent of infectees,[184] making this issue a personal and hotly debated topic, especially in countries like South Africa.[185]

In July 2000, the International AIDS Conference was held in South Africa.[186] Highlighted at the conference were real life stories about those suffering with AIDS in South Africa.[187] In a broader stroke, the conference

---

ment, 29 VAND. J. TRANSNAT'L L. 735, 746-47 (1996).

180. *See id.* at 747.

181. *See* Claire Bisseker, *The Turning of the Tide*, FIN. MAIL (S. AFR.), July 21, 2000, at 59 (Current Affairs).

182. *See id.*

183. *See id.*

184. *See* Robert Cunningham, *Fine Points of Domestic Debate Dwarfed by Global AIDS Challenge*, MED. & HEALTH, vol. 54, issue 10, Mar. 6, 2000, at 1S1.

185. *See supra* notes 145-56 and accompanying text.

186. *See South Africa Hosts World AIDS Conference as Continent Reels From Diseases Effect*, JET, July 24, 2000, at 6.

187. *See* Chris McGreal, *Nkosi's Last Stand: He Was Born With AIDS. It Killed His Mother. It Will Soon Kill Him. But in His Short Life, 11-Year-Old Nkosi Johnson Has Become a Potent Symbol of the Fight Against Bigotry in South Africa*, THE GUARDIAN (LONDON), Jan. 11, 2001. This article draws out the realities of the eleven year old's final days:

> Nkosi's lip quivers as he tries to manage a smile but the virus eating at his brain and the seizures of recent days have taken their toll. The diarrhoea which set in a few weeks ago - the first harbinger - has drained the flesh from his tiny frame. Nkosi does not even have the energy to turn himself over anymore. He communicates mostly by squeezing Johnson's hand.

also highlighted two distinct ways of addressing the AIDS epidemic in developing countries.[188] They were: 1) Uganda attempting to reduce HIV transmission, and 2) Brazil choosing to reduce AIDS related mortality.[189] In Uganda, President Museveni and the government have provided proof that prevention efforts can reduce HIV transmission, with infection rates falling fifteen percent between 1991 and 1998, noting the biggest change in fifteen to nineteen year olds.[190] Uganda's modus operandi was to educate people and reduce risky sexual behavior, as well as offer health care packages to Ugandans who are voluntarily tested for HIV.[191]

While Uganda's success in reducing HIV transmission rates is heartening, their results do not indicate that those who are already infected with HIV have access to treatment.[192] Brazil addressed the issue of treatment by providing all individuals infected with HIV better access to anti-retrovirals.[193] Brazil licensed local manufacturers to produce anti-retrovirals by compulsory license for the government.[194] Now the government supplies five AIDS drugs from its laboratories which reach 81,000 people, as opposed to 20,000 people reached four years earlier when less money was spent on creating anti-AIDS drugs.[195] By the end of 1999, AIDS mortality in Brazil had dropped fifty percent, and the money Brazil spent on anti-retrovirals was almost completely recovered by a boost to the economy from the lower mortality rate.[196]

The price of the drugs is the most prohibitive factor to distribution in developing countries. It is estimated that the average South African earns less than $3,000 per year, while the cost of the anti-retroviral therapy for

---

Id. See also Rosie DiManno, *AIDS Gloom Lifted by Boy's Glee*, THE TORONTO STAR, July 14, 2000 (reporting on Nkosi's speech and press talks at the International AIDS Conference in South Africa).

188. *See* Bisseker, *supra* note 181, at 59.

189. *See id.*

190. *See Uganda Tackles AIDS from the Very Top Down*, AIDS ALERT, vol. 14 issue 8, Aug. 1, 1999, at 1S3 [hereinafter "AIDS ALERT"]; *see also* Bisseker, *supra* note 181, at 59. There are slightly varying percentages between the two articles, with Bisseker noting only a ten percent decrease. *See id.*

191. *See* AIDS ALERT, *supra* note 190, at 1S3.

192. *See supra* note 190 and accompanying text.

193. *See* Bisseker, *supra* note 188.

194. *See id.*

195. *See* Stephen Buckley, *Brazil Becomes Model in Fight Against AIDS; Government, Activists Team to Defy Epidemic Through Distribution of Drugs*, THE WASH. POST, Sept. 17, 2000, at A22.

196. *See* Bisseker, *supra* note 188. *See also* African Growth and Opportunity Act of 2000, *supra* note 76, §§ 129(a)(4), (5)(A), (B) (stating HIV/AIDS crisis in sub-Saharan Africa threatens the future workforce, is weakening families and societies, and is hitting the age group most likely to support the economy in that region).

one person in the United States approaches $15,000 per year.[197] In India, however, a month's supply of the anti-retroviral AZT costs $48 compared to $239 per month in the United States: India can provide AZT at this low cost because pharmaceutical production is currently unregulated.[198] Cipla, an Indian Pharmaceutical Company, has offered to sell its generic version of an anti-retroviral drug therapy to Medecins sans Frontieres for $350 per therapy instead of the international price of around $10,000-$12,000.[199] This offer depends on the response Cipla receives from multinational pharmaceutical companies, to whom Cipla offered to sell generic versions of the pharmaceutical companies' HIV/AIDS drugs in return for five percent royalty on net sales.[200]

On the other hand, the people who create these drugs, and their interests, must be considered in this debate. Drug creators (or innovators) are influenced by state-granted property rights on innovation because otherwise, they would have little incentive to engage in research if other people could "commercialize the results of the inventor's research without consent or approval."[201] A strong patent system, such as TRIPS, could possibly correct the disincentive that comes from competitors who "free-ride" on the work of innovators.[202]

2. Arguments in favor of pharmaceutical compulsory licensing

The arguments in favor of pharmaceutical compulsory licensing come in a variety of forms from straight forward immediate life-saving, to farsighted development of policy for consistent and affirmative life-saving. In providing for compulsory licensing where there exists a "national emergency," the TRIPS Agreement has not set out a precise definition of national emergency, and many factors should be considered.[203] This section examines numerous factors that contribute to a possible national emergency standard, and makes the argument that compulsory licensing of pharmaceuticals is imperative. There are economic matters to consider

197. *See* Sarah Boseley, *At the Mercy of Drug Giants: Millions Struggle with Disease as Pharmaceutical Firms Go to Court to Protect Profits*, THE GUARDIAN (LONDON), Feb. 12, 2001, at 2 (basing South African salary approximations on employed people).

198. *See supra* notes 165-80 and accompanying text.

199. *See Cipla Talks with MNCs Over HIV Drug Offer*, MARKETLETTER, (Marketletter Publications Ltd., London, U.K.), Feb. 19, 2001.

200. *See id.*

201. *See* Gutterman, *supra* note 26, at 89, 91.

202. *See* Martin J. Adelman & Sonia Baldia, *Prospects and Limits of the Patent Provision in the TRIPS Agreement: The Case of India*, 29 VAND. J. TRANSNAT'L L. 507, 508 (1996) (stating that patent systems correct the defect of allowing competitors to "free-ride" by protecting against it in all market structures).

203. *See* TRIPS Agreement, *supra* note 52, art. 31(b).

which, like incentives for further invention if withheld from innovators, might force a national emergency into an even deeper emergency. Furthermore, there are social justice issues surrounding the serious health issues that compulsory licensing could help alleviate.[204]

The fact that many nations have distinct goals for compulsory licensing[205] brings to the forefront the idea that a "national emergency" could have a broad interpretation. For example, while the United States may grant a compulsory license for pollution control, or nuclear materials, Russia finds "any invention 'of special importance to the state'" to be subject to compulsory licensing.[206] In comparison, Switzerland may order a compulsory license for a patent simply to lower prices.[207]

Fauver states that compulsory licensing "enables the government granting the patent to force the patentee to license the invention if the government does not approve of the patent's use."[208] His article compares economic theories behind compulsory licensing and concludes that compulsory licensing advances a nation's technological growth.[209] Furthermore, justification for these public interest compulsory licenses is based on an increase in public access to these inventions, which increases competition, and will in turn advance technology.[210] "In the long run, the public interest is served by widening the availability of drugs through reduced prices while ensuring profits to maintain incentives for innovation."[211] Patents of pharmaceuticals are actually harmful to countries that are still developing their own domestic and technical infrastructure, and the patent laws become an impediment to economic growth.[212] This is the exact opposite result that patents are supposed to achieve.[213]

There is yet another argument that favors compulsory licensing. The compulsory licensing issue, as witnessed by the countries who invent, and those who are trying to compel licenses, raises thoughts of "neocolonialism."[214] Trade liberalization, as the protection of patent rights is sometimes called, disproportionately benefits the industrialized nations, whereas most

---

204. *See infra* notes 222-26 and accompanying text.

205. *See* Fauver, *supra* note 72, at 670-71.

206. *Id.*

207. *See id.* at 667.

208. *Id.*

209. *See id.* at 685.

210. *See id.* at 667.

211. Griffin, *supra* note 16, at 404 (suggesting in her article that the government could act to reduce the cost of AIDS therapy through direct price controls).

212. *See* Carroll, *supra* note 17, at 2471.

213. *See* Oddi, *supra* note 8, at 856-57.

214. *See* Susan Demske, Note, *Trade Liberalization: De Facto Neocolonialism in West Africa*, 86 GEO. L.J. 155, 173 (1997); *see also supra* notes 183-85 and accompanying text.

developing countries have few patents to protect.[215]

Another economic reason that supports compulsory licensing where pharmaceuticals are concerned is that lives are saved by making drugs more accessible.[216] In Brazil, for example, where the government licensed anti-retrovirals in 1996,[217] the costs incurred by the compulsory licensing were, by 1999, practically cancelled out by the overall savings to the economy in productive lives saved.[218] Furthermore, because 35 million people are infected worldwide, production is reduced, thereby harming the economy.[219] In South Africa, the population is currently around forty-three million, although it was projected to be at around fifty million.[220] This low growth rate has resulted from high mortality rates caused by HIV/AIDS.[221]

Aside from economic arguments, compulsory licensing for the sake of public interest is valid on social justice grounds as well.[222] South Africa's population of forty-three million includes more than 3.5 million people infected with HIV.[223] It is also alleged that nearly 1600 new HIV infections occur daily in South Africa.[224] These stark realities inspired a TIME article to query, "but at what point does the human benefit to desperate, destitute countries outweigh strict adherence to patents and profits?"[225]

In the United States, after the September 11th terrorist attack, an anthrax

---

215. *See* Demske, *supra* note 214, at 173. *See also* ABBOTT, *supra* note 172, at 711 (reminding reader that granting of compulsory licenses is at times favored "to maintain a balance of interests among industrialized and developing countries. Though developing countries do not appear to grant compulsory licenses with any frequency, industry groups have argued that the mere possibility that a compulsory license may be granted affects negotiations on the transfer of technology").

216. *See infra* note 223.

217. *See id. Supra* notes 198-201 and accompanying text.

218. *See* Bisseker, *supra* note 188, at 59.

219. *See* Exec. Order No. 13,155, 65 Fed. Reg. 30,521 (2000); Eric Foner, *AIDS in Africa; AIDS Conference in South Africa; The High Cost of AIDS Drugs*, THE NATION, Aug. 7, 2000 ("The sheer scale of the AIDS epidemic has forced Western governments to see the situation in Africa as more than a mere medical emergency. With so many dead or dying, the implications for food production, education, international economics and global security are grave.").

220. *See* BBC Worldwide Monitoring, *Report Says Population Growth Shrinking Due to AIDS* (Feb. 5, 2001).

221. *See id.*; African Growth and Opportunity Act of 2000, *supra* note 76.

222. *See* Foner, *supra* note 219 (stating economic reasons for calling AIDS a crisis, but "above all, it is a moral issue").

223. *See* Zackie Achmat, *We Can Use Compulsory Licensing and Parallel Imports: A South African Case Study, AIDS Law Project, available at* http://www.hri.ca/partners/alp/tac/license.shtml (last visited August 5, 2001).

224. *See id.*

225. Johanna McGeary, *Paying for AIDS Cocktails; Who Should Pick up the Tab for the Third World?*, TIME, Feb. 12, 2001; *see also* McGreal, *supra* note 187.

scare made the utility of compulsory licensing hit home. During the height of the scare, in October 2001, U.S. politicians were calling for patents to be overturned and Canada (although later reversing this decision) overrode the patent on Cipro and ordered 1 million pills.[226] The anthrax cases numbered no more than a dozen.

Many so-called AIDS activists are pushing for compulsory licensing, including such groups as Medecins sans Frontiers, Health Action International and Act Up USA.[227] These groups make action plans, lobby legislators, and protest in order to bring attention and promote action in fighting AIDS.[228] For example, AIDS activists have been responsible for "slamming" an offer by the United States to give loans to sub-Saharan African nations to purchase reduced-price drugs.[229] Activists, joined by the South African Department of Health, have responded cynically toward an offer by pharmaceutical companies to cut AIDS drugs prices in South Africa and other developing countries.[230] The cynicism has arisen from distrust of the pharmaceutical companies.[231]

AIDS activists have received some encouragement for their activism.[232] The European Union Trade Commissioner, Pascal Lamy, announced that the "European Commission can help poor nations facing health emergencies to override international rules on patent protection, which makes anti-AIDS and other remedies too expensive for Third-World citizens."[233] This type of announcement offers hope, albeit slight, but it is recognized as a step in the right direction because it "acknowledg[es] a link between pat-

---

226. *See* Geoff Dyer, *Cipro Case Poses Threat to Patent Law*, FINANCIAL TIMES (Nov. 6, 2001).

227. *See* Daniel Pruzin, *U.S., EU Resistance to Compulsory Drug Licensing Draws Health Care Fire*, BNA INTERNATIONAL BUSINESS & FINANCE DAILY, Mar. 29, 1999, News (activist groups accusing U.S. government in particular of pressuring developing countries to "refrain from adopting provisions which would allow compulsory licensing"); Achmat, *supra* note 223 (Treatment Action Campaign).

228. *See generally* Sabin Russell, *New Crusade to Lower AIDS Drug Costs*, THE SAN FRANCISCO CHRONICLE, May 24, 1999; Sternberg, *supra* note 1.

229. *See AIDS Groups Slam US $1 Billion for Africa*, MARKETLETTER (Marketletter Publications Ltd., London, U.K.), July 31, 2000 (offering loans financed by Congress for five-year terms at 7% interest in order to purchase drugs).

230. *See South Africa Not Impressed*, FIN. MAIL (S. AFR.), May 19, 2000, Current Affairs, at 36.

231. *See id.* (agreeing that prices would have to be less than a tenth of the world prices to be affordable, infrastructure in South Africa might still be too much of a barrier to distributing such drugs, and assuming that a condition will be placed on the offer such as abandoning the fight to introduce compulsory licensing).

232. *See* Shada Islam, *Hope is at Hand*, FIN. MAIL (S. AFR.), Oct. 6, 2000, Economy & Business, at 40.

233. *Id.*

ent protection and prices of key medicines."[234]

Disputes continue to arise over complaints of abusing the TRIPS Agreement. These complaints are lodged at the WTO's Dispute Resolution Board (DRB).[235] The DRB received a complaint from the European Communities [hereinafter "EC"] and their member States arguing that Canada's Patent Act violated the TRIPS Agreement and that the violations were causing the European Community's pharmaceutical industry to suffer economic losses.[236] One issue that was highlighted by WTO's Dispute Resolution Board in its Report of the Panel was whether the compulsory licensing that Canada's Patent Act allowed was within Article 30 of the TRIPS Agreement.[237] The EC argued that the provisions were not within Article 30 because the "curtailment of patent rights under Canadian legislation did not constitute 'limited exceptions to the exclusive rights conferred by a patent.'"[238] Furthermore, the EC's dominant argument was that the Act was discriminatory to the pharmaceutical industry in the first place and so it was already invalid.[239]

Canada's response was that the Act did not conflict with any "normal exploitation" of a patent, nor did it unreasonably prejudice any patent or "the 'legitimate interests' of a patentee taking account of the 'legitimate interests' of third parties."[240] Canada argued that Article 30 required that countries allow for only "limited exceptions to the exclusive rights con-

---

234. *Id.* (accepting that a radical shift in policy must be monitored and encouraged by activist groups such as Medicins sans Frontiers, and hopeful of the promise that the EU would "explore the flexibility provided by the TRIPS agreement to provide compulsory licensing and technical help to address public health concerns and emergency crises"). Another movement fighting the AIDS scourge has come from the African trade unions. *See* Judith Achieng, *Health – Africa: Trade Unions Declare All-Out Anti-AIDS War*, INTER PRESS SERVICE (Nairobi), Oct. 20, 2000. The movement brings a $6.3 million campaign to the table and such fighting words as: "'The greatest scandal of our time is that drug companies are holding dying people at ransom.'" *See id.* (the International Conference of Free Trade Unions' Secretary General, Bill Jordan spoke those words, stating further that "'[i]f there is anyone as bad as they are, it is the government leaders who defend the right by these companies to do so through punitive intellectual property rights regulations'").

235. *See generally* WTO, *Dispute Settlement, at* http://www.wto.org/english/tratop_e/dispu_e.html (last visited Aug. 5, 2001).

236. *See* WTO Dispute Resolutions *Report of the Panel: Canada – Patent Protection of Pharmaceutical Products*, WT/DS114/R, 15 (Mar. 17, 2000) WTO, *WTO Dispute Reports, available at* www.wto.org/english/tratop_e/dispu_e/7428d.doc (last visited Aug. 5, 2001).

237. *See id.*

238. *Id.* at 14.

239. *See id.* at 16.

240. *Id.*

ferred by a patent," to which they complied.[241] Canada's exceptions for regulatory review[242] and stockpiling[243] were limited to pharmaceuticals, the exceptions did not conflict with the life of the patent, they were within Canada's national interest according to Article 7, and did not prejudice the legitimate interests of the patent owners.[244]

Numerous countries (third parties) responded to the issue by preparing comments for the DRB on this matter.[245] India, for example, was in favor of finding Canada's Patent Act consistent with TRIPS.[246] The crux of India's argument was that a "systemic issue" was at stake: "[T]he balance in the TRIPS Agreement between the rights and obligations of the producers of technical knowledge on the one hand and the users of that knowledge on the other."[247] India cited specific provisions of the TRIPS Agreement, arguing that it was full of textual language that aimed to reconcile the rights and obligations of patent owners.[248] India argued that any other finding "would alter the delicate balance in the TRIPS Agreement between the rights conferred on the owner of the patent and the legitimate interests of all those other than the patent owner."[249] The Panel concluded that Canada had to bring only the stockpiling exception part of its disputed Act into conformity with the TRIPS Agreement; the regulatory review exception of the Act was consistent with TRIPS.[250]

---

241. *Id.*

242. Regulatory review requires a pharmaceutical to be approved before entering the market. The issue here was *when* the generic manufacturer could start the regulatory review process, before or after the original patent term expired. Generic manufacturers argued that if they had to wait until the patent term expired to go through regulatory review, then the actual patent term would be de facto extended between one to two-and-a-half years, the average length of regulatory review of generic pharmaceuticals in Canada. *See id.* at 146-47.

243. Stockpiling refers to the practice of manufacturing pharmaceuticals while the patent term exists in order to enter the market in full force immediately after the patent term expires. *See id.* at 148.

244. *See* WTO Dispute Resolutions *Report of the Panel: Canada – Patent Protection of Pharmaceutical Products*, WT/DS114/R, 17 (Mar. 17, 2000) WTO, *WTO Dispute Reports*, *available at* www.wto.org/english/tratop_e/dispu_e/7428d.doc (last visited Aug. 5, 2001).

245. *See id.* at 94.

246. *See id.* at 119.

247. *Id.* at 119.

248. *See id.* at 119-20 (using Art. 7, Objectives, Art. 8 Principles, as well as Art. 30 Exceptions to Rights Conferred of the TRIPS Agreement).

249. *Id.* at 119.

250. *See* WTO Dispute Resolutions *Report of the Panel: Canada – Patent Protection of Pharmaceutical Products*, WT/DS114/R, 15 (March 17, 2000) *available at* http://www.wto.org/english/tratop_e/dispu_e/7428d.doc (last visited Nov. 18, 2000).

3. Arguments against pharmaceutical compulsory licensing

"In general, analysts and policymakers who propose remedies such as compulsory licensing demonstrate a failure to understand the full implications of the economic model of induced innovation, a model that was implicitly supported by the framers of the United States Constitution."[251] This "failure to understand the implications" is not completely dependent on the actual occurrence of compulsory licensing to date, in fact, the fear of compulsory licensing is enough to affect the trade relations between countries.[252] Pharmaceutical companies may prefer not to risk entering a country where they are aware that internationally recognized patent law will not be respected.[253]

Likewise, pharmaceutical companies consider the potential for profit when deciding on which drugs to focus their research.[254] If certain diseases, like AIDS, remain more serious in developing countries that do not protect patents, companies may be more likely to curtail or even withdraw research for those drugs.[255] With the cost of creating drugs extremely high, pharmaceutical companies must consider expenses beyond the simple manufacturing of a product.[256] Only under a monopoly (by use of a patent) can pharmaceutical companies raise the product's price at will and recover for the expense of research and development.[257]

Furthermore, there is concern that the introduction of pharmaceutical drugs by compulsory licensing will not solve problems as hoped, but will instead lead to its own unique problems.[258] For example, currently there is

---

251. Khan, *supra* note 101, at 156. "The Congress shall have Power . . . to promote the Progress of Science and useful Arts, by securing for limited Times to Authors and Inventors the exclusive Right to their respective Writings and Discoveries." U.S. CONST. Art. I, § 8.

252. *See* ABBOTT, *supra* note 172, at 711. Compulsory licensing is likened to piracy. *See* McNeil, *supra* note 164; James M. Silbermann, Comment, *The North American Free Trade Agreement's Effect on Pharmaceutical Patents: A Bitter Pill to Swallow or a Therapeutic Solution?*, 12 J. CONTEMP. H.L. & POL'Y 607, 607-08 (1996).

253. *See* Nash, *supra* note 19, at 498.

254. *See* Adelman & Baldia, *supra* note 202, at 511; Oddi *supra* note 8, at 849 (going further to state that there are other, probably more important, factors than patent protection for pharmaceutical companies to balance when deciding whether to enter a country; such as, political stability and availability of materials and labor force).

255. *See* Fisch, *supra* note 7, at 302-03.

256. *See id.* at 303-04 (stating that only one out of 4,000 discovered compounds will become a product functional to be sold in the pharmaceutical market).

257. *See id.* at 307.

258. *See* Sternberg, *supra* note 1 (quoting Edmund Nminyem Delle, President of the African Commission on Health and Human Rights: "If they bring drugs into a

no drug which eradicates AIDS or HIV, and prolonging people's lives increases the number of infected people who will have even lower tolerance to infections.[259] In order to *not* create drug-resistant strains of HIV, the present drug regimen must be properly delivered, and must be taken according to a strict schedule.[260] For AIDS drugs to be most useful, even at a reduced price, these and other factors must first be addressed by the developing nations.[261]

As discussed above,[262] Canada's Patent Act has been debated for years.[263] In one recent instance, Apotex, Inc., Canada's largest manufacturer of generic drugs, offered to provide affordable drugs in sub-Saharan Africa.[264] The "Big Pharma" pharmaceutical companies fear that Apotex will succeed in getting the Canadian government to grant a short-term compulsory license for emergency reasons under the Patent Act and the TRIPS Agreement.[265] An Apotex representative hopes that they will be granted the short-term license immediately so they can begin delivery of the drug to Africa.[266] Dr. Mark Wainberg, director of the McGill University AIDS Centre, however, argued that this attempt may terrorize big pharmaceutical companies, which would not be beneficial.[267] Dr. Wainberg argued that patent protection creates profit for these companies, which is the only incentive for the pharmaceutical industry to continue AIDS research.[268]

Pharmaceutical patents are thought by some to be the most important patent rights to protect.[269] In May 1999, Clinton administration officials

---

country without educating doctors on what to do, the world will waste a lot of money. The world is about to make a big mistake.")

259. *See* Nash *supra* note 19, at 499.

260. *See Poor Developing Countries are Offered Cheaper AIDS Drugs by Manufacturers, But Other Costs Regarding the Infrastructure are Prohibitive*, (NPR: All Things Considered, Radio Broadcast July 10, 2000) 2000 WL 21471147.

261. *See id.*

262. *See supra* notes 157-63 and accompanying text.

263. *See id. See also* Tesher, *supra* note 18 (varying opinions have kept the debate lively and ongoing in Canada, as well as keeping it in litigation).

264. *See id.* (Apotex, Inc. now manufacturers the AIDS-fighting drug, AZT, in Canada and the patent rights of the product they are trying to give to Africa is currently under litigation regarding its lawful use).

265. *See id.* (allowing Apotex to make drugs and "sell them at cost to the Canadian government, which would distribute them to African countries as foreign aid").

266. *See* Ellie Tesher, *Africa Waits as Drug Firms Work it Out*, TORONTO STAR, Aug. 29, 2000, at A19.

267. *See* Tesher, *supra* note 18.

268. *See id.*

269. *See* Doane, *supra* note 92, at 470-71. In the article, Doane stated that:

> Many nations deny patent protection to pharmaceutical products which

reminded the world of the importance of its trade policy, especially for pharmaceuticals, stating that tampering with AIDS drug patents threatens the intellectual property rights protection which ensures the search for new drugs.[270]

The pharmaceutical industry maintains that compulsory licensing erodes patent rights.[271] A Bristol-Myers Squibb associate director of tax and trade policy said that, "[t]he only beneficiary of an erosion of patent rights . . . is the generic industry, which does not subsidize through its sales the costly process of research—research that makes new top-of-the-line therapies available."[272]

4. The Attempts at Meeting on Common Ground

Besides the pro and con arguments, there are many proposals for reaching common ground between proponents and opponents of compulsory licensing. And, although the Clinton administration found compulsory licensing a threat, it also suggested that a solution can be found without destroying the patent system.[273] Some of these middle-ground solutions are those that the pharmaceutical companies have suggested but that have not been accepted well by the cynical AIDS activists or even by government officials.[274] Glaxo-Wellcome's (a major pharmaceutical manufacturer) medical director for sub-Saharan Africa assessed an announcement of a major policy shift by the Clinton administration toward a "co-operative approach on health-related intellectual property matters," by saying, "[I]f there is a co-operative agreement that provides less expensive drugs while also respecting companies' intellectual property rights, we welcome it."[275] Even people like PharMA's Chief Executive Officer indicate that offers to cut drug prices are not conditional on forsaking compulsory licensing, but rather on removing the need for compulsory licensing with a better solu-

---

> by their nature require considerable time and expense to develop and bring to market. Consequently, some pharmaceutical companies face foreign competitors who misappropriate information with the active assistance and encouragement of their governments to produce inexpensive and potentially ineffective or dangerous imitations.

Id. at 470.

270. *See* Sabin Russell, *New Crusade to Lower AIDS Drug Costs: Africa's Needs at Odds with Firms' Profit Motive*, S.F. CHRON., May 24, 1999, at A1.

271. *See id.*

272. *Id.*

273. *See id.*

274. *See generally*, Claire Bisseker, *South Africa Not Impressed*, FIN. MAIL (S. AFR.), May 19, 2000, at 36; *see also supra* note 232-34 and accompanying text.

275. *See* Aaron Nicodemus, *Clinton Gives in on AIDS Drugs for 'Poor' Countries*, AFRICA NEWS, Dec. 3, 1999.

tion.[276]

Making gifts of drugs was, however, chosen by the pharmaceutical companies over a request to allow a generic manufacturer a license.[277] Those gifts include a German drugmaker giving away an unlimited supply of a drug that reduces an infected mother's risk of transmitting AIDS to her child, a focus program on Botswana with $100 million, and an unlimited amount of money for funding and medical support to Tanzania and Burkina Faso.[278]

However, another attempt to meet on common ground would be to use experience and standards from antitrust compulsory licensing and to apply it to national emergency compulsory licensing.

5. Comparison Between Pharmaceutical Compulsory Licensing and Antitrust Compulsory Licensing

Under the TRIPS Agreement, compulsory licensing[279] for circumstances of national emergency and compulsory licensing for remedying anti-competitive behavior are very similar.[280] The authorization for each type of compulsory licensing must be individually reviewed, limited in use, allow for adequate remuneration, and be subject to judicial review.[281]

Compulsory licensing is invoked in either case to remedy a problem that is either harming, or will harm the national welfare.[282] In either case the subsequent increase in availability and public access increases the market value of the product.[283] Although prices may be reduced in this process, the opening up of the market to more consumers and greater competition

---

276. *See id. See also supra* notes 156-59 and accompanying text.

277. *See* Steve Sternberg, *AIDS Activists Discount Big Drugmakers' Gifts Generic Producers Not Allowed to Fight Epidemic for Less*, USA TODAY, July 11, 2000, at 9D.

278. *See id.*

279. *See* GERVAIS, *supra* note 65 and accompanying text (compulsory licensing is the same as TRIPS' "Other Use Without Authorization of the Right Holder").

280. *See supra* notes 68-73 and accompanying text. However, compulsory licensing for a national emergency may waive the necessity of obtaining authorization but must notify the right holder as soon as reasonably practicable. *See* TRIPS Agreement, *supra* note 52, at (b). Compulsory licensing to remedy anti-competitive behavior is also able to waive the authorization requirement, but *is not bound to use in the domestic market. See* TRIPS Agreement, *supra* note 52, at (k).

281. *See* TRIPS Agreement, *supra* note 52, art. 31.

282. *See supra* notes 62, 108, 169 and accompanying text; *see also* United States v. Glaxo Group Ltd., 410 U.S. 52, 58 (1973) ("The licensee was free to challenge the patent in these circumstances because the 'federal courts must, in the public interest, keep the way open for the challenge of patents which are utilized for price-fixing.'").

283. *See supra* notes 212-13 and accompanying text for a discussion of pharmaceutical compulsory licensing, and notes 126-44 and accompanying text for a discussion of antitrust compulsory licensing.

creates a wider profitability.[284] As patents are supposed to increase innovation, at times their undoing increases innovation more so.[285]

These methods are so similar, in fact, that the Director of the Consumer Project on Technology, James Love, queried whether developing countries would be able to use U.S. antitrust cases which granted compulsory licenses for their own compulsory licensing endeavors.[286] In rebuke of this idea, one response was that the antitrust cases involve consent decrees in which the parties "agreed to license some intellectual property rights if the Department Of Justice or FTC would not challenge their merger."[287] However, Love maintained his query in light of an FTC Staff report that discusses compulsory licensing as a remedy for anti-competitive problems arising from intellectual property.[288]

Furthermore, the differences are not that severe when one considers that there is a high burden to attain the compulsory licensing in the first place.[289] Antitrust compulsory licensing is a remedy that is rarely and reservedly used;[290] it would be likewise with patent compulsory licensing of pharmaceuticals.[291] In addition, there is still a burden that must be met before compulsory licensing can be granted. Antitrust compulsory licensing must only happen when there is an action which is "in restraint of trade or commerce among the several States, or with foreign nations."[292] And compulsory licensing of a patent under TRIPS will only be granted to protect public health and promote public interest waiving the requirement to get the patent holders authorization "in the case of a national emergency or other circumstances of extreme urgency."[293]

Nonetheless, the Order of the FTC in *In re Digital Equipment Corporation*[294] stated that "this remedy has the effect of ensuring ... production by

---

284. *See id.*

285. *See supra* note 126-44 and accompanying text for a discussion of antitrust compulsory licensing.

286. *See* Coalition for Networked Information, *Helping Developing Countries with IP Issues, available at* http://www.cni.org/Hforums/cni-copyright/1997-02/1084.html (last visited Feb. 22, 2001).

287. *See* Coalition for Networked Information, *Helping Developing Countries with IP Issues, available at* http://www.cni.org/Hforums/cni-copyright/1997-02/1228.html (last visited Aug. 5, 2001) (rebuking Love's query was Mark Lemley, of University of Texas Law School).

288. *See id.*

289. *See supra* notes 72-74 and accompanying text.

290. *See supra* note 144 and accompanying text.

291. *See supra* notes 172-73 and accompanying text (stating that India rarely uses their compulsory licensing provisions, which is notable because India is one of the most liberal countries to apply compulsory licensing).

292. Sherman Act, 15 U.S.C. § 1 (1994).

293. TRIPS Agreement, *supra* note 52, at 86, 95.

294. 2000 FTC LEXIS 28 (2000).

requiring Compaq ... to develop alternative sources of supply ... by licensing" its technology to other firms.[295] This can hardly be called a simple "agreement" to license by parties; even writers for the *FTC Watch* recognize compulsory licensing as an antitrust remedy.[296]

The use of compulsory licensing to remedy antitrust problems is a credit of the TRIPS Agreement, as is its use for public health emergencies.[297] The U.S. judicial test for violation of antitrust law, unless the actions are per se unlawful or per se lawful, is the rule of reason.[298] The rule of reason determines an action as lawful "only if its procompetitive effects outweigh its anticompetitive effects."[299] This rule of reason test could be used to analyze the problem of delivering AIDS drugs to developing countries' populations who are suffering from the disease, but who cannot access the drugs, particularly because it is cost prohibitive.[300]

However, if an attempt was made to apply this type of test in a national emergency situation, the government attempt should use the language of TRIPS, making patent protection lawful if its pro-public health effects outweigh its anti-public health effects.[301] Under this type of test, the argu-

---

295. *See id.* at *4.

296. *See generally Reconciling Antitrust Enforcement with Intellectual Property Rights*, FTC: WATCH, (Washington Regulatory Reporting Associates, Washington, D.C.), Sept. 25, 2000 ("compulsory licensing tends to encourage price-based competition"); *Intellectual Property*, FTC: WATCH, (Washington Regulatory Reporting Associates, Washington, D.C.), Dec. 5, 1995 ("[c]ompulsory licensing may be necessary in some [antitrust] cases, [but] it can stifle innovations in others"). Referring to *supra* note 287, where compulsory-license-for-antitrust remedy was called an "agreement" to license.

297. *See* Marschall, *supra* note 9, at 1189-90 (disparaging the TRIPS Agreement for not delineating clearly where compulsory licensing should be used: "By explicitly entitling developing countries to justify limiting the exclusive rights of patentees based on *economic development concerns*, the provisions go beyond what is required to ensure that developing countries are not abused." (emphasis added)).

298. *See id.* at 1170-71; *see also supra* note 112-21 and accompanying text.

299. *See* Marschall, *supra* note 9, at 1171.

300. *Compare* Griffin, *supra* note 16; Oddi, *supra* note 8; Fauver, *supra* note 72, at 667-70; Stefan Kirchanski, *Protection of U.S. Patent Rights in Developing Countries: U.S. Efforts to Enforce Pharmaceutical Patents in Thailand*, 16 LOY. L.A. INT'L & COMP. L.J. 569 (1994); *with* Nash, *supra* note 19; Fisch, *supra* note 7; Adelman & Baldia, *supra* note 202. See also *supra* notes 197-99 and accompanying text describing how cost prohibits many from using drugs. *See also* David K. Tomar, Note and Comment, *A Look into the WTO Pharmaceutical Patent Dispute Between the United States and India*, 17 WIS. INT'L L.J. 579, 600 (1999) (querying whether it is ironic that the United States is a strong proponent of free-trade, and yet develops international laws that restrict free-trade in India).

301. Although this is not a recognized application of the rule of reason test, the author is applying language of the TRIPS Agreement, art. 8 (1) to the rule of reason test. Art. 8 allows members to adopt measures "necessary to protect public health."

ments for and against compulsory licensing raised in the prior sections[302] would be applicable, and judicial review would simply have to weigh the public health effects.[303] The patent system has prompted important pro-public health factors from which society benefits, including: "greater ... investment for research and development[,] greater 'transfer of pharmaceutical technology' ... [, and] greater availability of drugs."[304] Notwithstanding these benefits, the costs of submitting to patent rights in situations of severe public health emergencies may outweigh the more abstract benefits.[305]

The rule of reason does not necessarily take the reasoned approach of weighing costs and benefits to society.[306] The test simply determines whether the action "promotes or suppresses" competition.[307] Adding this to the pharmaceutical equation, it now would be tested something like this: the court must determine whether the restraint of drugs alleviates the national emergency or exacerbates the national emergency. The first step in using this test would be to determine whether there is a national emergency, just as competition has been recognized by Congress to be the most important objective in antitrust cases.[308]

The factors to consider, as explained earlier, and applied with the TRIPS Agreement in mind, are the condition of the country before and after the denial of compulsory licensing, the nature of the denial, the effect of the denial (probable or actual), the history of the denial (probably the future

---

TRIPS Agreement art. 8, (1).

302. *See supra* notes 203-50 *and* notes 251-72 and accompanying text.

303. Pro-public health effects could include: alleviating the suffering of 3.5 million people in South Africa alone living with HIV, preventing further transmission of AIDS to children of AIDS victims, and reducing AIDS mortality rates. *See supra* notes 145-56 and 203-50 and accompanying text. Anti-public health effects could include: a decrease in research and development targeted at eradicating HIV/AIDS, and an inadvertent development of drug resistant strains of the virus. *See supra* notes 251-72 and accompanying text.

304. Snyder, *supra* note 151, at 188-89.

305. *See id.* Using South Africa as an example, the benefits of research and development are abstract for a nation with almost ten percent of its population infected with HIV/AIDS, *see supra* notes 181-85 and accompanying text; *see also* Carroll, *supra* note 17, at 2469-474, but not so abstract for U.S. pharmaceutical companies. *See* Nash *supra* note 19, at 499 ("The most obvious and worrisome effect of the Act on U.S. interests is decreased revenue for American pharmaceutical manufacturers").

306. *See supra* notes 111, 115 118-21 and accompanying text.

307. *See id.*

308. *See* Arizona v. Maricopa County Medical Soc., 457 U.S. 332, 354 (1982) ("Our adherence to the per se rule is grounded not only on economic prediction, judicial convenience, and business certainty, but also on a recognition of the respective roles of the Judiciary and the Congress in regulating the economy," (citing United States v. Topco Associates, Inc., 405 U.S. 596, 611-12 (1972))).

denial too), and even the intentions of the pharmaceutical company and the country attempting compulsory licensing.[309] The rule of reason is an economic standard of analysis,[310] which does not prevent the application to pharmaceutical compulsory licensing, because the same concepts could apply.[311]

IV. CONCLUSION

When the next dispute arises in the WTO over pharmaceutical patent rights, the WTO should mirror U.S. antitrust cases and apply the rule of reason test. A dispute will likely arise soon where possible violations are easy to find. For example, one orphanage in Kenya has made a request to Cipla (the Indian pharmaceutical company) for AIDS drugs, which is in possible violation of TRIPS.[312] The orphanage director has determined that the children are "on the brink," and that the children can "wait no longer."[313] If this is the type of "national emergency" that the TRIPS Agreement refers to, then the request might not be in violation of TRIPS. This type of situation, a result of the impact AIDS is having on families, workforces and the population of sub-Saharan Africa, has caused AIDS to rise to the level of "national emergency" in some nations. That being the case, if one simply asks whether the denial of compulsory licensing of drugs by patent alleviates or exacerbates the national emergency, the answer can be made on a balance of factors peculiar to the AIDS crisis and the protection of pharmaceutical intellectual property rights. Most likely, the AIDS crisis is at such à level in some countries that protection of intellectual property will not promote any of its intended goals, therefore, compulsory licensing should be demanded of the pharmaceutical industry

---

309. *See supra* note 120 and accompanying text.

310. *See Arizona*, 457 U.S. at 332.

311. *See* Continental T.V. v. GTE Sylvania, 433 U.S. 36, 54 (1977) (reminding in discussion of the importance of considering the marketplace consequences in antitrust cases. "Competitive economies have social and political as well as economic advantages, *see e.g.*, Northern Pac. R. Co. v. United States, 356 U.S., at 4, but an antitrust policy divorced from market considerations would lack any objective benchmarks. As Mr. Justice Brandeis reminded us: 'Every agreement concerning trade, every regulation of trade, restrains. To bind, to restrain, is of their very essence.' Chicago Bd. of Trade v. United States, 246 U.S., at 238."); *see also supra* notes 203-21 and accompanying text.

312. *See* BBC News: World: Africa, *Kenyan Challenge to AIDS Drugs Prices, available at* http://news/bbc.co.uk/hi/english/world/africa/newsid_1182000/1182633.stm (last visited Feb. 22, 2001).

313. *See id.* The orphanage has apparently ordered the drugs from Cipla, but the Kenyan government is frustrating their efforts to import the drugs. *See AIDS Drugs: Govt Ignores Nyumbani Plea*, THE EAST AFRICAN (March 26, 2001).

in an attempt to alleviate the AIDS crisis in those countries.

Dora Kripapuri

SYMPOSIUM

Responding to Rogue Regimes: From Smart Bombs to Smart Sanctions

November 9, 2001

Welcome Address by John F. O'Brien Dean of New England School of Law

It is my honor to welcome you today to the annual International Conference organized by our Center for International Law and Policy. Given the topic of today's Symposium, I think it is appropriate to begin our proceedings with a moment of silence for those who lost their lives in the tragic events of September 11th.

Plans for the "Responding to Rogue Regimes" conference began a year ago, shortly after the U.S. State Department announced that it would no longer use the term "rogue states" to describe pariah regimes that resort to armed aggression, terrorism or gross violations of human rights to further their agendas. Countries like Afghanistan and Iraq were transformed overnight into mere "countries of concern."

Our idea at the time was to put together a conference that would analyze the continuing threat to U.S. national security posed by these radical regimes, and explore the pros and cons of a variety of options for responding to that threat. Now, post-September 11th, the threat posed by rogue regimes and the terrorist organizations that they support is evident to every American. But the question of how best to respond to rogue regimes is more controversial and significant than ever.

The co-chairs of this Conference, Michael Scharf, who is the Director of our Center for International Law and Policy, and Scott Silliman of Duke Law School, have put together a tremendous line-up of experts and high-level officials for today's event. We are also grateful to the American Society of International Law and the International Law Students Association for co-sponsoring this important conference.

Many of the conference speeches will touch upon the burning issue of the day -- What to do about Afghanistan and Osama bin Laden's al-Qaeda

organization? And we anticipate that there will be many thoughtful proposals hotly debated over the next few hours. The insights of the conference panelists will also be available for a larger audience through the publication this summer of the *New England Law Review's* 2002 symposium issue, containing the remarks, essays, and articles submitted by all of the conference speakers.

This conference will have added importance because it kicks off the work of a Blue Ribbon Committee of Experts on Rebuilding Post-war Afghanistan. This is a joint effort of our Center for International Law and Policy, the American Bar Association's Blue Ribbon Working Group on Terrorism, and the Public International Law and Policy Group. It is being funded as part of a $338,000 grant from the Carnegie Corporation. The committee being established today includes many of the fifteen panelists of this Conference, as well as a handful of other leading experts in the field. It will meet on November 30, 2001, in Washington, D.C. [The Report of the Committee of Experts, which was completed on December 10, 2001, is included at the end of these remarks].

It is becoming clear that the question is no longer "if" the Taliban will fall, but "when." And the bigger question is "what" will replace it? Without a comprehensive plan for rebuilding Afghanistan, the United States may well win the military conflict but fail to achieve its long-term security objectives. The mandate of this new committee will therefore be to help the United States quickly come up with a sound plan to create a stable, neutral, and law-abiding Afghan State. New England School of Law is proud to be part of this effort.

I hope each of you will gain new insight and resolve as you participate in today's exciting and extraordinarily timely conference. Thank you.

Report of the Committee of Experts on Nation Rebuilding in Afghanistan

December 10, 2001

INTRODUCTION

With the destruction of the Taliban regime, the international community is turning its attention toward the establishment of an interim government consistent with the Bonn Accords, and the identification of a process for selecting a more long-term governing arrangement. As is well known, these first steps toward establishing a new government in Afghanistan are the beginning of a long and difficult process for re-establishing peace. Absent a comprehensive and attainable plan for nation rebuilding in Afghanistan, the United States may find that despite its victory on the battlefield, it may be unable to adequately achieve its long term security objectives.

While many groups and organizations were working within the shadow of the Bonn Conference to rapidly consider how best to accomplish a myriad of important objectives, such as reinstating the rule of law, promoting economic development, designing a program of agricultural reform, and creating a transparent form of democratic government, on November 30th, 2001 the Public International Law and Policy Group and the New England Center for International Law and Policy convened a Committee of Experts to consider broader questions which might properly shape the international community's response to nation rebuilding in Afghanistan. In particular, the Committee of Experts examined the opportunity for constructive international community involvement arising from the recent development of contemporary norms of intermediate and evolving sovereignty developed from the recent precedents of Haiti, Bosnia, Kosovo, Sierra Leone, and East Timor, and the utility of creating a framework of conditionality to guide international involvement in Afghan nation rebuilding.

The November 30 meeting of twenty-five prominent foreign policy, military, and international legal experts was held as part of the "Intermediate Sovereignty" Project, which is described at www.intermediatesovereignty.org, and is funded by a grant from the Carnegie Corporation of New York. Participants at the meeting included the former Deputy Assistant Secretary of State for South Asian Affairs, President Clinton's Special Envoy to Yugoslavia, the former Deputy Legal

Counsel at the United Nations, a former legal adviser to the National Security Council, a former Ambassador, the Chairman of the American Bar Association's Section of International Law and Practice, and five former members of the Office of the Legal Adviser of the U.S. Department of State. Also participating were several leading academics, journalists, and experts from the U.S. Institute of Peace, the International Peace Academy, the Coalition for International Justice, the Public International Law and Policy Group, and the American Foreign Policy Council.

This report, written by Professors Michael Scharf and Paul Williams, is the product of the November 30 Meeting of Experts. The report should not be taken to reflect the view of any particular participant on the Committee, all of whom served in their individual capacity.

The report is divided into the following sections: relevant facts, the U.S. interest in participating in Afghan nation rebuilding, general goals for nation rebuilding, specific objectives for nation rebuilding, and the role of the concepts of intermediate sovereignty and phased recognition in achieving these goals and objectives.

I. RELEVANT FACTS

The ability of the international community to structure effective nation rebuilding in Afghanistan will be heavily influenced by the current state of political, economic and civil devastation in Afghanistan, as well as by a number of important internal and external factors. As noted by one participant, any plan for nation rebuilding in Afghanistan must acknowledge the fact that "Afghanistan is a mean place in a rough neighborhood . . .."

Over the past twenty-two years the foundations of the Afghan state have been eroded by near continuous warfare, and periodic drought and famine. It is estimated that as a consequence over two million Afghans have died, and this year over seven million face famine. In addition, two-thirds of the population of twenty-six million are illiterate, and large numbers of children do not live past the age of five. There is also virtually no infrastructure: no paved roads, electricity, or clean water, and there has been no effective central government for the past three decades. The war and the rule of the Taliban have also destroyed many of the basic elements of civil society and good governance.

Compounding the problems created by the sheer destruction of nearly a quarter-century of war, are unique internal and external economic and political factors. Unless the internal and external complications are addressed in tandem, it is unlikely that the primary strategic objectives of the United States and its allies can be achieved.

The internal factors include Afghanistan's lucrative trade in poppies, and the de facto ethnic and political partition of much of the country. Afghanistan currently produces eighty percent of the world's supply of pop-

pies, which bring a profit up to 250 times greater than the profit generated by other crops. The organized crimes and social destruction associated with the poppy export enterprise will challenge even the most basic objectives of nation rebuilding in Afghanistan.

As a result of the ethnic make-up of Afghanistan, its colonial past, the Soviet occupation, and the evolution of traditional modes of governance, most governance in Afghanistan occurs at the local level, where ethnic and tribal political structures dominate the political bargaining process. In fact, in some areas of Afghanistan central authorities have never exercised any effective control. Interestingly, while many outside commentators see this form of decentralized government as a basis for the possible partition or cantonization of Afghanistan, the Afghans themselves express little interest in either option.

The external factors that will influence any nation rebuilding effort relate primarily to the long history of political and economic meddling by Afghanistan's neighbors. From the day of the creation of the Afghan state, its neighbors and other interested states have sponsored proxies within Afghanistan in order to promote their unique interests. To structure a successful nation rebuilding effort, there must be agreement among all the interested states, including Russia, China, Iran, India, Pakistan, Tajikistan and Uzbekistan to cease their efforts for predominate influence in Afghanistan.

Recently, Iran, Uzbekistan, China and Tajikistan have developed common objectives relating to reducing the flow of refugees, curtailing drug shipments, and preventing the export of radical Islam. And in response to the September 11 attacks, most states have heightened concerns relating to the export of terrorism from Afghanistan. A number of important differences remain, however, with Russia vying for a Northern Alliance-dominated government in order to protect its energy-related interests, and Pakistan working toward a Pashtun-dominated government in order to protect its security interests.

There then exists the separate external problem of Pakistan, which will likely soon host more Taliban members and sympathizers than Afghanistan. While Pakistan has played a crucial role in supporting American efforts to defeat the Taliban and destroy al-Qaeda, a number of powerful government institutions are highly supportive of the objectives of the Taliban and al-Qaeda, and favor the transformation of Pakistan into a fundamentalist Islamic state. The Pakistani security service, the Inter-Services Intelligence (which has been described as a state within a state), is of particular concern, as well as certain agencies and personnel involved in Pakistan's nuclear weapons program. Without continued and effective secular reform in Pakistan efforts to rebuild Afghanistan and to stabilize the region will unlikely succeed.

II. THE U.S. INTEREST IN AFGHAN NATION REBUILDING

Some have argued that the United States' primary interest in Afghanistan ends with the destruction of the Taliban and al-Qaeda network, and that other States and international organizations should play the leading role in rebuilding the Afghan state. Although the United States has no affirmative interests in Afghanistan, it will be unlikely to achieve its objectives of destroying and preventing the re-establishment of the al-Qaeda terrorist network, curtailing the export of heroin, and preventing destabilizing political and economic chaos in Afghanistan which could destabilize Pakistan and generate regional instability, unless it plays a far more than a peripheral role in the effort to rebuild Afghanistan. The United States, however, must be realistic in its expectations and should proceed with a set of well-established benchmarks in mind. The approach of the United States should thus be driven by four lessons.

First, the United States is now keenly aware of the consequences of its earlier failure to work toward the establishment of a responsible post-Soviet Afghan regime. While the United States must not in every case of post-conflict or post-peacekeeping assume the primary responsibility for nation building, in the case of Afghanistan it is clear that without a sufficient nation rebuilding effort, Afghanistan may slide back into unmanaged chaos which could again pose a threat to regional security and to the direct security interests of the United States. The United States must also be aware that even with adequate funding and responsible international engagement, there is no guarantee that the nation rebuilding efforts in Afghanistan will be fully successful. The United States must thus choose its objectives carefully.

Second, the United States and its Western European allies uniquely have the military and financial assets needed to successfully rebuild Afghanistan, and to induce its transformation into a responsible non-chaotic state. As experience in the former Yugoslavia teaches, U.N. mediation and peacekeeping efforts, without U.S./NATO military involvement and financial inducements, are almost certain to fail. Similarly, in Sierra Leone, we witnessed the failure of regional troops to restore peace and security, contrasted with the later success of a robust, professional British deployment.

Third, ridding Afghanistan of the Taliban and al-Qaeda will not, in itself, end the threat to America of terrorism and narcotics production posed by Afghanistan. As noted above, the neighboring states cannot be relied upon to take the lead in building a stable, secure, law-abiding Afghan State, since each of these countries traditionally seeks to promote its own strategic interests through its proxies in Afghanistan. In many instances, these interests, be they economic, political or military, conflict with long-term U.S. security interests in the region. Without sustained U.S. political

leadership, and a commitment to non-intervention by neighboring states, there is a real possibility that the unstable situation in Afghanistan will spawn the creation of new terrorists or other destabilizing networks.

Fourth, the American response to nation rebuilding in Afghanistan will be watched carefully by our allies and future potential coalition partners and will influence their willingness to support American intervention in other states where credible threats are identified. It is therefore in the United States' interests to ensure it plays a constructive and leading role, while setting clear and attainable objectives. The United States must also be careful not to support the creation of political structures, such as cantonal forms of government, that may be perceived as prejudging future arrangements or set precedents seen as applicable to American actions in other theaters, such as Iraq.

III. GENERAL GOALS FOR NATION REBUILDING

Establish Non-Chaotic (neutralized) State

The United States and its allies should set as their first general objective the creation of a neutral, stable, secure, and law-abiding Afghan state. The attainment of this objective is a crucial element of America's long-term campaign against terrorism, and is necessary in order to maintain political support for continued U.S. involvement in rebuilding Afghanistan. Without the articulation of this objective there is the real possibility that when the immediate threat from al-Qaeda is eliminated, the American people and Congress will perceive America's mission in Afghanistan to be accomplished, leading to calls for premature disengagement.

The United States Government must be cautious, however, in the extent to which it defines the terms of its commitment to nation-rebuilding and the objectives it seeks to accomplish. While it will be necessary to rebuild the Afghan state to the point where it does not pose a threat to the security interests of the its neighbors or to the United States, the United States must not be pulled into the minutia of nation building, and must not seek to rebuild a type of Afghan state which never before existed and is not possible to create. The United States must also guard against having its military and other resources become political hostages to the interests of its allies, such as in Bosnia and Kosovo where certain troop deployments may no longer be required for military reasons, but nonetheless the continued deployment is deemed to be politically indispensable.

Identify Appropriate Governing Structures

The second general objective must be to adequately consider how best to facilitate the Afghan peoples' ability to establish the most appropriate form of government. While America's European allies will be quick to

push for a form of government similar to a West European democracy, a process must be created whereby the Afghan people are able to craft their own form of government. The difficult task will of course be to determine which entities legitimately express the will of the Afghan people.

The United States must ensure that such deference continues, but must also act to prevent the re-emergence of destabilizing de facto fiefdoms, which may emerge under the rubric of traditional forms of Afghan governance. Any plan to pursue a highly decentralized Afghanistan may in the end present serious challenges to the international community. In the absence of the ruling Taliban, warlords have begun to re-assert control over the Afghan towns and cities. Most of those in power at the local level are not people who can be relied upon to guarantee peace or fairly enforce the rule of law. In the absence of a common enemy, the numerous groups making up the anti-Taliban alliance are likely to pursue their own myopic interests and those of their external sponsors.

Implement Intermediate Sovereignty

The third objective is to establish the level of sovereignty to be attained by Afghanistan. While Afghanistan is a state and a member of the United Nations and thus entitled to sovereignty, territorial integrity, and political independence, in reality it is a less than fully sovereign entity. While a roadmap for full sovereignty should be crafted, the United States and its allies should not rush to restore every last element of full sovereignty until a number of conditions have been met. It is entirely appropriate to treat Afghanistan in a manner somewhat similar to that of Austria in the post war period -- in the case of Afghanistan it would be subject to enforced neutralization in order to prevent its own collapse from undermining its neighbors' security interests. Afghanistan's sovereignty should also be measured in a similar manner to that of post war Germany and Japan -- with significant interim international involvement in the crafting and establishment of governing structures. More contemporary examples of intermediate sovereignty would be Bosnia and East Timor. The recommended conditions for the attainment of full sovereignty and the process for their establishment are set forth in greater detail below.

IV. SPECIFIC OBJECTIVES FOR NATION REBUILDING

The specific objectives that the United States and international community should seek to attain with respect to the new Afghan state should include (in order of priority):

1. Establishment of a neutral, peaceful, transitional government that does not threaten the States in the region. Nineteenth Century Belgium or 1955 Austria would be good historical analogies of what the international community should press for. This is the main objective of the Bonn process, which may unfortunately have begun too late to be entirely successful in its implementation given that events on the ground, rather than rational policy, are largely dictating the make-up of the transitional government. Attaining this goal will also require the negotiation of an agreement of non-interference by all of Afghanistan's neighbors, including a pledge not to provide backing to any Afghan faction or to aspire for predominate influence in Afghanistan.
2. The destruction of al-Qaeda and end of Afghan support for terrorist organizations. In addition to a continuing U.S. military presence actively engaged in the search for al-Qaeda members, this will require ongoing efforts of the new Afghan government since many subversive elements have temporarily gone underground. Members of al-Qaeda that are taken into custody will have to be prosecuted according to international standards of justice. Amnesty for those responsible for war crimes, crimes against humanity, or the September 11 attack against the United States must be strongly opposed. In the long term it may be possible to expand the jurisdiction of the existing international criminal tribunal at The Hague to cover the leading members of al-Qaeda and the Taliban. Importantly, any peacekeeping or policing operation, headed preferably by Turkey or Great Britain, should not have as its mandate the destruction of al-Qaeda. This task should remain within the purview of those forces currently engaged in the combat operation.
3. Suppression of poppy production. Eighty percent of the world's heroin comes from Afghanistan. More American deaths are attributable to Afghan heroin than to the September 11 attacks. Thus, the U.S. has a significant interest in doing whatever is necessary to extinguish the Afghan narcotics trade. This will be extremely difficult, however, as seventy percent of the population of Afghanistan are farmers, and the leading cash crop is poppy production. Unless farmers are adequately compensated to substitute other crops for poppies or presented with other economic opportunities, pursuing this objective through strictly military means will almost certainly undermine support for the fragile transitional government.
4. Establishment of a secure environment, including demilitarization of Kabul and other major cities. Convincing the Afghan people to disarm may prove particularly difficult, but de-mining and de-

commissioning heavy weapons may be attainable under U.N. monitoring. During the interim period, U.N. authorized forces can also play a role as domestic police and border guards. But an international force of West European and Islamic countries with a robust mandate likely will be necessary to ensure a secure environment and prevent renewed fighting between Afghan groups. While some American forces may be required for this operation, they should be deployed only out of military necessity, and not as tokens of political good faith.

5. Peaceful and safe return of refugees. Over two million Afghan refugees are currently in camps in Pakistan and Iran. As in Bosnia, this raises difficult questions about where they should be returned. Obviously moving them into camps in Afghanistan is not the answer, as refugee camps are notorious for producing radicals, revolutionaries, and terrorists. Efforts must be undertaken to rebuild existing communities and restart local economies.
6. Establishment of accountable and transparent governmental institutions and a workable justice system. There is no existing legal framework in Afghanistan. The challenge of creating one from scratch will be even greater than that faced in Rwanda after the 1994 Genocide. But this is an area in which international organizations have a great deal of experience and a fairly good track record, provided adequate financial support of this effort is available. While the United States should play an active role in the development of a strategy and in the funding for a transparent government, it should not divert significant military or civilian personnel into the operation.
7. Recognition and enforcement of human rights and protection of minority rights, including in particular basic rights of women. Since 1995, the Taliban was identified as the worst human rights-violating regime in the world. Women in Afghanistan, who held the majority of professional positions (lawyers and doctors) during the conflict with the Soviet Union, were stripped of their rights under the Taliban regime. Although pursuing this objective may come at the expense of popular support of the new government, the re-integration of women professionals and the protection of women's rights serves the strategic (and moral) interests of the United States; in large part because they have not taken part in the widespread human rights violations attributable to most of the leadership structures of many of the parties represented in the interim government.
8. Institution of anti-corruption procedures and an accountability process, such as establishment of a truth commission and the "lustration" of culpable war lords. As in the case of Bosnia, the

provision of financial assistance in Afghanistan will be subject to corruption and diversion unless culpable war lords are removed from positions of power. The various factions in Afghanistan have serious grievances with one another, stemming from war crimes and crimes against humanity committed throughout the 1990s. A truth commission which documents responsibility and victimization will facilitate reconciliation, assist in "lustration" (removal of perpetrators from positions of authority) and victim compensation, and deter vigilantism. Serious consideration should also be given to an international sponsored regime for the judicial prosecution of those responsible for war crimes and human rights violations.

V. THE ROLE OF THE CONCEPTS OF INTERMEDIATE SOVEREIGNTY AND PHASED

Recognition in Achieving these Goals and Objectives

While traditional concepts of international law tend to reflect a somewhat rigid perspective of sovereignty as an either/or proposition (a territory was either deemed sovereign or it was not), the political definition of sovereignty is more elastic and reflects a history of numerous precedents of colonies, mandate territories, protectorates, trust territories and a history of foreign or international administration of territories. Working from previous state practice and the long history of varied forms of sovereignty, the United States should expressly employ the concept of "intermediate sovereignty," or "transitional sovereignty," in the case of Afghanistan.

Formal adoption of the notion of intermediate sovereignty broadens the range of options open to the United States in its efforts to influence the nation rebuilding process in Afghanistan. The approach of intermediate sovereignty has recently been employed in Bosnia, Haiti, Kosovo, and East Timor.

An associated concept is "phased recognition," under which the international community incrementally bestows the attributes of sovereignty on a territory in return for its compliance with a series of stipulated benchmarks. The transitional administration approach for Afghanistan, which was negotiated at Bonn, gives the international community the opportunity to impose a series of benchmarks upon the transitional government. The attainment of each would be accompanied by certain diplomatic or financial rewards, and conversely the failure to attain them would be accompanied by certain diplomatic or financial sanctions.

Taken together, intermediate sovereignty and phased recognition amount to a form of "soft international trusteeship." Despite its colonial connotations, the term soft international trusteeship adequately describes many contemporary state building enterprises and forms a useful paradigm for understanding the future role of the international community in Af-

ghanistan. Most importantly it also helps to define the limits of international involvement and the notion that such involvement must work toward an end state of full sovereignty for Afghanistan.

To date, one of the defects of "conditional recognition" has been that provision of rewards and imposition of sanctions was an all-or-nothing proposition. Under such circumstances, the states and international institutions controlling the process were practically unable to withdraw recognition, withhold aid or impose sanctions in the face of a series of minor breaches.

Consistent with the concepts of intermediate sovereignty and phased recognition, the carrots and sticks which the international community can employ to induce the transitional Afghan government's adherence to the objectives set forth above include:

(1) Formal recognition of the Afghan government by the United States and other countries. There are a variety of legal benefits that flow from formal recognition of a new government that displaces an existing government after an internal conflict. If the United States formally recognizes the new Afghan government, for example, frozen assets of Afghanistan in the United States will automatically be made available to the new government. The situation in Afghanistan, however, is not as simple as Kosovo, where international recognition was the key objective of the parties. Moreover, this option may be limited in that at the conclusion of the Bonn process, the international community effectively recognized the transitional government as the legitimate Afghan government based on general pledges of cooperation, rather than actual achievement of the objectives set forth above.

(2) Recognition of the Afghan Delegation at the United Nations and other international organizations. In the face of challenges by a competing delegation representing the Taliban regime, since 1995 the United Nations has "provisionally accepted" the credentials of the Rabbinni government as the delegation of Afghanistan on the ground that it represented the last legitimate government. But "provisional acceptance" is not the same as full recognition, which could be used as an inducement to achieve attainment of the above objectives, unless the United Nations prematurely formally recognizes the new transitional government. Even if U.N. recognition is part of the deal that emerges from Bonn, other U.N. affiliated international organizations are autonomous when it comes to recognizing a country's credentials; therefore, recognition in those bodies could still be held out as an award for the government's cooperation.

(3) Provision of IMF/World Bank loans and reconstruction assistance by donor countries. The threatened withholding of large IMF/World Bank loans ultimately induced Croatia to cooperate fully with the Security Council established International Criminal Tribunal for the Former Yugo-

slavia. Similarly, the offer of $1.3 billion in foreign assistance to Serbia conditioned on its surrender of Slobodan Milosevic to the Tribunal proved effective in inducing Serbia's cooperation. Rather than a grand lump sum payment, a series of large loans and foreign assistance grants for Afghanistan, each tied to the attainment of one of the objectives set forth above, would be the most powerful mechanism of attaining the full cooperation of the national and local Afghan authorities. Politically, however, the United States and other major powers are likely to lose interest in Afghanistan before all of the goals have been attained as the war against terrorism turns to other countries. To ensure that adequate financial resources are available, a trust fund could be created for Afghanistan, with release of money over time pegged to attainment of the stated objectives.

CONCLUSION

The lesson to be learned from previous attempts at nation building is not that all such efforts should be eschewed, but that they must be well-funded, subject to enforceable objectives via conditionality of aid and sanctions, and subject to strong and focused leadership. There is no question that rebuilding failed states such as Afghanistan is politically and economically costly and fraught with unforeseen peril. However, if the United States fails to assume a leading role in such a mission it runs the risk of winning the military conflict, but failing to secure a meaningful or lasting peace.

ABOUT THE AUTHORS

Michael P. Scharf (A.B./J.D., Duke) is a professor of Law and Director of the War Crimes Research Office at Case Western Reserve University School of Law, and formally Director of the of the Center for International Law and Policy at New England School of Law. During the Bush and Clinton Administrations, he served in the office of the Legal Adviser of the U.S. Department of State, where he held the positions of Counsel to the Counter-Terrorism Bureau, Attorney-Adviser for Law Enforcement and Intelligence, Attorney-Adviser for United Nations Affairs, and delegate to the United Nations General Assembly and the U.N. Human Rights Commission. In 1993, he was awarded the State Department's Meritorious Honor Award "in recognition of superb performance and exemplary leadership in support of U.S. policy initiatives regarding the former Yugoslavia." Scharf's publications include "Balkan Justice," which was nominated for the Pulitzer Prize in 1998, "The International Criminal Tribunal for Rwanda" which was recognized as the Outstanding Book of 1999 by the American Society of International Law, and casebooks on "The Law of International Organizations" and "International Criminal Law."

Paul R. Williams (Ph.D., Cambridge; J.D., Stanford) is a professor at American University, where he holds a joint appointment in the School of International Service and the Washington College of Law and directs the MA/JD Joint Degree Program. He has served as Attorney-Adviser for European and Canadian Affairs at the U.S. Department of State, as a Fulbright Scholar at the University of Cambridge, and as a Senior Associate with the Carnegie Endowment for International Peace where he directed the Public International Law and Policy Program. He is the Managing Director of the Public International Law and Policy Group and has served as Legal Counsel on the Bosnian Delegation to the Dayton Peace negotiations, the Kosovar delegation to the Rambouillet/Paris peace talks, and the Armenian delegation at the Key West peace talks concerning Nagorno Karabagh. With Michael Scharf, he is the co-author of "Peace With Justice?," soon to be published by Rowman and Littlefield Publishers, Inc. Williams has testified before the Senate Foreign Relations Committee and has appeared frequently as an expert commentator on CNN's "Burden of Proof" and other television and radio news shows.

LIST OF PARTICIPATING EXPERTS

November 30 Meeting on Rebuilding Post-conflict Afghanistan

Mr. Jason Abrams,
Consultant, International Peace Academy
Former Attorney in the United Nations
Office of Legal Affairs

Ms. Nina Bang-Jensen,
Coalition for International Justice

Mr. Mark Bartolini,
International Rescue Committee

Mr. Ilan Berman,
American Foreign Policy Council

Ms. Suzy Blaustein,
Coalition for International Justice

Professor Hiram Chodosh,
Case Western Reserve University Law School

Professor Willie Curtis,
U.S. Naval Academy

Mr. Stephen Del Rosso, Jr.,
Senior Program Officer
Carnegie Corporation of New York

Mr. Edison Dick, Esq.,
American Bar Association Blue Ribbon
Task Force on Responding to Terrorism

Mr. Mike Dziedzic,
United States Institute of Peace

Ms. Nancy Furman Paul, Esq.,
Wilmer, Cutler, & Pickering

Mr. Roy Gutman,
Newsweek

Mr. Jim Hooper,
Public International Law and Policy Group
Former State Department Foreign Service Officer

Ms. Heather Hulbert,
International Crisis Group

Prof. Larry Johnson,
University of California at Davis School of Law
Former Principal Legal Officer at the United Nations Office of Legal Affairs

Mr. Gibson Lanpher,
Former Deputy Assistant Secretary of State for South Asia

Professor Robert Lutz,
Chairman of the American Bar Association Section
of International Law and Practice

Ms. Shibani Malhotra, Esq.,
Dickstein and Shapiro

Ms. Sarah Means,
Washington Times

Prof. Ved Nanda,
University of Denver College of Law

Mr. Jim O'Brien,
Former Special Envoy to Yugoslavia

Mr. Peter Rosenblatt, Esq.,
Heller & Rosenblatt
Former U.S. Ambassador to the Freely Associated States

Ms. Laura Rozen,
United States Institute of Peace

Professor Michael Scharf,
Case Western Reserve University School of Law
Former Attorney-Adviser for U.N. Affairs U.S. Department of State

Mr. John Terzano,
Vietnam Veterans of America Foundation

Professor Geoffrey Watson,
Catholic University School of Law
Former Attorney-Adviser for Law Enforcement and Intelligence and for Near East Asian Affairs U.S. Department of State

Professor Paul Williams,
Washington College of Law at American University
Former Attorney-Adviser for European Affairs U.S. Department of State

Professor David Wippman,
Cornell Law School
Former Director in the N.S.C. Office of Multilateral and Humanitarian Affairs

Rapporteurs

Megan Akers
Washington College of Law

Cassandra Capobianca
Washington College of Law

Patricia Taft
American University

Keynote Address

After September 11th

Ruth Wedgwood*

The common sense of shock has lingered since September 11th. On that infamous morning, I was flying from New Haven to Washington in a small plane, and did not know initially why we were suddenly diverted to the airport in Baltimore. Only on arrival did we learn of the al Qaeda attacks on the World Trade Center in New York City and the Pentagon. We watched the trade center towers collapse, over and over again, on the television news broadcast. The sense of dread was unavoidable. The terrorist career of Osama bin Laden has shown an unremitting ambition over the last decade, and a gargantuan appetite for violence. In 1999, when I looked closely at bin Laden's handiwork,[1] it was apparent that the al Qaeda chief was seeking a nuclear device. On September 11, 2001, no one knew whether this ultimate weapon of mass destruction might already be in his arsenal. The New York City attacks could have killed 25,000 people or more, had the building evacuations not succeeded, and al Qaeda's attack on Washington had other targets beyond the Pentagon. Bin Laden has passed beyond any escalatory threshold, unchecked in his hostility to the West. With full-throated nihilistic violence, bin Laden defies the possibility of compromise.

The events have made plain the fragility of any liberal society. The qualities we take most for granted and cherish most dearly -- such as privacy and freedom of movement and the diversity of a multi-ethnic society -- are the very sources of our vulnerability. Economic globalization, touted as the answer for the next century, has been challenged by a globalized skein of cooperating terrorist networks, skilled at moving men, money, and materiel from region to region, from the Philippines to Afghanistan to Europe and now to North America. Al Qaeda framed its attacks against buildings that stood as symbols of international commerce,

---

* Ruth Wedgwood is the Edward B. Burling Professor of International Law and Diplomacy at the School of Advanced International Studies, Johns Hopkins University, and Professor of International Law at Yale University. She is a former federal prosecutor.

1. *See generally* Ruth Wedgwood, *Responding to Terrorism: The Strikes Against bin Laden*, 24 YALE J. INT'L L. 559 (1999).

as if to trumpet the globalization of violence against the more secular view of a world economy.

Bin Laden and al Qaeda pose unique problems as adversaries. To be sure, dealing with non-state actors is nothing new for veterans of the 1990's. We watched the ethnic communities in the African Great Lakes and the Balkans tear their countries apart, and occasionally we applied the law of war to their activities, through the ad hoc international criminal tribunals. Yet these communities had programmatic demands. The Bosnian Serbs, for example, wanted union with Serbia and contiguous territory for Republika Srpska. But here we have an adversary who does not have a set of worldly demands. Holding fast to martyrdom, al Qaeda members cannot be deterred in an ordinary fashion.

The single-minded violence of al Qaeda is also pernicious because the network has no other competing interests. States usually have other fish to fry. States have a territory and population to protect, as well as commercial relationships. But here we face an adversary who really does not seem to care about earthly things, only about the triumph of an eschatological ideology.

What this means operationally is less clear. We will have occasion to act multilaterally and rally against terrorism. Some of the older treaty projects embarked upon by the U.N. are useful in this effort. The terrorism treaties of the 1970's condemned aircraft hijacking and aircraft bombing and attacks on diplomats, and then used universal jurisdiction and the idea of international crimes to encourage their prosecution. The Terrorism Working Group of the United Nations General Assembly Sixth Committee has lately produced other useful treaties on the financing of terrorism and on terrorist bombing. These were initially intended to address the problems of civil wars such as the fight between the Tamils and Sinhalese in Sri Lanka, fueled by financial booty from expatriates around the world. But the treaties have broader application, against the financial networks that sustain al Qaeda. The U.N. has pressed for these two new conventions, trying to keep fuel away from the fire.

The decentralized structure of al Qaeda is a particular challenge. We will have to rely in large part on police work to root out its scattered groups in Europe, Asia, and elsewhere. With the member countries of the U.N., the G-8, OECD and NATO, we are going to need an active campaign of police work, to unravel the threads and burst the cells of the network.

Bin Laden underestimated American resolve to carry the fight to his center of gravity in Afghanistan. His second strategic error is in placing even the Europeans in doubt of their own safety and stability. The planned attack on the Strasbourg market and the chemical plant explosion in Toulouse have frightened the French. The discovery of cells in Milan, Hamburg, and elsewhere have shown that Europe's tradition of guarding data

privacy and reticent police tactics has been exploited by al Qaeda.

Yet our erstwhile European allies became faint of heart in the 1990's when it came time to counter Saddam Hussein's Iraqi program to develop weapons of mass destruction. The Europeans have commercial ambitions in the region and diverse immigrant populations. One hopes for real-time cooperation in policing, after these earlier defections that served narrower interests.

The great debate at the moment poses a paradigm choice -- under what regime of law should we think about and act against al Qaeda? Should we think of the network in criminal justice terms, as we have throughout the 1990's, or in war and armed conflict terms? Lawyers have a vested interest, some might suppose, in choosing criminal justice. But the law speaks as loudly in armed conflict as it does in criminal justice. The scope of the damage caused on September 11th makes the language of war seem apropos, and this has some important consequences.

To intercept al Qaeda's plans for future attacks, it is crucial that we go beyond retrospective historical reconstruction. Criminal charges look backwards and include only selected individual offenders for the sake of a manageable trial. Criminal justice supposes that deterrence is possible. We need to think about the anticipatory moves that one makes in war. We need, and have newly obtained, the pooling of information between the domestic criminal justice agencies and the intelligence agencies. The Patriot Act now allows domestic criminal justice agencies to share with intelligence agencies both grand jury information and criminal justice wiretaps that often yield important information on network plans.[2] The effective sharing of foreign counterintelligence wiretap information will also be encouraged. We have resisted both in the past. The thought in the 1970's in the aftermath of the Church Committee hearings was to segregate intelligence and military matters from the domestic realm. Intelligence agencies could operate only beyond the water's edge. But on a homeland battlefield, that bright line has faded. When you are fighting a transjurisdictional globalized network, a coastal Maginot line doesn't work. It will take a network to fight a network. If we have the CIA operating overseas and the FBI conducting onshore investigations and they cannot marry up their data, there will be little ability to anticipate the adversary's moves. Following network activities, foreseeing and disrupting al Qaeda's plans, involves matching up thousands of pieces of data -- car rental agreements, hotel leases, student ID's, double endorsements, ambiguous remarks on wiretaps, surveillances of physical movement -- and this cannot be done

2. *See* Uniting and Strengthening America by Providing Appropriate Tools Required to Intercept and Obstruct Terrorism (USA PATRIOT) Act of 2001, § 203, P. L. 107-56, 115 Stat. 272 (2001).

through a single liaison officer in the CIA counter-terrorism center. You have to have the equivalent of joint terrorist task forces, with intelligence people and domestic criminal investigators working side by side to reconstruct in real time the RICO style architecture of these networks. For those reared in civil libertarian households, where intelligence agencies were to be confined and exiled overseas, this is a bit of a startle. But in the exigencies of our time, a rank ordering of relative dangers is necessary. In my own view, it was long overdue to have the sharing provisions of the Patriot Act, and indeed the statute should be further extended to allow some limited sharing with foreign intelligence officials and with state and local officials. The *right* to share also has to be *acted upon* through a changed culture in all affected agencies.

We also need to think creatively about the modalities of trial, if we catch people who are alleged to belong to al Qaeda. In debates after the attacks, I have had a constructive disagreement with my good friends Professor Anne-Marie Slaughter of Harvard and Professor Harold Koh of Yale. America chose to use full dress civilian trials in the Southern District of New York for the first World Trade Center bombing case in 1993 and the East African embassy bombings in 1998. We obtained convictions, but al Qaeda did not slow down. Now we have to be discerning about what the structure of ordinary criminal trials may mean for our efforts to interrupt and intercept future al Qaeda plans. In a federal trial, the government is traditionally expected to have almost open-file discovery. Any evidence offered against the defendant, even from a highly classified source, must be presented in open court. The transparency is desirable in ordinary circumstances, but in the middle of a conflict, the adversary can read the trial record along with the public. This might argue for putting criminal justice aside until the conflict is over and detaining combatants as prisoners of war or unlawful combatants, straight up. In the alternative, we may take a page out of Justice Jackson's book and adapt the forms of trial to the situation. We may wish to consider a liberal version of the military commission, a forum under the law of war used by the Allies in the Second World War to try war crimes in Europe and the Far East. A military commission was used for the trial of German saboteurs captured on Long Island and Florida, German nationals in China, and the Japanese commander in the Philippines, General Yamashita. These commission proceedings were approved by the Supreme Court on three separate occasions.[3]

A military commission allows two things: some latitude in the admission of reliable evidence, and some ability to close limited portions of a

3. *See Ex parte* Quirin, 317 U.S. 1 (1942), Johnson v. Eisentrager, 339 U.S. 763 (1950), and *In re* Yamashita, 327 U.S. 1 (1946).

trial to avoid the gaze of the adversary. Our sensibility is different from the 1940's. In the Supreme Court review of habeas corpus petitions from the commission trials, some justices dissented from perceived deficiencies of the trials. It will require a sensitive brush to adapt the picture of 1945 to our own political and philosophical aspirations. But we must keep in mind that fifty years ago, civil court procedures were also quite different from our present view. We surely should not educate an adversary about the engineering stability of skyscrapers or the source of our wiretaps of al Qaeda operations. In the first World Trade Center trial, the question was asked whether the Trade Center could withstand certain kinds of shocks to its system. We all learned, and al Qaeda learned too, that the Trade Center could withstand the impact of a Boeing 707 jet airplane, but the implication lingered -- a larger plane might exceed the buildings' engineering tolerance.

A greater range in admitting evidence for evaluation is consistent with civilized methods of trial. Both European trials, international trials, and American bench trials admit reliable forms of hearsay. The European doctrine of *mittelbarkeit* says that one can admit hearsay, and critically evaluate its appropriate weight. Eyewitness testimony is preferred, needless to say. But even hearsay evidence may be entitled to some weight. Osama bin Laden is said to have telephoned his mother before September 11th to warn her of the event. Bin Laden's mother surely would not testify. Yet, say that she told a good friend about the call. In federal district court the jury would not get to hear of the matter at all, even where the friend is willing to testify and has no apparent reason to distort. The Anglo-American system has distrusted the capacity of juries to give appropriate weight to varied forms of evidence, but other tribunals would permit its consideration. The Federal Rules of Evidence are not the only model of fairness.

Our ability to adapt civilian courts to handle these practical problems is limited by three decades' constitutionalization of civilian criminal procedure. Beginning in the 1960's, the Supreme Court embarked on a long project of reforming federal and state criminal procedure, using the due process clause of the Fifth and Fourteenth Amendments as the instrument. The reforms may be judged a success for ordinary cases. But the use of the Constitution as a vehicle for change now means there is less sure room to adapt American civilian trials to these radically changed circumstances. At a minimum, it would warrant the consideration of the Congress to judge whether federal district court trials can be adapted to the felt problems of intelligence protection and wider-ranging admission of available evidence, within the bounds of our civil constitutional norms. For the moment, the older venue of military commissions, authorized by the Congress in Title 10 and long serving as the trial venue for war crimes, provides an alternative forum.

Apart from trials, we need to understand the limits of criminal justice as an alternative to national security decisionmaking, when it comes to planning our responses. The Washington Post published an extraordinary report on October 3rd in an article by Bart Gellman -- asserting that there was a bypassed chance to seize Osama bin Laden in 1996.[4] The Khartoum government was reportedly willing to force bin Laden to leave his lair in Sudan, and hand him over either to America or the Saudis. We did not push the Saudis to take him captive. We did not try to intercept his airplane. Rather, we let bin Laden fly to Afghanistan where he reestablished his enterprise with the visible results of September 11th. The reason for this lapse was the misapplication of criminal justice standards. The director of the Federal Bureau of Investigation reportedly advised the White House that there was not yet a fully-formed criminal case against bin Laden and hence no way to detain him.[5] The White House wrongly assumed that criminal justice was the only standard by which one could act against an adversary, and let bin Laden fly off to Afghanistan where he was able to operate with impunity.[6] But self-defense and military rules of engagement are different from civilian criminal court. One need not have trials before self-defense. The National Command Authority does not report to the FBI.

In the 1998 response to the Embassy bombings, the United States used Tomahawk missiles to attack bin Laden's camp at Khost, Afghanistan, and a pharmaceutical plant in Khartoum that was believed to have links to bin Laden and the manufacture of chemical weapons. My colleague Jules Lobel of the University of Pittsburgh School of Law has argued in debate that it was improper to use military force until the criminal investigation was complete and that all justificatory evidence should be made public.[7] In my view, this is a highly impractical standard, when faced with a dangerous foe and continuing threat.

On the military side of things, we have some hard choices. The United States has the largest military budget in the world – as high as 3.5 percent of our Gross Domestic Product. The Europeans consistently lag behind – generally limiting their commitment to international security to amounts that range from 1.5 to 2 percent. In the debate on United Nations dues,

---

4. Barton Gellman, *U.S. Was Foiled Multiple Times in Efforts to Capture Bin Laden or Have Him Killed; Sudan's Offer to Arrest Militant Fell Through After Saudis Said No*, WASH. POST, Oct. 3, 2001, at A1.

5. The former United States Attorney for the Southern District of New York, Mary Jo White, was not informed about the Sudanese offer.

6. *See* Ruth Wedgwood, *The Law at War: How Osama Slipped Away*, NAT'L INT., Jan. 1, 2001, at 69.

7. *See generally* Jules Lobel, *The Use of Force to Respond to Terrorist Attacks: The Bombing of Sudan and Afghanistan*, 24 YALE J. INT'L L. 537 (1999).

American policymakers could not help but notice that we paid the lion's share of NATO expenses, the deterrent that provides authority to many U.N. Security Council resolutions.

But even our military might is challenged by the task that we have at hand. The Pentagon began the current Quadrennial Defense Review asking how one could both modernize the force, go into space and cyberspace, and rejuvenate old equipment. The proposition was seriously entertained before September 11th that we might demobilize two active army divisions, four reserve divisions, a marine expeditionary force, and a carrier battle group. The civilian leadership wisely polled the views of the Joint Chiefs of Staff who argued that the operations tempo of military personnel was already at the breaking point. With the future that now lies ahead, with a limited land conflict in Asia, any idea of going below ten ready divisions is even more questionable. So in the future, without significant increases in spending, what do you give up? Modernization? Going into space? Or the chance to renew aging equipment? None of these alternatives is attractive. Liberals and conservatives alike have a duty to consider that, even with our philosophy of limited government, it may be necessary to make a substantially higher investment in international security over the long term. For many decades the U.S. has been the keystone in the architecture of deterrence in Asia, the Middle East, and Europe. This will not change soon.

The American academy and political leadership also have a separate burden, in particular, a short-term and long-term engagement with Islam. In my mother's expatriate girlhood, the primary historical event taught in French schools was the victory of Charles Martel at the Battle of Tours, beating back the Saracens so they could not pass beyond the Pyrenees, saving France for Christianity. This was a central event in French self-identity. Even now the Europeans have not liked to recognize how intermingled they are and always have been with Islam. The glory of Cordovan Spain was as a center of Islam. The Omayyeds of Spain were not a peripheral suburb of Middle Eastern Islam, but the powerhouse of Islamic scholarly thought. Abd ar Rahman the Third made Cordova an intellectual center of Europe for all scholars -- Christian, Jewish and Islamic alike.[8] The Islamic scholar Averroes was recognized by Thomas Aquinas as a crucial philosopher in the one truth of God. Yet the embrace between Islam and the West now seems to have fallen by the side. The fault lies with both parties, dare I say, but one part of it is the adoption of an almost truculent closed identity by Europe, as well as the decay of Islamic-Arab intellectual independence.

---

8. Maria Rosa Menocal, *A Golden Reign of Tolerance*, N.Y. TIMES, Mar. 28, 2002, at A31.

Europeans uneasily recall the role of the Turks in Central Europe. The Ottomans were at the gates of Vienna in 1683, and lingered in Budapest until 1686. The Ottoman Empire was a key actor in the Balkans throughout the nineteenth century (hence the Serb sobriquet that all Bosnian Muslims are "Turks"). With the collapse of the Ottoman Empire in World War One, the Allies were inclined to punish Turkey by carving it up, initially intending to neutralize the Dardanelles, consign the Aegean coastline to the Greeks and Italians, and leave Turkey as a pastoral upland plateau. This plan, of course, was countered by the resistance of Mustafa Kamal, and followed by ethnic bloodletting in the coastal city of Smyrna that prolonged Greek-Turkish antipathy long after the exchange of populations under the Treaty of Lausanne. The antipathy has not yet given way, despite Turkey's charter membership in NATO and its key alliance with the West in the Cold War. Turkey is a secularized administrative state. It has resisted Islamist fundamentalism. It has a close military relationship with Israel. Yet, even now, the self-identity of Europe as Christian marks the divide, even while individual Muslims have emigrated and make up large segments of Europe's population. A strange and unhappy resistance to Turkey's candidacy for the European Union has continued, even while Cyprus has been proposed for membership. Though there are sharp divisions within the Muslim world, still the example of Europe's inhospitality strikes home.

A cure for the current antipathy between the West and Islam also depends on Muslim intellectuals and political leaders, both at home and in Europe. It is startling to see that even after the events of September 11th, a whole generation of European-Arabs has remained silent. Living and studying in Europe for decades, nonetheless few have felt able or willing to take it as their own task to confront the militant brand of Shiite and Wahabi fundamentalism that so rudely attacks any embrace of modern civic nationalism or secular culture.

But the divide is our own challenge as well. Rather than seeing the study of Islam as the recondite specialty of a few scholars such as Bernard Lewis and Fouad Ajami and an occasional professor of religion, we need to understand enough to speak in resonant terms to a generation of Arab students and Turkish students and Persian students. We need to revive the study of Islam, both its historical engagement with the West, and though it might be presumptuous to try, the sources and arguments of Islamic theology. Our students should be able to speak intelligently about the world view of militant Islam and its violent account of jihad, as well as the Islamic world view that is more pacific and engaging. Though moderate Islamic intellectuals may feel intimidated at the moment, even with greater courage, they need to have a cohort within which to revive a different conversation. They will need the indirect support of the West. And we also have the duty of speaking directly to the Arab street. We cannot

leave it to Saudi-funded charities to provide elementary education in Pakistan, but rather, need to find a way to provide secular education and scientific education to youngsters in Egypt, Pakistan, Afghanistan, and elsewhere, that does not leave them captive to mullahs who are destructive in their plans.

Some proposed cures remain wishful. Money alone is not the answer, for Egypt has shared generously in the five billion dollar annual aid package used to sustain the Camp David accords. The Saudi economy is not as wealthy as once it was, but Saudi Arabia's oil reserves could provide a regional Marshall Plan; many Muslim countries share resentment at Saudi unwillingness to invest in the region. Money alone is not going to change the governments that are undemocratic, that do not try to develop their economies, and that do not embrace their disaffected youth through any other mechanism than sheer repression.

America needs to be hard-headed and act robustly in self-defense, especially when the stakes are so high. But at the same time, we may address an intellectual challenge to our cultural institutions and also to foreign governments. For the sake of our own breadth of view, as well as for self-protection, we need to understand all the roots of our civilization, not only within Judaism and Christianity but within Islam as well.

International Law and Rogue States: The Failure of the Charter Framework

Anthony Clark Arend*

The concept of "rogue state" has been much maligned in the past. Appearing to some to be nothing more than an ideological tag with no content, the term was being eased out of much official discussion. As is well noted, the Department of State had changed its designation from "rogue states" to "states of concern." But in the wake of the terrorist attacks of September 11, 2001, the idea of a "rogue state" seems to have returned. Indeed, in his State of the Union Address in January 2002, President George W. Bush went so far as to place three states -- Iran, Iraq, and North Korea -- in an "Axis of Evil."[1] Without going to that extreme, the moniker "rogue state" nonetheless does seem to be useful for understanding the behavior of certain regimes in light of international law. In a sense, a rogue state is one that does not follow the rules. While almost all states may at times violate international legal rules, a rogue state would be a perennial violator. It would, in essence, be a state whose identity is to some extent defined by acting outside of the standard rules of international law.[2] But from the perspective of international jurisprudence an even greater problem is that while as such a state may be violating well-established international legal rules, it may also be engaging in behavior that was not anticipated when the rules were developed. In particular, a rogue state may use military force in ways that classic international law did not address. And here is the problem for international law. Since the rules developed to respond to certain classic uses of force, it is often unclear how the traditional rules of international law apply to the behavior of rogue states. How does international law respond to actions by

---

* Professor of Government and Adjunct Professor of Law, Georgetown University. The author wishes to thank Robert J. Beck for his collaboration on these issues over the years and Michael J. Glennon for his recent work on the use of force. This article draws upon Professor Glennon's approach to these crucial legal questions.

1. George W. Bush, State of the Union Address (January 29, 2002) *available at* http://www.whitehouse.gov/news/releases/2002/01/20020129-11.html ("States like these [Iraq, North Korea, and Iran], and their terrorist allies, constitute an axis of evil, arming to threaten the peace of the world").

2. An analogy could be drawn to a criminal organization. The very identity of such an organization is tied to the fact that it is acting outside established legal norms.

rogue states that fall outside the parameters of traditional international law? If a rogue state engages in covert actions against another state, supports rebels in another state, supports the work of terrorist groups, commits human rights abuses against its own citizens, or develops weapons of mass destruction, how does the existing international legal regime respond?

The purpose of this article is to explore several of the challenges posed by rogue states relating to the recourse to force and to explore some efforts to respond to these challenges. In order to do this, this article will do four things. First, it will discuss the United Nations Charter framework for the use of force -- the most widely accepted legal paradigm for contemporary international law relating to the recourse to force. Second, it will explain several problems that rogue states pose to that framework. Third, it will explore efforts by the international community to respond to those problems, arguing that the international community has not yet developed a clear approach to the challenges of rogue states. Fourth, it will conclude that the challenges posed by rogue states are merely a part of the challenges posed to the Charter framework by the community of states at large and suggest that efforts need to be made to establish a general rule-based regime for the use of force.

I. THE UNITED NATIONS CHARTER FRAMEWORK FOR THE RECOURSE TO FORCE

Over the years, customary international law has come to recognize two elements of the international law dealing with the use of force: the *jus ad bellum* and the *jus in bello*. The *jus ad bellum* deals with the legal rules concerning the recourse to force.[3] It asks the question of when an actor has the right to take recourse to military force against another actor. Can, for example, Iraq lawfully invade Kuwait? Can other states use force to respond to that invasion? Can the United States use force against Iraq for its non-compliance with the weapons inspections regime established by the United Nations Security Council? The *jus in bello* contains the rules dealing with the conduct of hostilities. It explores the rules concerning how all actors in a conflict have to behave, irrespective of who had the right to initiate hostilities in the first place. Thus, the *jus in bello* asks questions such as: what types of targets are permissible in combat? How are non-combatants to be treated? How are prisoners are war to be cared for by their captors? While both sets of rules would apply to the behavior of rogue states, this article

3. For a discussion of the law relating to the recourse to force, see generally ANTHONY CLARK AREND & ROBERT J. BECK, INTERNATIONAL LAW AND THE USE OF FORCE: BEYOND THE U.N. CHARTER PARADIGM (1993); CHRISTINE GRAY, INTERNATIONAL LAW AND THE USE OF FORCE (2000).

will focus on the *jus ad bellum*– the legal rules that address the question of when an international actor can permissibly use force.

Since 1945, most discussions of the law relating to the recourse to military force have begun with an exploration of the United Nations Charter framework. Coming out of two devastating world wars, the founders of the United Nations sought to establish an extremely restrictive approach to the use of force. To do this, the Charter begins with a fundamental presumption against the recourse to force. Article 2, paragraph 4 of the Charter provides:

All Members [of the United Nations] shall refrain in their international relations from the threat or use of force against the territorial integrity or political independence of any state, or in any other manner inconsistent with the Purposes of the United Nations.[4]

This provision seemed to constitute a basic proscription on both the actual use and the threat to use force against another state. It is even more restrictive than the much-criticized Kellogg-Briand Pact of 1928,[5] which only prohibited "war" not all uses and threats of force.

In the Charter, there are only two explicit exceptions to this Article 2(4) prohibition that are still applicable: self-defense, and force authorized by the United Nations Security Council. The first exception is contained in Article 51, which provides in part:

Nothing in the present Charter shall impair the inherent right of individual or collective self-defence if an armed attack occurs against a Member of the United Nations, until the Security Council has taken measures necessary to maintain international peace and security.[6]

Thus, if a state is experiencing an "armed attack," it is permissible for that state to use military force against the states that have instituted that attack. The victim state can also call upon other states to help it fend off the attacker in collective self-defense. So, for example, when Argentina attacked the Falkland Islands in 1982,[7] the United Kingdom responded in self-defense in response to an armed attack and could have asked other states for assistance in this effort.

The second exception to the Article 2(4) prohibition is contained in other portions of Chapter VII of the Charter. Under Article 39, the Security Council is empowered to determine if there has been a "threat to the peace,"

---

4. U.N. CHARTER, art. 2, para. 4.

5. Kellogg-Briand Pact, Aug. 27, 1928, Stat. 46, 2343, T.S. No. 796, 94 L.N.T.S. 57.

6. U.N. CHARTER, art. 51.

7. On the Falklands War, see generally LOWELL F. GUSTAFSON, THE SOVEREIGNTY DISPUTE OVER THE FALKLAND (MALVINAS) ISLANDS (1988); THE FALKLANDS WAR: LESSONS FOR STRATEGY, DIPLOMACY, AND INTERNATIONAL LAW (Alberto R. Coll & Anthony C. Arend eds., 1985).

a "breach of the peace," or an "act of aggression."[8] If the Council so determines, it is further authorized under Article 41[9] to impose diplomatic and economic sanctions and under Article 42[10] to impose military sanctions. This is precisely what happened in the Gulf War.[11] After the invasion of Kuwait on August 2, 1990, the United Nations Security Council met and immediately condemned the Iraqi invasion.[12] Soon thereafter, the Council adopted resolution 661,[13] which imposed sweeping economic and diplomatic sanctions on Iraq. Ultimately, the Council decided to adopt resolution 678,[14] authorizing states with troops deployed in the region to use force to remove Iraq from Kuwait.

II. ROGUE STATES AND THE CHARTER

The Charter framework seems quite well suited to address the types of force with which the framers of the Charter would have been most familiar -- clear, overt acts of aggression undertaken by regular military forces of states. No doubt, these were the kinds of uses of force that the framers were

---

8. Article 39 provides:

> The Security Council shall determine the existence of any threat to the peace, breach of the peace, or act of aggression and shall make recommendations, or decide what measures shall be taken in accordance with Articles 41 and 42, to maintain or restore international peace and security.

U.N. CHARTER., art. 39.

9. *Id.* art. 41.

10. Article 42 provides:

> Should the Security Council consider that the measures provided for in Article 41 [diplomatic and economic sanctions] would be inadequate or have proven to be inadequate, it may take such action by air, sea, or land forces as may be necessary to maintain or restore international peace and security. Such action may include demonstrations, blockade, and other operations by air, sea, or land forces of Members of the United Nations.

Id. art. 42.

11. For a discussion of international law and the Gulf War, see Oscar Schachter, *United Nations Law in the Gulf Conflict,* 85 AM. J. INT'L L.452, 452 (1991).

12. *See* U.N. S.C. Res. 660, U.N. SCOR, 45th Sess., 2932nd mtg., U.N. Doc. S/RES/660 (1990). .

13. U.N. S.C. Res. 661, U.N. SCOR, 45th Sess., 2922rd mtg.,U.N. Doc. S/RES/661 (1990). .

14. Resolution 678 authorized:

> Member States co-operating with the Government of Kuwait, unless Iraq on or before 15 January 1991 fully implements, as set forth in paragraph 1 above, the above-mentioned resolutions, to use all necessary means to uphold and implement resolution 660 (1990) and all subsequent relevant resolutions and to restore international peace and security in the area.

U.N. S.C. Res., U.N. SCOR, 45th Sess., 2963rd mtg., U.N. Doc. S/RES/678 (1990).

envisioning when the Charter was written in San Francisco in the spring of 1945.[15] The world had just been shaken by the aggressive actions of Germany, Japan, and Italy. Under the Charter system, such actions would be clearly prohibited, and states would be empowered to respond, either in self-defense or as authorized by the Security Council. The problem is that most of the uses of force that have taken place in the world since 1945 have not fallen easily into the neat definitions provided by the framers of the Charter. In particular, many of the challenges posed by rogue states are not clearly addressed by the Charter framework. Three challenges seem most acute. First, rogue states often do not engage in obvious armed attacks, but instead use force against other international actors in ways that do not squarely fall under the textbook definition of "armed attack." Second, rogue states frequently engage in actions against their domestic populations that do not constitute international aggression, but may nonetheless be horrible violations of international human rights law. Third, rogue states may possess or seek to possess or develop weapons of mass destruction. While the mere possession of these weapons, or potential possession of these weapons, does not constitute an actual armed attack, such actions by the rogue states may still threaten international peace.

The Armed Attack Threshold

When a state launches a conventional military attack against another state -- when armies cross the borders, bombs are dropped on the territory of a state, or naval vessels fire cruise missiles -- an armed attack has occurred. Hence, the Iraqi invasion of Kuwait clearly was an armed attack that engendered a right of self-defense under Article 51. But, while rogue states like Iraq are not strangers to this type of action, many of the uses of force undertaken by such states do not fall under the standard definition of armed attack. Several such uses of force include: covert action, intervention in support of rebels, and, perhaps most notably, various forms of support for terrorist actions.

Covert Action and Support of Rebels

A covert action is an act by a state in which it attempts to influence the political, economic, or other conditions of the state without that action being known. While commentators often only think of the most dramatic types of covert actions -- such as efforts to assassinate state leaders -- covert actions run the spectrum from the least forceful, such as propaganda, financial support of political parties, bribing of government officials, to the most force-

15. For a discussion of the United Nations Charter, see generally THE CHARTER OF THE UNITED NATIONS: A COMMENTARY (Bruno Simma ed., 1994).

ful, such as blowing up military installations in another states and forcible actions against regime elites. From the perspective of international law, the critical question is under what circumstances such actions would rise to the level of an armed attack.

A similar problem presents itself when a state intervenes -- either covertly or overtly -- in an on-going civil conflict to support rebel groups that are challenging the existing regime. This type of involvement in internal conflict has occurred repeatedly during the period since World War II. During the Nicaragua conflict in the 1980's, for example, the United States accused the Sandinista government of Nicaragua of providing various types of financial and other support to the Farabundo Marti de Liberacion Nacional (the "FMLN") rebels in El Salvador.[16] While such actions would constitute impermissible interference in the internal affairs of a state, when does assistance– money, arms, training, logistical support, etc. – reach the point of being tantamount to an armed attack? When could the aggrieved state take arms against the intervening state?

A. Terrorism

In light of the September 11th attacks and the response of the U.S. and its allies, one of the most pressing questions regarding the armed attack threshold relates to terrorism. While efforts to "define" terrorism have been undertaken numerous times over the years,[17] terrorism will be defined here as *the threat or use of force, undertaken for political purposes, by a non-state actor that intentionally targets civilians and other non-combatants in violation of existing law relating to the conduct of hostilities*. Hence, actions such as the attacks on the World Trade Center would clearly meet the definition of terrorism. Not surprisingly, rogue states are frequently involved in supporting, in one way or another, the efforts of terrorist groups. During the recent action in Afghanistan, for example, the Taliban regime was found by the Security Council to have provided support to the Al Qaeda terrorists in their efforts against the United States and other targets throughout the world.[18]

---

16. For an exploration of the American position, *see generally* JOHN NORTON MOORE, THE SECRET WAR IN CENTRAL AMERICA (1987); ROBERT F. TURNER, NICARAGUA V. UNITED STATES: A LOOK AT THE FACTS (1987).

17. *See* Walter Gary Sharp, Sr., *The Use of Armed Force Against Terrorism: American Hegemony or Impotence?*, 1 CHI. J. INT'L L. 37, 37-39 (2000) ("Although the international community began a concerted effort to control international terrorism in the late 1920s, it has never been able to agree on a definition of international terrorism.") (footnote omitted).

18. In Resolution 1390, the Council condemned

> the Taliban for allowing Afghanistan to be used as a base for terrorists training and activities, including the export of terrorism by the Al-Qaida network and other terrorist groups as well as for using foreign mercenaries in hostile actions in the territory of Afghanistan . . . .

But once again, the interesting legal question is: when does state support to terrorists rise to the level of an armed attack under Article 51? Clearly, there are a number of different levels of support that states could be providing to terrorists. At the high end of the spectrum, states could be directing and controlling terrorist groups. In other words, states could essentially be using terrorists as proxies in the international system. This level of state involvement might be termed *state sponsored terrorism.*[19] At a lower level, states could be providing financial and military assistance to terrorists. This level could be termed *state supported terrorism.* And at yet a lower level, states could be tolerating and otherwise harboring terrorists. This would be a circumstance where "a state does not sponsor or support terrorist groups within its borders, but knows of their existence and fails to suppress them."[20] This level could be termed *state tolerated terrorism.* At what level would state involvement in terrorism rise to the level of an armed attack? Clearly, different observers could draw the line at different places.

B. The Internal Use of Force

While rogue states may very well become pariahs because of their actions abroad, such states may also engage in the use of military force against their own citizens. The notorious actions of the Khmer Rouge in Cambodia, the horrors of Rwanda, and the ethnic cleansing of the Serbian regime immediately come to mind. Prior to World War II, international law, did not address the matter of how a government treated its own people. It was not until the abomination of the Holocaust that the international community decided to make human rights norms a major part of international law. In 1945, the drafters of the U.N. Charter proclaimed their determination "to reaffirm faith in fundamental human rights, in the dignity and worth of the human person, in the equal rights of men and women . . . ."[21] Following the adoption of the Charter, the United Nations proceeded to adopt a host of instruments dedicated to the establishment of an elaborate regime of international human rights law. These include the Genocide Convention,[22] the Covenant on Civil and Political Rights,[23] and the Covenant on Economic,

---

U.N. S.C. Res. 1390, U.N. SCOR, 54th Sess., 4452nd mtg., U.N. Doc S/RES/1390 (2002)..

19. This follows the categories suggested in AREND & BECK, *supra* note 4,at 142.

20. *Id.* (footnote omitted).

21. U.N. CHARTER, preamble.

22. Convention on the Prevention and Punishment of the Crime of Genocide, Dec. 9, 1948, 78 U.N.T.S. 277 [hereinafter Genocide Convention].

23. International Covenant on Civil and Political Rights, G.A. Res. 2200A, U.N. GAOR, 21st Sess., Supp. No. 16, at 1, U.N. Doc. A/6546 (1966), *reprinted in* 999 U.N.T.S. 171 [hereinafter Covenant on Civil and Political Rights].

Social, and Cultural Rights.[24] When states violate applicable human rights law , they are clearly breaching their international legal obligations, yet they are not engaging in an armed attack against another member state. If the Security Council determined that such abuse of human rights would constitute a threat to international peace and security, and subsequently, authorized the use of force under Article 42 of the Charter, military action against the state would seem to be lawful. But in the absence of Security Council action, the Charter framework would seem to outlaw forcible action by states against the offending regime, even if the purpose of such force was to prevent massive violations of human rights.

C. Weapons of Mass Destruction

When the Charter was drafted in the spring of 1945, only a few people were aware of the potential power of the atom. Consequently, the Charter framework was, as one advisor to the United States delegation, John Foster Dulles, would later say, a "pre-atomic" document.[25] While chemical and biological weapons were certainly known to the delegates at San Francisco, their use as "weapons of mass destruction" was probably not on the minds of the Charter framers. After all, the use of chemical weapons in the First World War had generally been regarded as a strategic mistake and, indeed, such use had been prohibited by the 1925 Geneva Gas Protocol.[26] But in the world after Hiroshima and Nagasaki and with increasing evidence of the production and stockpiling of various chemical and biological agents, the unique threat posed by weapons of mass destruction presents a particularly acute problem in the wake of rogue regimes.

At first blush, the problem may seem unremarkable. These weapons, like any others, would be subject to standard laws of armed conflict. Several conventions prohibit the possession and use of chemical and biological weapons, and while the use of nuclear weapons is not clearly prohibited under customary international law,[27] the use of such weapons in an indis-

---

24. International Covenant on Economic, Social and Cultural Rights, adopted Dec. 21, 1966, 993 U.N.T.S. 3 [hereinafter Covenant on Social and Cultural Rights].

25. *See,* ANTHONY CLARK AREND, PURSUING A JUST AND DURABLE PEACE: JOHN FOSTER DULLES AND INTERNATIONAL ORGANIZATION 231 (1988).

26. Protocol for Prohibition of the Use in War of Asphyxiating, Poisonous or Other Gases, and of Bacteriological Methods of Warfare, June 17, 1925, 26 U.S.T. 571, 94 L.N.T.S. 65 [hereinafter Geneva Gas Protocol].

27. The International Court of Justice in its advisory opinion on *The Legality of the Threat or Use of Nuclear Weapons* declined to rule that the use of nuclear weapons would be unlawful in all circumstances. Legality of the Threat or Use of Nuclear Weapons, 1996 I.C.J. 226 (July 8, 1996). For a discussion of the case, *see generally* Dale Stephens, *Human Rights and Armed Conflict -- The Advisory Opinion of the In-*

criminate way or in a manner that causes unnecessary suffering would be proscribed.[28] But the problem with rogue regimes lies at a deeper level. If certain states have a history of violating acceptable norms of international behavior, how should the international community respond to any efforts on the part of these states to develop such weapons of mass destruction? Following the Gulf War, the United Nations Security Council was so troubled by the prospect of Iraq acquiring weapons of mass destruction, that it imposed strict disarmament requirements and ordered Iraq to destroy chemical, biological, and nuclear weapons facilities and submit to an inspection regime.[29] Since Iraq was a party to both the Nuclear Non-Proliferation Treaty and the 1972 Biological Weapons Convention,[30] it was already prohibited from possessing nuclear and biological weapons. At the time of the Security Council action, however, there was no explicit prohibition on the *possession* of chemical weapons.[31] Iraqi compliance with these requirements has been less than satisfactory and on several occasions, the United States and some of its allies have used force against Iraq to secure compliance with the inspections regime.[32]

These uses of force raise the question that weapons of mass destruction pose for the Charter framework: Does the mere possession of weapons of mass destruction present such a threat to international peace and security that states can unilaterally intervene to abate this threat? If so, at what point, can states take such action? In 1981, Israel bombed the Osarik Reactor in

---

ternational Court of Justice in the Nuclear Weapons Case, 4 YALE HUM.RTS. & DEV. L.J. 1 (2001).

28. The Hague Convention on the Laws and Customs of War on Land prohibits the use of "arms, projectiles, or material calculated to cause unnecessary suffering." Convention (no. IV) Respecting the Laws and Customs of War on Land, with Annex of Regulations, Oct. 18, 1907, 36 Stat. 2277, T.S. No. 539, 1 Bevans 631, Annex, art 23, e.

29. *See* U.N. S.C. Res. 687, U.N. SCOR, 46th Sess., 2983rd mtg., U.N. Doc. S/RES/687 (1991).

30. Convention on the Prohibition of Development, Production, and Stockpiling of Bacteriological (Biological) and Toxin Weapons and on Their Destruction, Apr. 10, 1972, 26 U.S.T. 583, 1015 U.N.T.S. 163.

31. The Chemical Weapons Convention was not signed by the U.S. until 1993 and did not enter into force until April 29, 1997. *See* Convention on the Prohibition of the Development, Production, Stockpiling and Use of Chemical Weapons and Their Destruction, Jan. 13, 1993, S. TREATY DOC. NO. 21, (1993), 1974 U.N.T.S. 3.

32. The United States used force in 1993 and again in 1998 to secure Iraqi compliance with the weapons inspections regime established by the Security Council under Resolution 687. For a discussion of the lawfulness of these actions, see generally Jules Lobel & Michael Ratner, *Bypassing the Security Council: Ambiguous Authorizations to Use Force, Cease-Fires and the Iraqi Inspection Regime,* 93 AM. J. INT'L L. 124 (1999); Nigel D. White & Robert Cryer, *Unilateral Enforcement of Resolution 687: A Threat too Far,* 29 CAL. W. INT'L L. J. 243 (1999).

Iraq, claiming that the reactor was to be used to produce weapons-grade material that would produce weapons that would be used to threaten Israel. While the members of the Security Council, including the United States, condemned Israel for its actions, the history of the Iraqi nuclear weapons program and its regular rejection of arms inspectors, raise similar questions today.

III. ADDRESSING THE PROBLEMS POSED BY ROGUE STATES

In light of these challenges that rogue regimes pose for the United Nations Charter framework, scholars and international actors have attempted to apply the basic Charter principles to these problems not anticipated by the framers of the Charter. Accordingly, a series of arguments have been developed about the questions posed in the section above.

A. The Armed Attack Threshold

1. Covert Action and the Support of Rebels

In the scholarly literature, there seems to be general agreement that both covert actions and support of rebels can rise to the level of an armed attack, thus engendering the right of self-defense. The difficulty is determining precisely at what point the armed attack threshold is reached. The most instructive case on this matter is the *Nicaragua* case. As noted earlier, in that case, the United States alleged that Nicaragua was providing various forms of support to the FMLN, a rebel group in El Salvador. While the International Court of Justice did not find any evidence of assistance from the Government of Nicaragua to the FMLN, the court did comment on the circumstances under which assistance to rebels would rise to the level of an armed attack. It explained that:

> [I]t may be considered to be agreed that an armed attack must be understood as including not merely action by regular armed forces across an international border, but also "the sending by or on behalf of a State of armed bands, groups, irregulars or mercenaries, which carry out acts of armed force against another State of such gravity as to amount to" (*inter alia*) an actual armed attack conducted by regular forces, "or its substantial involvement therein."[33]

This was a crucial statement by the court for two reasons. First, it was the first judicial acknowledgment that "indirect aggression" could rise to the

33. Military and Paramilitary Activities (Nicar. v. U.S.), 1986 I.C.J. 14 (June 27) para. 195 [hereinafter Nicaragua Case].

level of an armed attack. At some point, the support for rebels could be so great that it would justify a forcible response in self-defense against the intervening state. Second, the court's decision was also a comment on the point at which support would be tantamount to an armed attack. For the court, this occurred only in sending in "armed bands, groups, irregulars or mercenaries" that use force at a level equivalent to an armed attack by regular military forces. Despite American arguments to the contrary, the court explicitly rejected the notion that anything short of sending in such armed bands would ever rise to the level of an armed attack. The court noted that while providing arms and other assistance to rebel groups may indeed violate Article 2(4),[34] the court did "not believe that the concept of 'armed attack' includes not only acts by armed bands where such acts occur on a significant scale but also assistance to rebels in the form of the provision of weapons or logistical or other support."[35]

This second conclusion by the court has raised a great deal of scholarly concern.[36] The court seems to suggest that a state could be in a most untenable position. An outside state could be providing significant military, logistical, financial, and other assistance to rebels within the territory of a state. While such assistance might violate the Charter provisions on the use of force, as long as assistance did not actually involve the sending in of armed bands, the victim state would be unable to respond in self-defense. In essence, an outside state could do as much damage by providing weapons and money to rebels as it would if it had launched an outright attack, and yet Article 51 would not be triggered. Such a position– especially in light of the potential threat from rogue regimes– seems to contradict the very essence of a right of effective self-defense established in the Charter. Indeed, as will be noted below, subsequent decisions by the Security Council on terrorism would seem to reject such a restrictive interpretation of armed attack.

---

34. *See id.* The court noted that: "such assistance may be regarded as a threat or use of force, or amount to intervention in the internal or external affairs of other States." *Id.*

35. *Id.*

36. Professor John Norton Moore, for example, has strongly argued against this "high threshold" approach to armed attack. Following the decision of the court, he concluded that the court had "announced, contrary to the Charter, a restrictive interpretation of the right of defense that could deny individual or collective defense against secret warfare, the most serious contemporary threat to world order." John Norton Moore, *The Nicaragua Case and the Deterioration of World* Order, 81 AM. J. INT'L L. 151, 152 (1987).

2. Support for Terrorism

Until the events of September 11th, most of the discussion about how assistance to terrorist groups related to the armed attack threshold came in a variety of scholarly debates. Not surprisingly, scholarly opinion was all over the map. Some scholars supported a "high threshold" approach, arguing that a terrorist attack should only be considered to be tantamount to an armed attack if such attack occurs within the territory of the aggrieved states. Presumably, these scholars would want the victim states to prove direct state sponsorship of the terrorist group. At the other extreme, other scholars argued for a "low threshold." Under such a conception, terrorist actions could rise to the level of an armed attack even if they did not occur within the territory of the victim states. It would also make sense that low threshold supporters would also require less than full sponsorship from a state aiding terrorism to be held responsible for an armed attack. Somewhere in between these two would be a "moderate level" threshold.

Using the logic of the *Nicaragua* case, it would seem that the International Court of Justice was in the high threshold camp. Indeed, this was the conclusion reached by Michael Glennon in his recent analysis of the American response to the September 11thactions.[37] The act in question would need to occur in the territory of the state and the state aiding the terrorist would have to have be engaging in *state-sponsored* terrorism. As Glennon notes, quoting from the *Nicaragua* case, "[a]ctive, not passive, support– an actual 'sending' of 'armed bands, groups, irregulars or mercenaries' or 'substantial involvement therein' -- is necessary to meet the armed attack requirement."[38] "Indeed," as Glennon continues, "the entire approach of the United States in fighting terrorism -- refusing to distinguish between terrorists and those who harbor them, which has come to be called the 'Bush Doctrine' -- is outlawed by this precept to the extend that it precludes any use of force against states that only passively provide a safe harbor for terrorists and avoid substantial involvement in terrorists' activities."[39]

Yet, despite this approach by the International Court of Justice, in recent months it appears that the Security Council articulated a bit lower threshold. In Resolution 1373, the Council was very strong in its condemnation of providing any kind of assistance to terrorist groups.[40] The Council also reaf-

---

37. *See generally* Michael J. Glennon, *The Fog of Law: Self-Defense, Inherence and Incoherence in Article 51 of the United Nations Charter*, 25 HARV. J.L. & PUB. POL'Y 539 (2002).

38. *Id.* at 542 (citation omitted).

39. *Id.* at 543-44.

40. *See* U.N. S.C. Res. 1373, U.N. SCOR, 56th Sess., 1385th mtg., U.N. Doc. S/RES/1373 (2001). In this Resolution, the Council:

firmed the "inherent right of individual or collective self-defence as recognized by the Charter of the United Nations as reiterated in resolution 1368 . . . ."[41] As noted earlier, in Resolution 1390, the Council went on to condemn the Taliban regime for its support for Al Qaeda. In light of these actions by the Council, and in the wake of American actions against the harboring of terrorists, its seems that the Council has been willing to go further than the International Court of Justice in expanding the definition of an armed attack. But neither the Security Council nor any other authoritative decision-maker has articulated a clear definition of the armed attack threshold.

Decide[d] also that all States shall:
(a) Refrain from providing any form of support, active or passive, to entities
or persons involved in terrorist acts, including by suppressing recruitment of
members of terrorist groups and eliminating the supply of weapons to terrorists;
(b) Take the necessary steps to prevent the commission of terrorist acts, including by provision of early warning to other States by exchange of information;
(c) Deny safe haven to those who finance, plan, support, or commit terrorist
acts, or provide safe havens;
(d) Prevent those who finance, plan, facilitate or commit terrorist acts from
using their respective territories for those purposes against other States or their
citizens;
(e) Ensure that any person who participates in the financing, planning, preparation or perpetration of terrorist acts or in supporting terrorist acts is brought
to justice and ensure that, in addition to any other measures against them, such
terrorist acts are established as serious criminal offences in domestic laws and
regulations and that the punishment duly reflects the seriousness of such terrorist
acts;
(f) Afford one another the greatest measure of assistance in connection with
criminal investigations or criminal proceedings relating to the financing or support
of terrorist acts, including assistance in obtaining evidence in their possession
necessary for the proceedings;
(g) Prevent the movement of terrorists or terrorist groups by effective border
controls and controls on issuance of identity papers and travel documents, and
through measures for preventing counterfeiting, forgery or fraudulent use of identity
papers and travel documents.

Id.

41. *Id.* Resolution 1368 was the first that condemned the September 11th attacks. *See* U.N.S.C. Res. 1368, U.N. SCOR 56th Sess., 4370th mtg., U.N. Doc. S/RES/1368 (2001).

B. The Internal Use of Force

For many years, there have been extensive debates about the lawfulness of the use of force to prevent a state from committing genocide or other massive human rights violations. Using the rubric of "humanitarian intervention" scholars and public officials alike have deliberated upon the possible legal justification for such actions. Most commentators can be understood as falling into one of two categories -- restrictionists and counter-restrictionists.[42] The restrictionists argue that the language of the Charter is clear on this point. Unless the Security Council authorizes the use of force, states cannot intervene in another state to stop a government from engaging in human rights violations. Counter-restrictionists, however, present three primary arguments, or variants thereof, to claim that in cases where a state engages in genocide or other violations of fundamental human rights, intervention should be lawful -- even in the absence of an explicit Security Council authorization.

The first of these arguments relates to the protection of human rights. While it is true that the Charter sought to establish a restrictive regime relating to the recourse to force, it also endeavored to create a legal regime to protect human rights. As Professor Fernando Tesón has noted, "the promotion of human rights is as important a purpose in the Charter as is the control of international conflict."[43] As a consequence, it could be argued that in certain extreme cases, the protection of human rights should trump the Charter's prohibition on the use of force. Professors Reisman and McDougal explain that the Charter's "repeated emphasis upon the common interest in human rights indicates that the use of force for the urgent protection of such rights is no less authorized than other forms of self-help."[44]

A second argument presented by some counter-restrictionists contends that there has been a revival of a customary right of humanitarian intervention. Here, counter-restrictionists argue that because the United Nations Security Council has been, for the most part, unable to act effectively to prevent human rights abuses, a pre-existing right of self-help has been revived.

A final argument in support of the lawfulness of humanitarian interven-

---

42. We have used these categories previously. *See* AREND & BECK, *supra* note 4, at 131-36. The discussion of humanitarian intervention here draws upon our previous discussion.

43. FERNANDO TESÓN, HUMANITARIAN INTERVENTION : AN INQUIRY INTO LAW & MORALITY 131 (1988).

44. Michael Reisman & Myres S. McDougal, *Humanitarian Intervention to Protect the Ibos*, *in* HUMANITARIAN INTERVENTION AND THE UNITED NATIONS 172 (Richard B. Lillich, ed. 1973).

tion claims that using force solely for humanitarian purposes falls below the "Article 2(4) threshold." This argument is based on a close reading of the language of Article 2(4). The text of that provision does not literally prohibit all threats or uses of force, but only those directed against the "territorial or political independence" of states or which in some other way transgresses the "Purposes of the United Nations." Hence, it is contended, a use of force to protect human rights is lawful, since "[a] genuine humanitarian intervention does not result in territorial conquest or political subjugation."[45] Indeed, it could be argued that the true political independence of a state lies in the people of the state, not in the ruling elite. Using force to protect these suffering people, thus, would actually be affirming the political independence of the state.

Despite these scholarly arguments, one of the real problems with the concept of humanitarian intervention has been the difficulty finding examples where states made unambiguous claims of engaging in humanitarian intervention, and the international community responded to those claims. Only with the NATO action in the Kosovo crisis has there been a relatively clear case of humanitarian intervention. Even in that case, the United States did not explicitly justify its actions based on that doctrine.[46] Nonetheless, Kosovo may be a turning point. Given the many positive reactions to NATO's actions, seeming support from the Secretary-General, and muted criticism from the rest of the world, Kosovo may be the crucial stage in the emergence of a clear doctrine of humanitarian intervention. In a sense, it may still be too soon to tell, but the reaction of the international community to a subsequent use of force to promote human rights may well indicate that states are now prepared to accept humanitarian intervention.[47] The law, however, remains unclear at present.

---

45. TESÓN, *supra* note 42, at 131.

46. The United States was rather inconsistent in the justifications it gave for the NATO action. *See Contemporary Practice of the United States Relating to International Law,* 93 AM. J. INT'L L. 628, 631 (Sean D. Murphy, ed. 1999) "The U.S. Government based its position on the lawfulness of the air strikes under international law on various factors, which changed in emphasis over time." *Id.* While the United States did, at times, make reference to humanitarian reasons, it also noted other violations of international law by Yugoslavia and previous Security Council resolutions. *See id.*

47. For a lengthier exposition of the current state and possible future of humanitarian intervention, see Christopher C. Joyner & Anthony Clark Arend, *Anticipatory Humanitarian Intervention: An Emerging Legal Norm?*, 10 U.S. A.F. ACAD. J. L. STUD. 27 (1999/2000).

C. Weapons of Mass Destruction

As noted above, weapons of mass destruction in the hands of a rogue state pose a unique threat to the international system. But would such possession or efforts to gain possession of such weapons give rise to a right of forcible self-help on the part of an outside state?

Under customary international law that pre-dated the Charter, many scholars would argue[48] there was a right of anticipatory self-defense. Based on the famous *Caroline* test of 1837, a state could use force in self-defense prior to an actual armed attack if two criteria were met: necessity and proportionality. First, the state seeking to use force would need to demonstrate that some attack was imminent. Second, the response by the soon-to-be victim state would have to be proportionate to the threat. Such guidelines are fine when applied to conventional warfare in the 1800's. They become more problematic when applied to modern warfare -- in particular, in the face of the threat of weapons of mass destruction. The necessity prong is particularly troublesome. In the original formulation given by the American Secretary of State in the *Caroline* case, a state using force in anticipatory self-defense would have to demonstrate that "the necessity of that self-defense is instant, overwhelming, and leaving no choice of means, and no moment of deliberation."[49] Yet with nuclear weapons or other weapons of mass destruction, by the time this test is met, it is probably too late. As noted earlier, the Security Council criticized Israel following the bombing of the Iraqi reactor for failure to meet the test of necessity. The reactor was not yet on-line, it was unclear that when it did come on-line that it would produce nuclear weapons -- since Iraq was a party to the Nuclear Non-Proliferation Treaty -- and the mere possession of weapons would not constitute an imminent threat of their use. While hindsight may judge Israel's action differently, the legal questions posed at the time remain the same today. Certainly, the strict test of necessity that Daniel Webster described does not make a great deal of sense in the age of weapons of mass destruction. But what is the appropriate test? The international community does not seem to have developed a clear alternative.

---

48. For an examination of the scholarly debate on anticipatory self-defense, see AREND & BECK, *supra* note 4, at 71-79. Scholars supporting the existence of such right include William V. O'Brien, Derek W. Bowett, Myres S. McDougal, and Julius Stone.

49. J.B. Moore, *Destruction of theCaroline*, *in* 2 DIGEST OF INTERNATIONAL LAW 412 (1906) (quoting M. Webster, Sec. of State, to Lord Ashburton, British plen., Aug. 6, 1842).

IV. CONCLUSIONS AND RECOMMENDATIONS

In light of the preceding discussion of the challenges posed by rogue states, several conclusions become clear. First, the United Nations Charter framework for the use of force did not anticipate the ways in which rogue states would use force. As a consequence, it is unclear how the literal language of the Charter can be applied to various actions of rogue states. Second, while the international community has attempted to respond to some of these challenges through international organs– such as the International Court of Justice and the Security Council -- and the efforts of international legal scholars, international law is still very much in dispute. When does a covert action or efforts to support rebels rise to the level of an armed attack? What level of involvement with terrorist groups is necessary for a rogue state to have itself committed an armed attack? Can force ever be used to prevent a rogue state from engaging in genocide or other human rights violations? Can weapons of mass destruction in the hands of a rogue state give others the right to preemptively use force? The answers to all of these questions are by no means clear. So what does this mean for the Charter framework?

Most scholars have sought to salvage the United Nations Charter framework in the wake of these and other challenges. They have attempted to preserve the core of the Charter provisions on the use of force, modifying them based on their understanding of customary international law.[50] Such approach, however, seems to miss the fundamental point. While it is indeed true that the actions posed by rogue states challenge the efficacy of the Charter regime, so too do the actions of non-rogue states. The true test of whether a putative rule of international law is actually law is whether that rule has both authority and control.[51] First, is the rule authoritative? Is it perceived by the decision-making elites in states to be the law -- does it have *opinio juris*? Second, is the rule controlling of the behavior of international actors? Is it reflected in the *practice of states*?[52] If the rule in question is the Charter framework for the use of force -- *no threat or use of force unless undertaken in self-defense following an armed attack or as authorized by the Security Council* -- my conclusion is that it fails the test.

---

50. Elsewhere, we have called these individuals "core-interpretists." They argue that the "core" of Article 2(4) is still good law, but contend that a reading of 2(4) as prohibiting *all* threats or uses of force is not an accurate understanding of that provision. *See* AREND & BECK *supra* note 4, at 182-84.

51. For a more elaborate examination of the "authority-control" test, see ANTHONY CLARK AREND, LEGAL RULES AND INTERNATIONAL SOCIETY (1999).

52. As Professor Mark Weisburd has noted: "A putative rule cannot survive contrary practice no matter how that practice is characterized." A. Mark Weisburd, USE OF FORCE: THE PRACTICE OF STATES SINCE WORLD WAR II 24 (1997).

First, the rule of the Charter is not authoritative. While states may give lip service to the Charter framework, since 1945 states from all quarters of the international system have consistently claimed the right to use force in circumstances that fall outside of any reasonable interpretation of the Charter.[53] Second, the rule is not controlling of state behavior. Despite protests to the contrary, states have actually used force hundreds of times in ways that are inconsistent with the Charter rules.[54] As Professor Glennon aptly concludes, "[s]o many states have used force with such regularity in so wide a variety of situations that it can no longer be said that any customary norm of state practice constrains the use of force."[55]

If this is indeed the case, if the Charter framework for the use of force does not reflect existing international law, then what does this mean for rogue states? Does it mean that there is nothing that can be said about the lawfulness of their actions? To me, it says that now is the time to begin creating new rules about the behavior of rogue states. Democratic states must unite to develop a set of rules relating to the use of force that will truly criminalize the behavior of rogue regimes. But in doing so, those states must understand that a rule-based regime must apply to them as well. As Glennon notes, "[n]o rules will work that do not reflect underlying geopolitical realities."[56] "The use-of-force regime set out in the U.N. Charter failed," he explains, " because the Charter sought to impose rules that are

---

53. *See* AREND & BECK, *supra* note 4, at 181. (" For obvious political reasons, states have not overtly argued that the Charter norms are invalid. But as noted earlier, states have on numerous occasions claimed the right to use force in circumstances that are, nevertheless, clearly antithetical to the principle enshrined in Article 2(4). Given these claims, it seems incorrect to contend that states still hold 2(4) in very high esteem. Admittedly, the provision may command some perceptions of legitimacy, but they seem to be far below those required for a healthy rule of law.") (footnotes omitted).

54. In his brilliant work, *Limits of Law, Prerogatives of Power*, Michael Glennon provides the most recently exploration with respect to state practice. He examines numerous uses of force since the Charter was adopted and then asks this question -- "What lesson is to be drawn from this sorry record?" MICHAEL J. GLENNON, LIMITS OF LAW, PREROGATIVES OF POWER: INTERVENTIONISM AFTER KOSOVO 84 (2001). He answers the question by concluding:

> It is impossible to avoid the conclusion that use of force among states simply is no longer subject – if it ever was subject – to the rule of law. The rules of the Charter do not today constitute binding restraints on intervention by states. Their words cannot realistically be given effect in the face of widespread and numerous contrary deeds.

Id.

55. Glennon, *supra* note 38, at 554. This is consistent with the conclusion that Professor Beck and I reached in 1993. Although at the time, we argued that there did still seem to be an authoritative and controlling prohibition on the use of force for pure territorial aggrandizement. AREND & BECK, *supra* note 4, at 187-88.

56. Glennon, *supra* note 38, at 557.

out-of-sync with the way states actually behave."[57] Consequently, "[a] new use-of-force regime that does work will have to rest far more firmly upon actual patterns of practice that reveal, with solid empirical evidence, what regulation of force is possible and what is not."[58]

The starting point for this new regime, its "core," could be "rogue behavior." Rather than attempting to create a full-blown legal system to restrict the use of force, if the United States and its allies that seek to establish a rule-based regime would begin by prohibiting "rogue" actions, they may be able to develop rules that would truly be authoritative and controlling. What would such a regime look like? That remains to be seen. But whatever regime the United States and other democratic states seek, it must be a regime that they are willing to follow. Otherwise, such putative legal regime will be just as illusory as the Charter regime has become.[59]

---

57. *Id.*

58. *Id.*

59. Commenting as early as 1970, Professor Thomas M. Franck observed that "[t]he practice of these states has so severely shattered the mutual confidence which would have been the *sine qua non* of an operative rule of law embodying the precepts of Article 2(4) that, as with Ozymandias, only the words remain." Thomas M. Franck, *Who Killed Article 2(4)? Or: Changing Norms Governing the Use of Force by States*, 64 AM. J. INT'L L. 809, 809 (1970).

Lessons from the 1986 Libya Airstrike

W. Hays Parks*

Before turning to the principal portion of my remarks, I believe it is necessary to preface them with a few points.

First, my remarks are offered as a participant in the events I will describe.[1]

Second, it should be noted that my official curriculum vitae notes that I was a legal adviser in 1986, not the legal adviser. Working at a higher level was Judge Abraham D. Sofaer, a panelist this afternoon, who was the Legal Adviser at the Department of State at the time, and Michael J. Matheson, his Principal Deputy Legal Adviser. Due to security compartmentalization, we had no contact with one another with respect to the primary events I am about to describe during the course of their planning.

Third, consideration of this mission must be taken in the context of the time in which it was conducted. I will elaborate on this later in my presentation, but it is made to emphasize that one must measure a mission's degree of success based upon what it was intended to accomplish, and what it accomplished at the time.

Finally, the law does not operate in a vacuum. It operates within the political and military structure of our national security system. Hence my presentation will provide a summary of events leading up to the attack, a look at the legal issues with which I was concerned, a summary of the

---

* Special Assistant for Law of War Matters to The Judge Advocate General of the Army. Remarks offered to the Regional Meeting of the American Society of International Law, "Responding to Rogue Regimes: From Smart Bombs to Smart Sanctions," New England School of Law, Boston, MA, November 9, 2001. The comments are the personal views of the author and do not necessarily represent official positions of the Department of the Army or any other agency of the United States Government.

1. The primary source for these remarks is the author's personal participation in those events, previously recorded for the most part in W. Hays Parks, *Crossing the Line*, U.S. NAV. INST. PROC. (Nov. 1986), at 40-52. Secondary sources include DAVID C. MARTIN & JOHN WALCOTT, BEST LAID PLANS (1988); BRIAN L. DAVIS, QADDAFI, TERRORISM AND THE ORIGINS OF THE U.S. ATTACK ON LIBYA (1990); and ROBERT E. VENKUS, RAID ON QADDAFI (1992).

attack, a few words on the immediate aftermath, and my view of the attack's effects.

I. EVENTS LEADING UP TO THE ATTACK

Terrorism has been with us for some time. Modern terrorism as a direct threat to the United States and its citizens began in the 1960s. It is a long story told elsewhere which, due to time constraints, cannot be detailed here.[2]

A leading and outspoken proponent of the use of terrorism was Libyan dictator Moammar Ghadafi. Ghadafi gained power on September 1, 1969, dramatically reversing the good relations that had been enjoyed between Libya and the United States. From 1972 until President Ronald Reagan formerly severed diplomatic relations in 1981, the United States declined to enjoy an ambassadorial relationship with Libya. In the years between his assumption of power and the airstrike, Ghadafi provided economic aid, training, and logistical support for dissident factions, terrorist groups, or insurgent movements in at least twenty-six nations, including the United States. He joined with the Palestine Liberation Organization (PLO) to provide praetorian guards for the murderous regimes of Emperor Bokassa in Central Africa and Uganda's Idi Amin. He also provided Amin refuge following his ouster. Ghadafi established strong ties with terrorist groups such as George Habash's Popular Front for the Liberation of Palestine (PFLP) and began a long, close relationship with Ramirez Sanchez, the Venezuelan terrorist known as "Carlos."

Ghadafi also provided the funds, arms, and training for the terrorists who committed the 1972 Olympic Games massacre in Munich.[3] The terrorists departed for Munich from Libya, and were awarded bonuses by Ghadafi upon their return. Connections were found between Libya and the Black September attack on the Saudi Arabian Embassy in Khartoum on March 2, 1973, in which U.S. Ambassador Cleo A. Noel, Jr., and Charge

---

2. As terrorism relates to the 1986 airstrike, see generally Edward Weisband & Damir Roguly, *Palestinian Terrorism: Violence, Verbal Strategy, and Legitimacy*, *in* YONAH ALEXANDER, INTERNATIONAL TERRORISM: NATIONAL, REGIONAL, AND GLOBAL PERSPECTIVES 258-310 (1976); EDGAR O'BALLANCE, LANGUAGE OF VIOLENCE: THE BLOOD POLITICS OF TERRORISM (1979); CLAIRE STERLING, THE TERROR NETWORK, THE SECRET WAR OF INTERNATIONAL TERRORISM (1981); AMIR TAHERI, HOLY TERROR: INSIDE THE WORLD OF ISLAMIC TERRORISM (1987); and NEIL C. LIVINGSTONE & DAVID HALEVY, INSIDE THE PLO (1990).

3. *See* DAVIS, *supra* note 1, at 11, 13. The lengthy record of Libyan support for terrorism is provided in DAVIS, at 33-91, and in chronological order at 181-90. A recent account of the 1972 Munich Olympics massacre is discussed in SIMON REEVE, ONE DAY IN SEPTEMBER (2000).

d'Affairs George C. Moore were murdered.[4] Ghadafi established terrorist training camps in Libya and attracted terrorists from the PLO and PFLP, as well as every other major terrorist group.[5] His investments paid the dividends he sought. For example, on December 17, 1973, Palestinian terrorists attacked a Pan Am flight in Rome's Fiumicino Airport, killing thirty-one passengers. "Italian authorities determined that the terrorists had obtained their weapons and money from Libya, and traveled from [Libya] to carry out the attack."[6] In July 1984, eighteen merchant vessels struck sea mines in the Suez Canal and Red Sea, apparently laid by the Libyan cargo ferry Ghat.[7] On October 7, 1985, PLO terrorist Abul Abbas hijacked the Italian cruise liner Achille Lauro. A 69-year-old, wheelchair bound American, Leon Klinghofer, was murdered in the process.[8] The operation was attributed to Ghadafi.[9] On November 23, 1985, Abu Nidal terrorists hijacked an EgyptAir airliner. When the aircraft landed in Malta, the terrorists shot Israeli and American passengers.[10] Fifty-seven other passengers died in a subsequent failed hostage rescue operation.[11] Three of the four terrorists had embarked on their mission aboard a Libyan airliner, and the weapons they used to kill the passengers were traced to Libya.[12] Ghadafi's open reliance on and support for terrorism continued unabated through the 1970s and up to the time of the 1986 U.S. airstrike.

Parallel to Libya's overt support of terrorism, another event occurred relevant to U.S. action. On October 11, 1973, in response to U.S. re-supply of Israel's Yom Kippur War, Ghadafi declared that portion of the Mediterranean Sea south of 32° 30′, known as the Gulf of Sidra (or Sirte), to be Libya's internal waters. In a demarche′ dated February 11, 1974, the U.S. Department of State characterized the Libyan claim as "unacceptable

---

4. *See* DAVIS, *supra* note 1, at 13, 35, *see also* DAVID KORN, ASSASSINATION IN KHARTOUM 48-49 (1993).

5. Including the Japanese Rengo Sekigun (Red Army), Irish Republican Army (IRA), Basque Euzkadi Ta Askatasuna (ETA), West Germany's Red Army Faction, Italy's Red Brigades, Nicaragua's Sandanistas, Argentina's Montoneros, Uruguay's Tupomoro's, Peru's Shining Path, and the Ecuadorian Alfaro Vive, Carajo. *See* Parks, *supra* note 1, at 41.

6. Parks, *supra* note 1, at 41.

7. *See* Scott C. Truver, *Mines of August: An International Whodunit*, U.S. NAV. INST. PROC., at 42-48 (Aug. 1986).

8. *See* YOSSI MELMAN, THE MASTER TERRORIST: THE TRUE STORY OF ABU NIDAL 173-74 (1986).

9. *See* DAVIS, *supra* note 1, at 76; Parks, *supra* note 1, at 44 *and* MELMAN, *supra* note 8, at 182-89.

10. *See* MARTIN & WALCOTT, *supra* note 2, at 267.

11. *See id.*

12. *See* MARTIN & WALCOTT, *supra* note 1, at 266-67; DAVIS, *supra* note 1, at 77.

as a violation of international law."[13] In the course of the United Nations Conference on the Law of the Sea, the United States maintained a policy and program of challenging excessive sea claims. A naval challenge to Libya's illegal claim was considered during the Carter Administration, but was shelved due to concerns for the safety of U.S. citizens working in Libya, and a fear that it would result in the loss of support of moderate Arab nations, such as Tunisia, who were acting as intermediaries with Iran for release of the U.S. hostages (held mainly by Libyan-trained militants) held in Tehran.[14] President Carter's reluctance to challenge the illegal Libyan claim was met by an unsuccessful attack on September 16, 1980, by two Libyan MiG-23 Floggers on a U.S. Air Force RC-135 operating in international airspace, well outside the area Libya claimed.[15]

Challenge to the illegal Libyan claim was an early priority of the Reagan Administration. Such a challenge was intended first as a warning to Ghadafi, regarding his continued emphasis on training terrorists and exporting terrorism. It also was the start of a vigorous program to challenge the illegal claims to the high seas around the world that continue to this day. On the second day of the U.S. freedom-of-navigation (FON) challenge to the illegal Libyan claim, which began on August 18, 1981, one of two Libyan Su-22 Fitters fired an air-to-air missile at two U.S. Navy F-14 Tomcat aircraft. The F-14s maneuvered to avoid the missile before downing the two Libyan fighters with Sidewinder missiles.[16] The FON exercise concluded the following day without further incident. During the next four years, U.S. military forces would enter the challenged area four more times.[17]

Libya's actions in support of terrorism brought matters to a head when,

---

13. The note continued:

> The United States Government views the Libyan action as an attempt to appropriate a large area of the high seas by unilateral action, thereby encroaching upon the long-established principle of freedom of the seas . . .. The [U.S.] Government reserves its rights and the rights of its nationals in the area of the Gulf of Sirte affected by the action of . . . Libya.

DIGEST OF UNITED STATES PRACTICE IN INTERNATIONAL LAW at 293-94 (Arthur W. Rovine, ed., 1974).

14. *See* Parks, *supra* note 1, at 42; *See also* DAVIS, *supra* note 1, at 37. For discussion of the Iranian hostage crisis of 1979-1980, see Warren CHRISTOPHER, ET AL., AMERICAN HOSTAGES IN IRAN (1985).

15. *See* Parks, *supra* note 1, at 43.

16. *See* Dennis. Neutze, *The Gulf of Sidra: A Legal Perspective*, U.S. NAV. INST. PROC. at 26-31, (Jan. 1982).

17. *See* Parks, *supra* note 1, at 44. A separate incident would conclude the Reagan Administration, when on January 21, 1989, Navy F-14 Tomcats shot down two Libyan MiG-23s manifesting hostile intent during their attempt to penetrate a U.S. carrier battle group's defensive perimeter.

on December 27, 1985, Abul Nidal terrorists attacked El Al Airlines ticketing counters in the Rome and Vienna airports, killing twenty civilians, including five Americans. Ghadafi publicly described the attacks as "heroic," and provided a safe haven for the raid's mastermind. Moroccan passports found on the gunmen had been taken from Moroccan laborers in Libya by Libyan officials.[18] The line had been crossed, at least with respect to the patience of the United States.

II. THE PATH TO THE AIRSTRIKE

While something needed to be done to quell terrorism, exactly what, remained to be determined. One endeavor begun during the Reagan Administration was to exclude terrorist acts from the "political offense" exception in extradition treaties.[19] Another step taken by the Reagan Administration was legislation establishing extra-territorial jurisdiction in United States Federal District Courts over any individual who seizes or detains, threatens to kill, assaults or murders a U.S. citizen anywhere in the world.[20] Of course, it was unlikely Ghadafi would turn over anyone to U.S. authorities for terrorist acts he supported and praised. Extraterritorial military force had been used in the past, generally, however, only to rescue hostages during an on-going incident.[21] Actions by Israel outside its

---

18. *See* MELMAN, *supra* note 8, at 182; *See also* DAVIS, *supra* note 1, at 78-79; MARTIN & WALCOTT, *supra* note 1, at 267-68; Abraham D. Sofaer, *Terrorism, The Law, and the National Defense*, 126 MIL. L. REV. 91, 103 (1989).

19. *See* Abraham D. Sofaer, *Terrorism and the Law*, FOREIGN AFFAIRS, 901, 906.(Summer, 1986) [hereinafter Sofaer, *Terrorism and the Law*]. 20. *See* Act for the Prevention and Punishment of the Crime of Hostage Taking, 18 U.S.C. § 1203 (1994). The law bore fruit on September 13, 1987, when FBI agents arrested Fawaz Yunis in international waters off Beirut. Yunis and four others hijacked Royal Jordanian Airlines flight 402 in Beirut, Lebanon, on June 11, 1985. Two of its passengers were U.S. citizens. The plane made two unsuccessful circuits of the Mediterranean before returning to Beirut, where the hijackers destroyed it. *See* U.S. v. Fawaz Yunis, 924 F.2d 1086 (D.C. Cir., 1991). Details of Yunis's capture are provided in OLIVER REVELL & DWIGHT WILLIAMS, A G-MAN'S JOURNAL 262-63, 265-72 (1998).

21. Classic examples are the July 4, 1976, Israeli rescue of the passengers and crew of Air France flight 139 in Entebbe, Uganda, and a mission by the West German Federal Border Guard Group 9 Special (GSG9) on October 18, 1977, in Mogadishu, Somalia, when the unit rescued the passengers and crew of Lufthansa flight 181 that had been hijacked by four Palestinian terrorists. *See generally* YESHAYAHU BEN-PORAT, ET AL., ENTEBBE RESCUE (1976); *see also* ROLF TOPHOVEN, GSG9: GERMAN RESPONSE TO TERRORISM 66-73 (2d ed., 1985); LEROY THOMPSON, THE RESCUERS 23-28, 30-33 (1986); BARRY DAVIES, FIRE MAGIC (1994); and Sofaer, *supra* note 18, at 107.

borders in response to terrorism were generally condemned – in all probability for political as much as legal reasons – and had not been supported by the United States.

Terrorist attacks against the United States prompted re-thinking in the early 1980s of the legal bases for use of force in self-defense. The right of self-defense, articulated in Article 51 of the Charter of the United Nations, was widely interpreted as condoning self-defense against an existing armed attack by conventional forces. It is safe to say that at the time the Charter was promulgated, its drafters did not contemplate that one state would resort to regular and continuous acts of terrorism, often covert, against the citizens of another state or other states. In the briefest of terms, the analysis developed at the national level reasoned that self-defense could be used against a continuing threat, such as that posed by Libya's actions. I suspect that I have not done fairness to this argument, but must refer you to the explanation by Judge Abraham Sofaer, State Department Legal Adviser at the time, due to time limitations.[22] We can explain what the airstrike was and what it was not intended to be. It was neither a peacetime reprisal nor an act of retaliation. It was a limited act of self-defense against a continuing threat for the purpose of deterring Libya's continued reliance upon illegal acts of terrorism as a foreign policy tool.

The perfect legal answer seldom resolves a problem, and this certainly was the case within the Reagan Administration with respect to gaining the support of other governments for a response to terrorism in general and Libyan terrorism in particular. In the interim, in the opening days of 1986, the national leadership directed the military to begin to develop military options ranging from a naval presence mission off Libya to an airstrike. The mission statement was straightforward: If directed, "inflict damage on target complexes that support terrorist operations." Simultaneously the President signed Executive Order 12543, terminating commerce between Libya and the United States, and prohibiting work by U.S. citizens in Libya.[23] Its intent was to encourage U.S. citizens working in Libya to depart, lest Libya take them hostage.

---

22. *See* Sofaer, *supra* note 18, at 92-98; *see also* Sofaer, *Terrorism and the Law*, *supra* note 19, at 911; DAVIS, *supra* note 1, at 84. Secretary of State George Shultz commented:

> The law is a weapon on our side and it is up to us to use it to its maximum extent. . .. [A] state which supports terrorist or subversive attacks against another state, or which supports or encourages terrorist planning and other activities within its own territory, is responsible for such attacks. Such conduct can amount to an ongoing armed aggression against the other state under international law.

Sofaer, *Terrorism and the Law*, *supra* note 19, at 910.

23. *See* Prohibiting Trade and Certain Transactions Involving Libya, 51 Fed. Reg. 875 (Jan. 7, 1986).

U.S. naval presence off the coast of Libya grew during the first three months of 1986 from a single carrier battle group to three carrier battle groups, culminating in a successful freedom-of-navigation exercise from March 23rd to 27th below the so-called "line of death" which Ghadafi had drawn across the Gulf of Sidra.[24] Diplomatic initiatives were not as successful. Efforts to enlist support for sanctions against Libya were rebuffed by European governments. Most gained substantial revenue and/or oil from trading with Libya, had citizens working in Libya, and/or feared retaliatory terrorist attacks by Libya.[25]

III. LEGAL ISSUES AT THE AIRSTRIKE PLANNING LEVEL

The planning for a possible airstrike began in the days immediately following the terrorist attacks on the Rome and Vienna airports, with the view that the more time permitted for planning and coordination, the better its execution. It remained an option rather than a foregone conclusion.

While lawyers at the national leadership level were focused on the jus ad bellum, that is, the law relating to when force may be used, the airstrike planning involved the jus in bello, commonly known as the law of war. At the outset I laid out some basic principles:

> Only legitimate targets may be attacked. But a legitimate target may be attacked at any time, wherever it is located.
>
> No nation possesses the means or capability of accurately forecasting likely civilian collateral casualties at a given target, under all circumstances.
>
> The inevitability of collateral civilian casualties and collateral damage to civilian objects is recognized as regrettable but not necessarily illegal.
>
> The obligation to reduce collateral civilian casualties and collateral damage to civilian objects is the responsibility of all parties, including the attacker, defender, and the civilian population. Civilians who choose to remain near a target do so at an assumed risk.
>
> The obligation of the attacker to reduce collateral civilian casualties and collateral damage to civilian objects is limited to those precautions that are feasible and consistent with mission accomplishment and security of the attacking force.
>
> A defender cannot use individual civilians or the civilian population as human shields to prevent the attack of a lawful target. At the same time, the inevitable intermingling of civilian objects and military objectives is not illegal.
>
> Not all collateral civilian casualties or collateral damage to civilian objects can be attributed to either the attacker or the defender. Some are the result of

---

24. *See* Parks, *supra* note 1, at 40-41, 44.
25. *See* DAVIS, *supra* note 1, at 85-88.

the friction of war, beyond the control of either the attacker or defender.

Target selection did not involve the determination of whether a target was "legal" or "not legal," but rather, in accordance with the mission statement, whether targets were terrorist-related. Targets also were considered on the basis of force risk and/or risk to innocent civilians. Many traditional military objectives or terrorist-related targets were not selected for one or both of these reasons. One terrorist-related facility – Al Arafiq, the Libyan intelligence service headquarters - was not selected for attack due to its proximity to densely populated areas, including diplomatic residences. Ultimately two targets were selected that were in populated areas. The following target nominations by the combatant commander were endorsed by the Joint Chiefs of Staff and Secretary of Defense, and approved by President Reagan:

> Murrat Sidi Bilal Terrorist Training Camp: a combat swimmer and naval commando school in the Tripoli area where PLO and other terrorist organization frogmen were trained.
>
> Tripoli Military Airfield: Il-76 Candid transports used to support Ghadafi's export of terrorism were the primary targets.
>
> Benghazi Military Barracks/Jamahiriyah Guard Barracks: an alternate terrorist command and control headquarters. Like Aziziyah Barracks, it was a billeting area for Ghadafi's elite Jamahiriyah Guard. It also contained a warehouse for storage of MiG aircraft components, housing for members of the Islamic Legion, and served as a terrorist training facility.
>
> Benina Military Airfield: Although not directly tied to Ghadafi's export of terrorism, Benina Military Airfield was selected for attack because it contained MiG-23 fighter aircraft manned by Libyan pilots capable of night operations against U.S. strike aircraft.[26]
>
> Tarabulus (Aziziyah) Barracks: Located in Tripoli, it was the principal command and control center for Ghadafi's worldwide terrorist program. It was also Ghadafi's residence on occasion. However, because he feared an internal coup, he stayed there infrequently. We did not know where he would be if an airstrike was ordered. He was not designated as a target, even though an attack on his person would have been lawful.

Means for mission execution were considered concurrent with selection of targets. A variety were available, including a special operations or Marine raid; use of AC-130 gunships, capable of precision fire; and a tactical airstrike. Naval gunfire from a battleship might have been considered, but none were available. Moreover, accuracy limitations would have limited its use against targets in heavily populated areas. Cruise missiles were not considered because of a lack of confidence in their capabilities at that time. Also, a decision was made not to employ the newly-operational F-

---

26. *See* Parks, *supra* note 1, at 48; DAVIS, *supra* note 1, at 136-38.

117 stealth fighter and reveal its existence. In considering the level of damage sought, reducing the risk to civilians, and risk to the attacking force, the decision to use tactical aircraft was made in about as much time as it has taken to summarize the force options available and factors considered. This decision process highlights another law of war principal: a nation is not legally obligated to use its most precise means to attack a target. A balancing process exists, and a reasonable commander is entitled to make his decisions based upon the mission, his assessment of force risk, and other information reasonably available to him at the time.

In addition to the carrier-based A-6E Intruder, the theater commander sought use of U.S. Air Force F-111F and EF-111 aircraft in order to achieve maximum effectiveness while limiting risk to the attack force. Accordingly, while Navy and Air Force units began joint planning, legal issues arose with regard to basing from, and overflight of, foreign nations. Quiet, diplomatic approaches to several foreign governments were not successful,[27] causing U.S. Air Force planners to resort to another international law principle – freedom of the high seas, and the air space above it – in establishing the air route to and from their targets for the F-111s and EF-111s that would launch from bases in Great Britain. Their use was conditional, however, based on British consent. In the first three months of 1986, such consent was not a foregone conclusion. On January 10, 1986, British Prime Minister Margaret Thatcher rejected economic sanctions against terrorism, arguing that they "don't work," and characterized retaliatory military attacks against terrorism as "against international law" and likely to cause "much greater chaos."[28] In the ensuing months, European governments lobbied Prime Minister Thatcher heavily to deny the United States use of British bases in England for any military action against Libya.[29]

Up to this time – early April, 1986 – a U.S. airstrike against Libya had been only an option. The only steps taken had been those of planning and coordination, and a number of unsuccessful diplomatic efforts to energize European governments to support the United States in discouraging Libya's unabated reliance upon terrorism.

Ghadafi was to force the issue when, eight days after the U.S. freedom-of-navigation exercise in the Gulf of Sidra, a bomb exploded aboard TWA flight 840 en route from Rome to Athens, killing four U.S. citizens. The Abu Nidal group claimed responsibility in retaliation to the U.S. naval exercise.[30] Two days later a terrorist bomb, detonated in the La Belle dis-

---

27. *See* DAVIS, *supra* note 1, at 125-26.

28. DAVIS, *supra* note 1, at 84; *see also* MARGARET THATCHER, THE DOWNING STREET YEARS 442 (1993).

29. *See* DAVIS, *supra* note 1, at 125.

30. *See* PARKS, *supra* note 1, at 45; *see also* DAVIS, *supra* note 1, at 110.

cotheque in Berlin, killed two people, including a U.S. serviceman, and seriously injured 230 others, including 80 Americans. One American soldier died from his wounds two months later. Evidence of Libyan responsibility, gathered by U.S., British and German intelligence, was "incontrovertible."[31] While the decision to strike was almost immediate, authorization was not given until a week later, as the U.S. leadership made one final but unsuccessful attempt to persuade Western European governments to go along with peaceful measures to isolate Libya, and to enhance antiterrorism capabilities in order to deter Libya and others from engaging in terrorism. Media speculation about the raid raised the legal issue as to whether Ghadafi's death would violate Executive Order 12333, which prohibits assassination (it would not have). Time was of the essence, as Ghadafi announced that he was taking hostage all foreigners in Libya in order to use them as human shields around key Libyan military bases. While most European leaders continued to express timidity, British Prime Minister Thatcher responded favorably to President Reagan's request to permit U.S. use of British bases by Air Force components.[32]

IV. SUMMARY OF THE ATTACK

The attack against targets in the vicinity of Tripoli was executed by F-111F aircraft, while Navy A-6E Intruder aircraft attacked the two targets in eastern Libya. Simultaneously, carrier-based Navy and Marine Corps aircraft engaged in suppression of enemy air defenses while protecting the strike force against any potential enemy air threat. Each target was damaged. One F-111 was lost in-bound, and not as a result of hostile action. Damage was not as severe as it might have been, particularly against Tarabulus (Aziziyah) Barracks, due to overly restrictive rules of engagement imposed by an intermediate command.[33] Errant bombs from one F-111F struck the French Embassy, one and one half miles away from its intended target of Tarabulus (Aziziyah) Barracks when the aircraft's crew selected the wrong off-set point.[34]

---

31. DAVIS, *supra* note 1, at 116. On November 13, 2001, four persons -- a German woman, a Palestinian man, a Libyan man, and a Lebanese-born German woman -- were convicted of murder or attempted murder by a German court for their involvement in the La Belle discothèque attack. Peter Finn, *4 Convicted in '86 Berlin Nightclub Bombing*, WASH. POST, Nov. 14, 2001, at A30; Geir Moulson, *4 Guilty in Disco Bombing; Ghadafi Link Rejected*, WASH. TIMES, Nov. 14, 2001, at A13.

32. *See* Parks, *supra* note 1, at 50; DAVIS, *supra* note 1, at 125-27; THATCHER, *supra* note 28, at 443-46.

33. *See* Parks, *supra* note 1, at 51; DAVIS, *supra* note 1, at 137-38.

34. *See* DAVIS, *supra* note 1, at 140.

V. IMMEDIATE AFTERMATH

As soon as the last U.S. aircraft was "feet wet," that is, safely over water and headed away from Libya, President Reagan addressed the American people on nationwide television. His statement was a careful articulation of the rationale for the attack, as well as the degree to which the rule of law influenced the mission and its planning.[35] Rather than do the President's address injustice by attempting to summarize it, I recommend it to you.

Following the attack, Libya claimed that thirty-six Libyans had died, some of whom were Libyan soldiers, many of whom may have been killed by Libyan anti-air munitions rather than U.S. bombs. Media tours arranged by the Libyan government to spread disinformation were less than successful, while U.S. officials were able to refute Libyan claims of U.S. losses through quick analysis of Libyan photographs. Of particular note was the claim that an adopted daughter of Moammar Ghadafi had been killed in the airstrike on Ghadafi's "home" – a Bedouin tent – located at Tarabulus (Aziziyah) Barracks. The Gucci-made tent was not Ghadafi's residence, but where Ghadafi held his press conferences; and there was no record that an adopted daughter existed, much less had been killed in the air raid.[36]

Finally, in somewhat of a humorous post script, in April 1987 former U.S. Attorney General Ramsey Clark and Illinois law professor Francis Boyle brought suits against President Ronald Reagan, British Prime Minister Margaret Thatcher, and all other officials involved in the planning or execution of the air strike ostensibly on behalf of the next of kin of those killed during the air strike. The suit was seen as lacking any merit, and was dismissed quickly.[37]

VI. ATTACK'S EFFECTS

The attacks brought both supporters and detractors. In light of the Libyan bombing of Pan American flight 103 over Lockerbie, Scotland, on December 21, 1988, it might be argued that the April 1986 attack did not teach Ghadafi a lesson.[38] Such an argument would miss the point. No one believed that a single limited air strike would defeat terrorism, or that it

---

35. President's Address to the Nation, United States Air Strike Against Libya, 1986 PUB. PAPERS 491 (Apr. 14, 1986).

36. *See* DAVIS, *supra* note 1, at 140-42.

37. *See* Saltany v. Reagan, 702 F. Supp. 319, 322 (D.C. 1988).

38. For discussion of the Lokerbie bombing and its aftermath, see Michael P. Scharf, *Terrorism on Trial: The Lockerbie Criminal Proceedings*, 6 ILSA J. INT'L & COMP. L. 355 (2000).

would serve as anything more than a temporary deterrence.[39] It did have several effects. First, it established a critical precedent that force could and would be used against those who support, plan, or finance terrorism, or who use their sovereign territory to harbor or train terrorists. Second, as British Prime Minister Thatcher acknowledged, it placed terrorism on the international agenda of Western governments.[40] Finally, while alliances in responding to terrorism and other acts by rogue regimes may be nice, they are not critical. U.S. actions on the night of April 15, 1986, indicated to others that their timidity to respond to terrorism would not restrain the United States where action was necessary.

39. *See, e.g.*, THATCHER, *supra* note 28, at 444 (containing President Reagan's statement).

40. *See id.* at 498.

The Iraqi Quagmire: Enforcing the No-Fly Zones

Scott L. Silliman*

In a relatively short period of time following the Iraqi invasion of Kuwait on August 2, 1990, the United Nations Security Council issued a series of resolutions that served as the international legal predicate for the use of force by the coalition. The first, Resolution 660[1], issued on that very day, determined that the invasion constituted a breach of international peace and security; and the Council, acting under Articles 39 and 40 of the Charter, condemned the invasion and demanded that Iraq withdraw immediately and unconditionally. Thereafter, within a span of less than four months, the Security Council issued eleven more resolutions calling for economic embargoes, blockades and finally, the use of armed force in its attempt to force Iraq to withdraw from Kuwait. The last of the twelve, Resolution 678[2], passed on November 29, 1990, authorized member states "to use all necessary means to uphold and implement resolution 660 and all subsequent resolutions and to restore international peace and security in the area" if Iraq did not, by January 15, 1991, comply with the previous eleven resolutions.[3] As we all know, Saddam Hussein did not comply, and Operation Desert Storm commenced in the early morning hours of January 17, 1991, with air strikes lasting for some thirty-nine days, followed by a 100–hour ground campaign that ended on February 28th with a declaration of an end to hostilities.[4] There can be no question of the legality of the use of force by the coalition to oust the Iraqi troops from Kuwait in Operation Desert Storm. It was specifically sanctioned by the Security

---

* A.B. 1965, J.D. 1968, University of North Carolina. After serving for 25 years as a uniformed judge advocate in the United States Air Force, Professor Silliman retired in the grade of colonel and joined the faculty at Duke University School of Law. He is the Executive Director of the Law School's Center on Law, Ethics and National Security.

1. S.C. Res. 660, U.N. SCOR, 45th Sess., 2932nd mtg., U.N. Doc. S/RES/660 (1990).

2. S.C. Res. 678, U.N. SCOR, 45th Sess., 2963rd mtg., U.N. Doc. S/RES/678 (1990).

3. *See id.*

4. *See generally Triumph Without Victory: The History of the Persian Gulf War*, U.S. NEWS & WORLD REPORT (1992).

Council acting under Chapter VII and invoking Article 42 of the Charter.[5] With respect to the clarity and certainty of the international legal predicate for this operation, I would suggest that it differs markedly from NATO's Operation Allied Force in Kosovo some eight years later.

THE LEGALITY OF THE NO-FLY ZONES

The so-called "no-fly" zones over Iraq were created soon after the cessation of hostilities, with the northern exclusion zone — north of the 36th parallel — established in June, 1991 incident to Operation Provide Comfort, and the corresponding southern zone — south of the 32nd parallel — established in August, 1992.[6] The southern zone was ultimately extended in 1996 to cover all the airspace south of the 33rd parallel.[7] The ostensible purpose of the zones was to protect the Kurds in the north and the Shiites in the south from Iraqi oppression. My concern regarding these exclusion zones has less to do with their stated purpose than with their foundation in law.

The legal argument put forth to justify the creation of the "no-fly" zones over northern and southern Iraq rests upon an interpretation of and necessary interaction between four Security Council resolutions. The first is one I have already mentioned — Resolution 678 of November 29, 1990, which authorized the coalition to use force in implementing the previous eleven resolutions and to restore peace and security in the area.[8]

The second is Resolution 686[9] of March 2, 1991, which noted the suspension of combat activities but reaffirmed that Resolution 678 and its eleven predecessors were still in "full force and effect" and went on to demand that Iraq implement its acceptance of all of them.[10] More particularly, the resolution demanded that Iraq rescind its purported annexation of Kuwait; accept its liability under international law for the loss, damage and injury to all countries as a result of the invasion and occupation of Kuwait; release all Kuwaiti and third-state nationals and prisoners of war; return all seized Kuwaiti property; cease any hostile or provocative actions; designate military commanders to meet with coalition commanders to arrange for military aspects of a cessation of hostilities; and provide information and assistance regarding the identification of mines, booby

---

5. *See* U.N. CHARTER art. 42.

6. *See* Alain E. Boileau, *To The Suburbs of Baghdad: Clinton's Extension of the Southern Iraqi No-Fly Zone*, 3 J. INT'L & COMP. L. 875, 881 (1997).

7. SIMON CHESTERMAN, JUST WAR OR JUST PEACE? 199 (2001).

8. *See supra*, note 2, and accompanying text.

9. S.C. Res. 686, U.N. SCOR, 46th Sess., 2978th mtg., U.N. Doc. S/RES/686 (1991).

10. *See id.*

traps, explosives, and chemical and biological weapons.[11] Finally, Resolution 686 stated that the specific authorization of force contained in the second paragraph of Resolution 678 would remain valid during the time required for Iraq to comply with these demands.[12]

The third resolution with relevance to the purported legality of the "no-fly" zones is Resolution 687[13] of April 3, 1991, which recalled and affirmed the prior thirteen resolutions "except as expressly changed . . . to achieve the goals of the present resolution, including a formal cease – fire".[14] It thereafter set forth certain conditions upon Iraq, notably the requirement for the destruction, removal or rendering harmless of all chemical and biological weapons and ballistic missiles. Paragraph thirty-three of the resolution stated that once Iraq notified the Security Council of its acceptance of the listed conditions, a formal cease – fire would be in effect.

The final resolution bearing on this issue is Resolution 688[15] of April 5, 1991. In it, the Security Council addressed and condemned the oppression of the Iraqi civilians, especially in Kurdish populated areas, which it said threatened international peace and security in the region. The Council demanded that Iraq end the oppression and insisted that international humanitarian organizations be granted immediate access to those in need in the country. Finally, it appealed to all member states and humanitarian organizations to contribute to relief efforts. Of particular significance is the fact that the Security Council did *not* state anywhere in the resolution that it was acting pursuant to Chapter VII of the Charter, a phrase oftentimes found when coercive measures are being authorized, nor did it contain any authorization for protection of humanitarian relief efforts.

Those who argue that the "no-fly" zones are clearly supported in law suggest that the authorization for the use of force contained in paragraph two of Resolution 678 has never been rescinded.[16] Accordingly, that provision remains viable and may be invoked whenever there is a perceived breach of the cease – fire conditions under Resolution 686, or any other incident threatening the peace or stability of the region, such as a violation of the conditions regarding the destruction or removal of

---

11. *See id.*

12. *See id.*

13. S.C. Res. 687 U.N. SCOR, 46th Sess., 2981st mtg., U.N. Doc. S/RES/687 (1991).

14. *Id.*

15. S.C. Res. 688 U.N. SCOR, 46th Sess., 2982nd mtg., U.N. Doc. S/RES/688 (1991).

16. *See* Timothy P. McIlmail, *No-Fly Zones: The Imposition and Enforcement of Air Exclusion Regimes Over Bosnia and Iraq*, 17 LOY. L.A. INT'L & COMP. L.J. 35, 51 (1994).

chemical and biological weapons under Resolution 687. Another variant of the argument is that the provisions of Resolution 688 imply an enforcement mechanism to ensure that Iraq ends its suppression of its people, and that the "no-fly" zones are just such an enforcement mechanism.[17] If the "no-fly" zones have a proper legal predicate under Security Council resolutions, the argument goes, then force can be used in "self-defense" in response to attacks or threats of attack by Iraqi air defense facilities. I disagree not with this last issue of the use of force in self-defense, but rather with the underlying argument supporting the "no-fly" zones themselves.

While the wording of Resolution 678 certainly authorizes the coalition to "use all necessary means to uphold and implement resolution 660 (1990) *and all subsequent resolutions* . . . ;"[18] I would contend the phrase "all subsequent resolutions" clearly refers only to those ten resolutions passed after 660 and before 678. To argue otherwise is to suggest that there is a continuing mandate for the coalition countries, individually or collectively, to use force under the auspices of Operation Desert Storm in dealing with any breach of future resolutions involving Iraq or any threat to peace in that area of the world. Since the Security Council indicated that *it* was remaining seized of the matter, I find this argument untenable.

With regard to the assertion that authority for the "no-fly" zones emanates from or is a legal complement to Resolution 688, I find this equally problematic. As mentioned before, this resolution does not mention that the Security Council is acting under Chapter VII, nor does it purport to sanction, explicitly or even implicitly, the use of enforcement action. Its primary and stated focus is to require that Iraq grant unfettered access to international humanitarian organizations in order to give assistance to those in need, and also to appeal to all member states and humanitarian organizations to contribute to the relief efforts. Granted, the resolution "condemns" the Iraqi suppression of its people and "demands" that it cease the suppression, but it falls far short of any possible grant of authority to impose "no-fly" zones for the purpose of ensuring Iraqi compliance with its terms.

In sum, even though the Security Council has not spoken out to challenge the exclusion zones being maintained by American and British fighter aircraft, I do not see in that silence an affirmation of the zones. Without the firm legal predicate for us to be within the sovereign airspace of Iraq, the justification for our use of force in self-defense is greatly weakened.

---

17. *See* CHESTERMAN, *supra* note 7, at 198.
18. S.C. Res. 678, *supra* note 2.

HAVE THE ZONES ACHIEVED THEIR PURPOSE?

With regard to what was accomplished in Desert Storm back in 1991, there can be no argument that it was a successful military campaign with a remarkably low casualty rate. The criticism of that operation, now more in vogue than ever before, focuses on the belief that the ground offensive was terminated too early — that we did not "finish" the job by removing Saddam Hussein from power and toppling his regime.[19] In the wake of the September 11th terrorist attacks on the World Trade Center and the Pentagon, and the assumption that Iraq may have been involved in some way in those attacks, or perhaps in the anthrax attacks through use of the mail, much is now being said about the need for the United States to "go to Baghdad" and complete the job that George Bush Sr. failed to do back in 1991. I was in uniform back in those days and there were several reasons why we did not press the attack in Baghdad. First of all, from a purely legal perspective, we did not believe that the Security Council resolutions, individually or collectively, gave us the authority to do anything other than dislodge Saddam Hussein from Kuwait. Second, we were concerned with the immense and protracted logistical effort which would have been required in any occupation of Iraq. Third, we were also concerned that toppling the Iraqi government at that time would have created a power vacuum upon which Iran could have easily capitalized. Finally, we believed, erroneously in hindsight, that we could always go back and get rid of Saddam Hussein whenever we wanted and through a variety of means. For whatever it is worth, that was our reasoning at the time.

In appraising the measure of success of the "no-fly" exclusion zones over the last decade, we need to ask whether they have achieved their stated purpose of protecting the Kurds in the north and the Shiites in the south. On balance, we would probably all agree that they have. The minority populations have been shielded from the type of Iraqi armed repression that immediately followed the cessation of hostilities in 1991.[20] In addition, having aircraft in the skies over northern and southern Iraq has probably yielded additional dividends in providing excellent intelligence on what Saddam is doing, as well as restricting his movement of military assets within the country.[21] I would suggest, however, that enforcement of the "no-fly" zones has eclipsed its original purpose and has become an

---

19. *See* Daniel Pipes & Jonathan Schanzer, *On to Baghdad?: Yes –The Risks Are Overrated*, N.Y. POST, Dec. 3, 2001; *see also* Editorial, *Good Cop, Bad Cop*, ST LOUIS POST-DISPATCH, Sept. 26, 2001, at B6.

20. *See Dual Purpose to No-Fly Zones Imposed on Iraq*, PATRIOT LEDGER (Quincy MA.), Dec. 30, 1998, at 3.

21. *See id.*

integral part of the much larger policy of containment; a policy, which in addition to the "no-fly" zones, also includes U.N. imposed inspections for weapons of mass destruction and a sanctions regime. With regard to this overall policy of containment, the record is probably more mixed. Saddam Hussein still appears to be fully in control and I believe there is little likelihood that he will be taken down from within. As to the inspections for weapons of mass destruction, inspectors have not been in the country since December 1998. Furthermore, conjecture has it that Saddam secreted away thousands of pounds of materials and that he is probably also rebuilding his nuclear, chemical and biological weapons programs.[22] The sanctions regime, originally imposed by Security Council resolution 661[23] on August 6, 1990, is linked to these weapons inspections since Security Council resolution 687[24] provides that the sanctions are to continue until Iraq has destroyed, removed or rendered harmless its weapons of mass destruction. But the sanctions regime is arguably ineffective at best and counterproductive at worst. Shipments of food, medicine and equipment to rebuild the country's essential infrastructure are immersed in massive amounts of red tape, and consequently delayed, while other products such as DVD players and electronic equipment bypass the "oil for food" inspection system and can be bought openly if one has enough money.[25] To a certain extent, the sanctions regime plays into Saddam's hand because it has provided him with an excuse for the suffering of his people.[26] That is why there have been constant calls for the use of "smart sanctions" to replace the current regime, sanctions that would more precisely target materials which could be used in weapons systems while alleviating to a much greater measure the plight of the Iraqi citizenry.

We, therefore, appear to be at an impasse with our interconnected policy of containment.[27] We have used fairly sizeable air strikes on occasion to try to force compliance — notably Desert Strike in 1996 where we sought to punish Saddam for sending his troops to the north to assist a pro-Iraqi Kurdish faction against a pro-western Kurdish faction, and Desert

---

22. *See* Robert A. Pape, Editorial, *Our Iraq Policy is Not Working*, N. Y. TIMES, Feb. 24, 2001, at A13.

23. S.C. Res. 661, U.N. SCOR, 45th Sess., 2933rd mtg., U.N. Doc. S/RES/661 (1990).

24. S.C. Res. 687, *supra* note 13, ¶ 22.

25. *See* Michael Slackman, *Response to Terror*, L. A. TIMES, Oct. 19, 2001, at A1; Larry Kaplow, *Iraq: Years of Sanctions Hurt Iraqis More than Regime*, ATLANTA J. & CONSTITUTION, Oct. 20, 2001, at 6A.

26. *See* Yahia Said, Letter, *Why I Oppose Sanctions*, GUARDIAN (London), Jan. 9, 2001, at 21.

27. *See Iraq Airstrikes: Unfinished Business*, HERALD SUN, Aug. 19, 2001, at A12.

Fox in 1998 when we tried to force him to allow access by UNSCOM inspectors to facilities within the country — but the results were marginal. Further, it is clear that Saddam Hussein has little incentive to subject his country to renewed weapons inspections without some assurance that sanctions might be lifted in return, which I believe to be highly unlikely in light of current events.[28] Finally, there is a high cost to maintaining the "no-fly" enforcement zones over the north and south. Some 150-200 aircraft are used on a daily basis to fly and support enforcement missions, and those assets are greatly in demand elsewhere in the world.[29] Also, remember that a U-2 surveillance aircraft, flying at 70,000 feet, came very close to being shot down last summer by a modified Iraqi SAM-2 missile; and the Iraqis were successful in shooting down an unmanned Predator aircraft shortly thereafter, a surveillance platform costing almost $3.2 million each.[30] With Saddam determined to continue to use his air defense systems to engage aircraft flying the enforcement missions, it is probably only a matter of time before we lose a piloted aircraft. These missions are costly in other ways. Turkey's Prime Minister has expressed concerns about the frequent clashes in the northern "no-fly" zone and has indicated it could undermine support in his country for our continued use of Incirlik Air Base for the missions.[31]

CONCLUSION

Simply put, I believe that the "no-fly" exclusion zones in northern and southern Iraq rest not only on a dubious legal predicate, but have also produced marginal results when weighed against the costs involved. There have been recent indications, albeit before the tragic events of September 11th, that the Bush administration was re-evaluating the exclusion zones with a view towards possibly reducing or even eliminating this aspect of its overall containment policy of Saddam Hussein. That would certainly be well-advised.

---

28. *See* Roula Khalaf, *Disarming Baghdad Attempt to Break Deadlock*, FINANCIAL TIMES (London), June 17, 1999, at 8.

29. *See* Pamela Hess, *Pentagon Weighing Iraq Options*, UPI, Mar. 28, 2001.

30. *See* Thom Shanker, *U.S. Tells Iraq It May Retaliate for Missile Attack on Spy Plane*, N. Y. TIMES, July 27, 2001, at A4; Editorial, *Drone, But Also an Omen*, COMMERCIAL APPEAL (Memphis, TN), Aug. 29, 2001, at A6.

31. *See Iraq: Defying "no-fly" Zones Continues*, UPI, Feb. 11, 1999.

Humanitarian Intervention and International Law in the Case of Kosovo

Major General William Moorman*

On March 24, 1999, without the benefit of a U.N. Security Council resolution expressly authorizing military action, North Atlantic Treaty Organization (NATO)[1] began a seventy-eight day air campaign[2] over the Former Republic of Yugoslavia (FRY). The air campaign concluded with Slobodan Milosevic[3] capitulating to the NATO demands. NATO's actions also laid the predicates for U.N. Security Counsel Resolution 1244,[4]

---

* The Judge Advocate General, United States Air Force.

1. The North Atlantic Treaty was signed in Washington, D.C. on 4 April 1949, creating an alliance of twelve independent nations committed to each other's defense. *See* North Atlantic Treaty, Apr. 4, 1949, 63 Stat. 2241, 34 U.N.T.S. 243. Four more European nations later acceded to the Treaty between 1952 and 1982. On 12 March 1999, the Czech Republic, Hungary and Poland were welcomed into the Alliance, bringing the total to nineteen members. The North Atlantic Treaty continues to guarantee the security of its members. Further information about NATO is available on its website at <http://www.nato.int/home.htm>.

2. The name of this operation was Operation Allied Force. The operation's objective was "to degrade and damage the military and security structure that President Milosevic (Yugoslav President) has used to depopulate and destroy the Albanian majority in Kosovo." Prepared Statement, William S. Cohen, Secretary of Defense, to the Senate Armed Services Committee (Apr. 15, 1999) http://www.defenselink.mil/specials/kosovo/ (last visited Apr. 19, 2002).

3. Slobodan Milosevic was the President of FRY at the time of Operation Allied Force. At the present time, Milosevic is facing the ICTY (International Criminal Tribunal for the Former Yugoslavia). *See* Indictments and Proceedings, *at* http://www.un.org/icty/ind-e.htm (last visited Apr. 19, 2002). "Following the transfer of Slobodan Milosevic to the custody of the ICTY on 29 June 2001, the Trial Chamber entered a 'not guilty' plea on behalf of the Accused at his Initial Appearance 3 July 2001." Status of Cases *at* http://www.un.org/icty/glance/casestatus.htm (last visited March 29, 2002). For general information regarding the Milosevic case and the ICTY, see the United Nations ICTY website *at* http://www.un.org/icty/ (last visited Apr. 19, 2002).

4. *See* S.C. Res. 1244, U.N. SCOR, 54th Sess., 4011th mtg., U.N. Doc. S/RES/1244 (1999).

According to the official NATO website:

adopted June 10, 1999, the day after the halt of NATO air strikes.

The conclusion of these NATO operations and adoption of U.N. Security Council Resolution 1244 resulted in the deployment of international security forces in Kosovo. The responsibilities of the forces were, and continue to be, to:

> (a) deter renewed hostilities; (b) maintain and enforce a cease-fire; (c) ensure the withdrawal of Yugoslav and Serb military, police and paramilitary forces from Kosovo; (d) establish a secure environment for refugees in which displaced persons can return home safely; (e) supervise de-mining; (f) conduct border monitoring; (g) demilitarize the Kosovo Liberation Army[5] and other armed Kosovo Albanian groups; and (h) "[e]nsur[e] the protection and freedom of . . . movement of the international civil presence, and other international organizations" in Kosovo.[6]

Implementation of these responsibilities through a Military Technical Agreement[7] between NATO and FRY put an effective halt to independent

---

> On 10 June the UN Security Council passed a resolution (UNSCR 1244) welcoming the acceptance by the Federal Republic of Yugoslavia of the principles on a political solution to the Kosovo crisis, including an immediate end to violence and a rapid withdrawal of its military, police and paramilitary forces. The Resolution, adopted by a vote of 14 in favour and none against, with one abstention (China), announced the Security Council's decision to deploy international civil and security presences in Kosovo, under United Nations auspices.

NATO and Kosovo: Historical Overview, *available at* http://www.nato.int/kosovo/history.htm (last visited Apr. 2, 2002).

5. The Kosovo Liberation Army (KLA) is an ethnic Albanian insurgent group. For an excellent background on the KLA see in general, Lt. Col. Michael E. Smith, *NATO, the Kosovo Liberation Army, and the War for an Independent Kosovo: Unlawful Aggression or Legitimate Exercise of Self-Determination?*, Feb 2001, ARMY LAW., at 1.

6. U.N. S.C. Res. 1244, U.N. SCOR 54th Sess., 4011th mtg. ¶ 9(h), U.N. Doc. S/RES/1244 (1999).

7. For a complete copy of the Military Technical Agreement, see http://www.nato.int/kosovo/docu/a990609a.htm (last visited May 9, 2002). According to the official NATO website:

> On 10 June 1999, after an air campaign lasting seventy-seven days, NATO Secretary General Javier Solana announced that he had instructed General Wesley Clark, Supreme Allied Commander Europe, temporarily to suspend NATO's air operations against Yugoslavia. This decision was taken after consultations with the North Atlantic Council and confirmation from General Clark that the full withdrawal of Yugoslav forces from Kosovo had begun.
>
> The withdrawal was in accordance with a Military-Technical Agreement concluded between NATO and the Federal Republic of Yugoslavia on the evening of 9 June. The agreement was signed by Lt. General Sir Michael Jackson, on behalf of NATO, and by Colonel General Svetozar Marjanovic of the Yugoslav Army and Lieutenant General Obrad Stevanovic of the Ministry of Internal Affairs, on behalf of the Governments of the Federal Republic of Yugoslavia and Republic of Serbia. The withdrawal was also consistent with the agreement between the Federal Republic of Yugoslavia and the European Union and Russian special envoys, President Ahtisaari of Finland and Mr. Victor Chernomyrdin, former Prime Minister of Russia, reached on 3 June.
>
> The NATO Secretary General announced that he had written to the Secretary-General of the United Nations, Mr. Kofi Annan, and to the President of the United

NATO operations, but important for the purposes of our discussion, did nothing to resolve lingering issues surrounding the legal basis for NATO's actions in the first place. In fact, these issues continue to be debated today.

It appears that the difficulty does not stem from NATO's inability to speak with one voice regarding the legal basis of its involvement in Kosovo and Serbia -- although that inability has certainly contributed to the debate. Rather, it seems that the debate is grounded in differing perceptions of the likely long-term implications of the Kosovo intervention as it potentially affects the development of customary international law.[8]

On one end of the spectrum is the argument that Article 2 of the U.N. Charter[9] prohibits intervention in the internal affairs of a sovereign nation

---

> Nations Security Council, informing them of these developments. The Secretary General of NATO urged all parties to the conflict to seize the opportunity for peace and called on them to comply with their obligations under the agreements which had now been concluded and under all relevant UN Security Council resolutions.

NATO and Kosovo: Historical Overview *available at* http://www.nato.int/kosovo/history.htm (last visited Apr. 3, 2000).

8. *See generally* John J. Merriam, *Kosovo and the Law of Humanitarian Intervention*, 33 CASE W. RES. J. INT'L L. 111 (2001); Louis Henkin, Comment, *NATO's Kosovo Intervention: Kosovo and the Law of "Humanitarian Intervention,"* 93 AM. J. INT'L L. 824 (1999); HENRY J. STEINER & PHILIP ALSTON, INTERNATIONAL HUMAN RIGHTS IN CONTEXT (2nd ed., 2000); Antonio Cassese, *Ex iniuria ius oritur:Are we moving towards International Legitimation of Forcible Humanitarian Countermeasures in the World Community?* 10 EUR.J. INT'L L. 23 (1999); Antonio D'Amato, *International Law and Kosovo*, 33 UN L. R. 112 (1999); Captain J Blackett, *The Legal Basis for the NATO Campaign in Kosovo, in* INTERNATIONAL SOCIETY FOR MILITARY LAW AND THE LAW OF WAR XV INTERNATIONAL CONGRESS (Working Session, Lillehammer, June 6-10, 2000); Adam Roberts, *Lessons From the 1999 Kosovo War and Other Cases, in* INTERNATIONAL SOCIETY FOR MILITARY LAW AND THE LAW OF WAR XV INTERNATIONAL CONGRESS (Working Session, Lillehammer, June 6-10, 2000); Dr. Peter Hilpold, *Humanitarian Intervention: Is There a Need for a Legal Reappraisal?, in* INTERNATIONAL SOCIETY FOR MILITARY LAW AND THE LAW OF WAR XV INTERNATIONAL CONGRESS (Working Session, Lillehammer, June 6-10, 2000); Simon Chesterman, JUST WAR OR JUST PEACE? HUMANITARIAN INTERVENTION AND INTERNATIONAL LAW (2001); Robert F. Turner, *Kosovo: Legal and Policy Considerations*, 10 US AIR FORCE ACAD. J. LEG. STUD. 67 (1999-2000); Daniel Lauretano, *Operation Allied Force: Was NATO's Use of Force Lawful and Did NATO Commit War Crimes?* War Crimes and Prosecution Seminar Paper, Georgetown University Law Center, (2000) (unpublished paper on file with author); Lt. Col. Michael E. Smith, *NATO, the Kosovo Liberation Army, and the War for an Independent Kosovo: Unlawful Aggression or Legitimate Exercise of Self-Determination?* 2001 ARMY LAW. 1 (Feb. 2001); George K. Walker, *Principles for Collective Humanitarian Intervention to Succor Other Countries' Indigenous Nationals*, (July 2001) (unpublished manuscript provided, Naval Justice School, Conference on Kosovo War) (on file with author); HUGO GROTIUS, DE IURE BELLI AC PACIS (1625); IAN BROWNLIE, INTERNATIONAL LAW AND THE USE OF FORCE BY STATES (1968); THOMAS BUERGENTHAL, INTERNATIONAL HUMAN RIGHTS IN A NUTSHELL (1995).

9. United Nations Charter Article 2 states;

> The Organization and its Members, in pursuit of the Purposes stated in Article 1, shall act in accordance with the following Principles:

based on humanitarian imperatives in the absence of express U.N. Security Council authorization. Those who make this argument fear, with some historically-based logic, that any action which creates an exception to the clear meaning of Articles 2(4)[10] and 2(7)[11] potentially weakens the very basis for the U.N. itself.[12] At the other end of the spectrum is the argument that Kosovo epitomizes the case for legitimate humanitarian intervention independent of U.N. authorization.[13] Many who articulate this view cite

---

1. The Organization is based on the principle of the sovereign equality of all its Members.
2. All Members, in order to ensure to all of them the rights and benefits resulting from membership, shall fulfill in good faith the obligations assumed by them in accordance with the present Charter.
3. All Members shall settle their international disputes by peaceful means in such a manner that international peace and security, and justice, are not endangered.
4. All Members shall refrain in their international relations from the threat or use of force against the territorial integrity or political independence of any state, or in any other manner inconsistent with the Purposes of the United Nations.
5. All Members shall give the United Nations every assistance in any action it takes in accordance with the present Charter, and shall refrain from giving assistance to any state against which the United Nations is taking preventive or enforcement action.
6. The Organization shall ensure that states which are not Members of the United Nations act in accordance with these Principles so far as may be necessary for the maintenance of international peace and security.
7. Nothing contained in the present Charter shall authorize the United Nations to intervene in matters which are essentially within the domestic jurisdiction of any state or shall require the Members to submit such matters to settlement under the present Charter; but this principle shall not prejudice the application of enforcement measures under Chapter VII.

U.N. CHARTER art. 2, para. 1-7.

10. U.N. CHARTER art. 2, para. 4 states: "All Members shall refrain in their international relations from the threat or use of force against the territorial integrity or political independence of any state, or in any other manner inconsistent with the Purposes of the United Nations." *Id.*

11. U.N. CHARTER art. 2, para. 7 states;

> Nothing contained in the present Charter shall authorize the United Nations to intervene in matters which are essentially within the domestic jurisdiction of any state or shall require the Members to submit such matters to settlement under the present Charter; but this principle shall not prejudice the application of enforcement measures under Chapter VII.

Id.

12. *See generally* CHESTERMAN, *supra* note 8. In his book, Chesterman argues:

> [T]hat such unilateral enforcement is not a substitute for but the opposite of collective action. Though often presented as the only alternative to inaction, incorporating a 'right' of intervention would lead to more such interventions being undertaken in bad faith, it would be incoherent as a principle, and it would be inimical to the emergence of an international rule of law.

Id. at 6.

13. *See generally* Merriam, *supra* note 8; *see generally* Henkin, *supra* note 8; *see generally* D'Amato, *supra* note 8; *see generally* Turner, *supra* note 8; *see generally* Lauretano, *supra* note 8; *see also* LT. COL. RICHARD J. ERICKSON, LEGITIMATE USE OF MILITARY FORCE AGAINST STATE-SPONSORED INTERNATIONAL TERRORISM 188 (1989) ("Jurists such as Grotius, Vattel, Wheaton, Heilberg, Woolsey, Bluntschli, Westlake, Stowell, Lawrence, and

customary international law predating the U.N. Charter. Proponents of this view assert that the right to humanitarian intervention has long existed and must continue to exist notwithstanding the U.N. charter.

And so, the debate is framed. The fact that legal scholars find themselves on either side of this argument only serves to illustrate the serious nature of the debate. Each scholar fears the logical or extreme outcome of acceptance of the other's view. On the one hand, one argument could lead to claims of legitimacy when one sovereign, for its own evil purposes, seeks to exterminate portions of its population without fear of outside interference. On the other hand, the opposing argument could legitimize the actions of a sovereign seeking to insert itself directly into the internal affairs of another state, perhaps cynically citing its right of humanitarian intervention as a cloak with which to disguise its aspirations for regional hegemony.

I believe that there is a sensible and legally defensible alternative approach that is consistent with long accepted legal principles. That middle ground is found in a fact-based analysis of the situation.[14] While such a fact-based analysis is somewhat dissatisfying because it clearly accepts the continued existence of ambiguity, I believe it leads to an acceptable analysis of the meaning of the Kosovo intervention. Certainly, it is more acceptable than the conclusion, reached by some, that NATO's intervention was a violation of international law, albeit a *necessary* violation.

By acknowledging the need for a fact-based analysis, we would also accept the fact that the Kosovo example was complicated and, even, exceptional. In the final analysis, any argument for the use of force in the absence of Security Council sanction might benefit from being treated as an exception rather than a rule.

I. WHAT MADE KOSOVO EXCEPTIONAL?

1. Kosovo Had a Unique History of U.N. Security Council Scrutiny

Beginning in 1998, Kosovo was the subject of a series of U.N. Security Council resolutions. U.N. Security Council Resolution 1160,[15] adopted

---

Borchard claim that customary international law recognizes the doctrine of legal intervention in situations where a state denies its people fundamental human rights in a way that shocks the conscience of mankind.").

14. For an excellent discussion and analysis of *possible* conditions for humanitarian intervention, see ERICKSON, *supra* note 13, at 191-193.

15. U.N. S.C. Res. 1160, U.N. SCOR 53rd Sess., 3868th mtg., UN Doc S/RES/1160 (1998). This resolution was issued pursuant to Chapter VII of the U.N. Charter, authority in light of reports of excessive force by Serbian police against civilians and peaceful demonstrators in Kosovo, and reports of acts of terrorism by the KLA.

pursuant to Chapter VII[16] on 31 March, took note of reports of use of excessive force by Serbian police against civilians in Kosovo.[17] It called upon the FRY to pursue a political settlement and to open a meaningful dialog with the Kosovar Albanians while recognizing the territorial integrity of FRY. At the same time, it called on the KLA to condemn terrorist actions and to pursue peaceful settlement of its goals. Finally, it clearly indicated the Security Council's intention to remain seized of the issues and to support the ICTY in gathering evidence of possible crimes.

Within six months, on 23 September 1998, the Security Council adopted Resolution 1199.[18] That Resolution expressed grave concern over the fighting in Kosovo, particularly citing the excessive and indiscriminate use of force by Serbian security forces and the Yugoslav military. The Security Council demanded of the FRY and the Kosovo Albanians: a cessation of hostilities, immediate action to improve the humanitarian situation to avert a humanitarian catastrophe, and to open a meaningful dialog aimed at political settlement. The Council specifically demanded of the FRY that it: call back its security forces and cease their actions affecting the civilian population; that it facilitate international monitoring; that it facilitate the safe return of refugees; and, that it make rapid progress in negotiations toward a political solution. In addition to other expressions along similar lines, it decided that if concrete measures were not taken to meet the dictates of the two then existing Resolutions, the Council would consider further action and additional measures to maintain peace and stability in the region.

Finally, on 24 October 1998, the Security Council adopted Resolution 1203,[19] restating many of its prior demands, specifically recognizing that the impending humanitarian catastrophe constituted a threat to continuing peace and stability in the region. It also endorsed the agreements of 15 and

---

16. U.N. CHARTER, art. 39-51 ("Action with Respect to Threats to the Peace, Breaches of the Peace and Acts of Aggression").

17. For a comprehensive list of human rights abuses in Kosovo and a complete chronology of events, see Kosovo Chronology, *available at* http://www.state.gov/www/regions/eur/fs_kosovo_timeline.html (last visited Apr. 19, 2002).

18. U.N. S.C. Res. 1199, U.N. SCOR, 53rd Sess., 3930th mtg., U.N. Doc S/RES/1199 (1998). This resolution was issued pursuant to authority under Chapter VII of the U.N. Charter out of grave concern over the fighting in Kosovo, and in particular the excessive and indiscriminate use of force by Serbian security forces and the Yugoslav Army, which resulted in numerous civilian casualties and the displacement of over 230,000 persons. *See id.*

19. U.N. S.C. Res. 1203, U.N. SCOR, 53rd Sess., 3937th mtg., UN Doc S/RES/1203 (1998). This resolution was issued pursuant to U.N. Security Council authority under Chapter VII of the U.N. Charter in response to the continuing grave humanitarian situation in Kosovo and the then impending humanitarian catastrophe, and in the belief that the unresolved situation constituted a continuing threat to peace and security in the region.

16 October between the FRY, NATO, and the Organization for Security and Cooperation in Europe (OSCE).

This history created a unique environment in which the demands of the U.N. had been clearly articulated, the dangers of the impending crisis had been analyzed, and the Security Council had expressed a willingness to consider further action. However, the threat of a veto from Russia and China precluded the Security Council from taking the next logical step – the authorization of the use of force.[20] Finally, following NATO's aerial campaign, the action eventually taken in U.N. Security Council Resolution 1244, authorized the placement of a U.N. force to ensure actions were taken which met the NATO terms for conflict termination.

The Security Council's actions demonstrated, first, its growing concern over the developing crisis and its belief that the situation posed a risk to regional peace and security. Secondly, the Security Council's failure to act, demonstrated that it was incapable of addressing the apparent need for international intervention; and by its overt endorsement of the result of the ensuing NATO intervention, expressed post-action approval of that effort. This last point may be critical to any legal analysis. When given the opportunity to, in effect, either decry or implicitly ratify the action taken by NATO, the Security Council clearly chose the latter.[21]

2. The Character of the Human Rights Abuses Was Severe and Immediate

By its Resolutions, the Security Council evinced a clear understanding of the grave crisis building in Kosovo. On 30 January 1999, the U.N. Secretary General forwarded a report to the Security Council describing the humanitarian situation in Kosovo as "grave" and stressing the need for action to remedy the looming humanitarian disaster.[22] This report cited numerous atrocities committed primarily by the FRY army and Kosovo Serbs, including widespread violence against Kosovo Albanians -- rape, torture, large-scale civilian displacements, abductions, unlawful deten-

---

20. *See* Merriam, *supra* note 8, at 144-45.

> Despite the lack of an official recorded vote, deadlock was readily apparent to all observers even before the crisis broke out. The UN had been increasingly ineffective at averting such tragedies since Somalia, with notable failures in Rwanda and Bosnia. A paralyzed Security Council was unable to act in either case to prevent a horrible humanitarian tragedy. The NATO powers had no reason to expect anything different in regard to Kosovo . . . .

Id.

21. *See generally* U.N. S.C. Res. 1244, U.N. SCOR 54th Sess., 4011th mtg. ¶ 9(h), U.N. Doc. S/RES/1244 (1999).

22. *See* Report of the Secretary-General Prepared Pursuant to Resolutions 1160 (1998), 1199 (1998), and 1203 (1998) of the Security Council at 1, U.N. Doc. S/1999/99 (1999) *available at* http://www.un.org/Docs/sc/reports/1999/s199999.htm (last visited Apr. 19, 2002).

tions, and massacres of civilians. At least two mass graves had by now been unearthed. Subsequent investigations have only served to confirm the catastrophic nature of the crisis at that time.[23]

3. There Was No Realistic, Peaceful Alternative to NATO Intervention in Kosovo

For months, diplomatic efforts had been pursued in the hope of resolution of the situation in Kosovo. The United Nations, the OSCE, NATO, the U.S., and the Balkans Contact Group all participated in increasingly urgent diplomatic efforts designed to halt the growing crisis.[24] This unique combination of world and regional bodies, coupled with the efforts of individual states, all sought to solve the crisis without negating the Former Republic of Yugoslavia's sovereignty. The last diplomatic initiative failed with the collapse of the Rambouillet peace talks in February of 1999.[25] Even then, Richard Holbrooke made a last ditch attempt to avoid NATO military action. When that attempt failed, Operation Allied Force commenced.

4. An Identifiable, Internationally Recognized, Threat to Peace and Security in the Region Existed

The Kosovo situation resulted in hundreds of thousands of refugees who flooded across its borders into neighboring states. These states lacked any realistic ability to accommodate the refugees' needs. This, in turn, threatened to increase the instability of those surrounding states while creating additional pressure on NATO allies Greece and Turkey. This was more than a mere NATO perception. The threat was specifically recognized by the U.N. Security Council as one which threatened regional peace and stability.[26]

5. Recent History had Demonstrated a Willingness by the Fry Leadership to Promote Regional Instability and Human Suffering in Apparent Preference to Political Settlement

For a period of years, the collection of states constituting Yugoslavia

---

23. *See generally* Situation in Kosovo, *available at* http://www.state.gov/www/regions/eur/kosovo_more.html (last visited Apr. 19, 2002).

24. *See* Merriam, *supra* note 8, at 143.

25. For an excellent overview of the Rambouillet peace talks, see DEPARTMENT OF STATE TRAVELS WITH THE SECRETARY, *available at* http://secretary.state.gov/www/travels/1999/990219trip_remarks.html (last visited Apr. 19, 2002).

26. *See* U.N. S.C. Res. 1199, U.N. SCOR, 53rd Sess., 3930th mtg., U.N. Doc S/RES/1199 (1998).

had been sliding into political disintegration and humanitarian disaster. NATO, acting under U.N. auspices, had been deeply engaged in Bosnia in an effort to quell the fighting and incredible brutality, which had been visited on the peoples of that region. The FRY and Serbia, in particular, had been a ready and apparent ally in the hostilities that threatened to consume that collection of states and would-be states. There was no reason to believe that a Milosevic-led government would pursue its objectives in Kosovo with less brutality if left unconstrained.

II. WHAT DOES THIS THEN MEAN?

First, taken together, this unique collection of factors led to a compelling need for international intervention to prevent a more serious and disastrous outcome. On the eve of conflict, Sandy Berger, then the President's National Security Advisor, wrote to the Congressional leadership this explanation for then likely NATO intervention:

> NATO would be acting to deter unlawful violence in Kosovo that endangers the fragile stability of the Balkans and threatens a wider conflict in Europe, to uphold the will of the international community as expressed in various U.N. Security Council Resolutions, as well as to prevent another humanitarian crisis, which itself could undermine stability and threaten neighboring countries.[27]

What is now certain is that modern states (and their international law advisers), will not allow themselves to be paralyzed into inaction whenever they confront a situation which does not neatly fit fully developed international law, particularly when faced with state-sponsored slaughter of innocents. They will, and should, adopt flexible but legitimate legal analyses to deal with emerging or emerged evils.

Secondly, and flowing from my first point, there is unfinished work ahead. Legal scholars and practitioners face the challenge of correctly describing what Kosovo means to international law. They must work on refining the set of criteria for conferring legitimacy on this kind of operation. Some of the Kosovo-specific facts I have talked about earlier might well be the basis for such criteria. If there are to be solid boundaries for such interventions, then these boundaries must not only be reasonable, but must also be acceptable to the community of sovereign states. Already there are multiple views and opinions circulating, and it is doubtful that any of the experts will give up their positions easily. However, this is a scholarly debate that should not be avoided.

Finally, the Kosovo operation was further evidence that this generation

---

27. Letter from Samuel R. Berger, Assistant to the President for National Security Affairs to The Honorable Trent Lott, (Mar. 23, 1999) (on file with author).

of public international lawyers is privileged to be examining international legal issues during a time when international law is so dynamic, and so interwoven with the political, diplomatic, humanitarian, and military dimensions of the 21st Century world order. Kosovo, and events since, including the tragedy of September 11th, have presented unique legal issues and challenges, and it will require continued special attention to emerging circumstances to navigate what may be rough and uncharted waters ahead.

Intentional Targeting of Regime Elites: The Legal and Policy Debate

Robert F. Turner*

This program was planned long before the events of September 11th, which have given my topic an added emphasis. For reasons that I trust will be obvious, I want to expand the topic to include not only "regime elites," but also leaders of terrorist organizations like Osama bin Laden.

I first became interested in this issue more than twenty-five years ago, when I was National Security Adviser to Senator Robert Griffin, the Assistant Senate Minority Leader and a member of the Foreign Relations Committee. As a result of the Church and Pike Committee hearings, Congress was considering enacting an "Intelligence Charter" — and one of the prohibitions was going to be a ban on "assassination."

I vividly recall cornering the Senator during a staff picnic one beautiful Saturday afternoon, and explaining in detail why I believed it would be a mistake to codify such a rule in law. He listened patiently, and then — reflecting the political mood of the time — responded: "Bob, you make a very impressive argument. Now, promise me something. Promise me you will never mention this again to anyone!"

Well, I kept my silence until October 1990, after Saddam Hussein invaded Kuwait and Defense Secretary Dick Cheney cashiered Air Force Chief of Staff Michael Dugan for suggesting to the press that the United States would intentionally target Saddam Hussein. Cheney justified the decision by noting that "assassination" was prohibited by Executive Order 12333, which provides in Section 2.11:

> "*Prohibition on Assassination.* No person employed by or acting on behalf of the United States Government shall engage in, or conspire to engage in,

---

* Professor, General Faculty, and Associate Director, Center for National Security Law, University of Virginia School of Law. B.A., 1968, Indiana University; J.D., 1981, S.J.D., 1996, University of Virginia. President, 1986-87, U.S. Institute of Peace; Principal Deputy Assistant Secretary for Legislative and Intergovernmental Affairs, 1984-85, U.S. Department of State; Special Assistant to the Under Secretary of Defense for Policy, 1981-82; National Security Advisor to U.S. Senator Robert P. Griffin, 1974-79.

assassination."[1]

I published an article shortly thereafter in the *Washington Post* that, I gather, helped to kindle some of the modern debate on this issue;[2] and in October 1998 I did a sequel in *U.S.A. Today* on the issue of intentionally targeting Osama bin Laden and other terrorist leaders.[3]

Every now and then a really nasty character surfaces somewhere around the world — a Stalin, a Hitler, a Saddam Hussein, or, more recently, an Osama bin Laden. And when we "academics" get bogged down in talking about the latest techniques of "dispute resolution," Bubba out in Peoria slams his hand down on the table and asks: "*Why don't you just shoot the SOB!*"

Well I am here to suggest that Bubba's instinct might just be right — we need to *rethink our targeting doctrine* in response to major acts of armed international aggression — be it from State leaders like Saddam Hussein or freelance terrorists like bin Laden who are determined to murder vast numbers of our people.

TARGETING DOCTRINE BASED ON OLD RULES

The fact is that our targeting doctrine has not kept pace with the dramatic transformation that has taken place in international law governing the initiation and conduct of aggressive war.

In some respects, we still approach war intellectually in terms of a set of rules devised by kings and their representatives in an era when aggressive war was viewed by the law as a "sovereign prerogative" of kings. Not surprisingly, one of the first rules was: "Don't hurt the King!"

Soldiers were viewed as the "subjects" or "property" of kings — to be disposed of just as the king might dispose of a piece of land or some mineral resource.

Consequently, very little moral value was placed on the lives of soldiers. Unless they were hospitalized or similarly disabled, international law considered it virtually "open season" on an enemy's military personnel.

From a legal perspective, it really did not matter if you killed a few hundred thousand of another State's soldiers. War, after all, was about killing soldiers.

---

1. Exec. Order No. 12,333, 46 Fed. Reg. 59941 (Dec. 4, 1981).

2. *See* Robert F. Turner, *Killing Saddam: Would it Be a Crime?*, WASH. POST, Oct. 7, 1990, at D1.

3. *See* Robert F. Turner, *In Self-Defense, U.S. Has Right to Kill Terrorist bin Laden*, USA TODAY, Oct. 26, 1998, at 17A.

THE "INNOCENT SOLDIER"

Now I have a strong personal prejudice against this rule. It dates back to 1968, when I first went to Vietnam. I was in Vietnam a total of five times between 1968 and the evacuation in 1975 as a journalist, soldier, and Senate staff member, and during that time I encountered a lot of soldiers on all sides — American, South Vietnamese, Viet Cong, North Vietnamese, and even a few Australians and South Koreans — and I came away from the experience thinking that most of them were pretty decent kids.

Since then, I have seen soldiers in Central America and other areas of the world as well — and I have come up with something I call the concept of the "Innocent Soldier." Most of the kids that died in Vietnam were not responsible for the policies that sent them there. Similarly, most of Saddam Hussein's troops were not morally responsible for his acts of aggression against Iran and Kuwait. They did not have the option of just crossing the line into Canada and starting anew — at least not if they cherished the lives of their family members. And yet hundreds of thousands of them died because of his "illegal" actions.

GENTILI & VATTEL

Ironically, even in the 16th and 18th centuries, thoughtful international lawyers were questioning the wisdom of the "don't hurt the king" rule. Alberico Gentili wrote in his *The Law of War* in 1588 that "The law bids us . . . to punish only the leaders, in order that the multitude may be spared."[4]

Gentili notes the emergence of a rule to the contrary, writing: "Now it is the common soldiers who are slain. The leaders, the rich, are saved . . . . Our worthy leaders consult for their own interests in this new fashion; for if they should come into the hands of the enemy, they would no longer have to fear for their own lives . . . ."[5]

In 1758, Vattel displayed a similar disdain for the "don't hurt the king" rule, writing in Book III of *The Law of Nations*:

In former times he who . . . kill[ed] the king or general of the enemy was commended and greatly rewarded. . . . Nothing could have been more natural . . . . [T]he death of the leader [often] put an end to the war. At the present day, a soldier would not dare . . . to boast of having killed the enemy's king. Thus tacitly agreed among sovereigns that their persons shall

---

4. ALBERICO GENTILI, DE IURE BELLI LIBRI TRES (1612), *reprinted in* 2 CLASSICS OF INTERNATIONAL LAW, at 322 (John C. Rolfe trans., James Brown Scott ed., William S. Hein & Co., 1995).

5. *Id.* at 325.

be held sacred.[6]

If there was ever any logic in giving special protection to the other side's leaders in war, it ceased when aggressive war ceased being recognized as a sovereign prerogative of kings and became instead an international crime.

The use of armed force as an instrument of policy was outlawed in the 1928 Kellogg-Briand Treaty,[7] and even more clearly by article 2(4) of the U.N. Charter.[8]

NUREMBERG PRINCIPLES

Since the London Charter of 1945 and the subsequent Nuremberg Principles, international law has recognized the planning and launching of "aggressive war" to be a criminal act.[9]

Indeed, Nuremberg stands for the principle that even heads of state—who are normally immune from foreign criminal process—may be arrested, tried, and even *executed* for the crime of aggressive war.[10]

Aggressive war is now a *crime*, and Regime Elites who engage in aggressive war are international *criminals*. And yet — despite that widely recognized face — the world continues to respond to aggression by slaughtering the children of the aggressor State. Where else in any civilized legal system does society direct its punishment for illegal conduct at the children, neighbors, or countrymen of the criminal? Who would call such an approach "just?"

But, if you really think about it, that is exactly what we are doing. We conclude that Saddam Hussein has launched an aggressive war against Kuwait, and then we send our kids out to slaughter his country's kids in the desert in the hope that he will rethink his ways and call a truce.

To be sure, the use of lethal force is, and under international law must be, a "last resort." We often begin — as we did in the Gulf Crisis — by imposing an economic embargo upon the "Aggressor State." What that

---

6. E. DE VATTEL, LE DROIT DES GENS, OU PRINCIPES DE LA LOI NATURELLE, APPLIQUES A LA CONDUITE ET AUX AFFAIRES DES NATIONS ET DES SOUVERAINS, [THE LAW OF NATIONS OR THE PRINCIPLES OF NATURAL LAW APPLIED TO THE CONDUCT AND TO THE AFFAIRS OF NATIONS AND OF SOVEREIGNS] (1758), *reprinted in* 3 CLASSICS OF INTERNATIONAL LAW, at 290 (Charles G. Fenwick trans., James Brown Scott ed., William S. Hein & Co., 1995).

7. *See* Kellogg-Briand Pact 1928, Aug. 27, 1928, *available at* www.yale.edu/lawweb/avalon/imt/kbpact.htm (last visited Feb. 6, 2002).

8. *See* U.N. CHARTER art. 2, para. 4.

9. Nuremberg Trial Proceedings, Charter of the International Military Tribunal, Aug. 8, 1945, *available at* www.yale.edu/lawweb/avalon/imt/proc/imtconst.htm (last visited Feb. 19, 2002).

10. *See id.*

means is that the Saddam Husseins of the world will take what they want by force and leave the women and children who have the misfortune to live under such a Regime to fend for themselves. The bulk of the suffering from the embargo of Iraq was not felt by Saddam or his cronies — they had enough food to eat — the real pain was imposed upon relatively innocent people who had no control over policy.[11]

This raises a *fundamental question*: Is it morally right to punish the people of Iraq for the criminal wrongdoing of a self-imposed dictator?

DETERRENCE IMPLICATIONS

Even if one accepts that it is morally permissible to direct the brunt of the defensive response to armed international aggression at relatively innocent soldiers and civilians, especially the kids, it is a bad policy because it *does not work*!

Saddam Hussein sent hundreds of thousands of Iraq's children off to die in his war of aggression against Iran, and he was clearly willing to sacrifice a few hundred thousand more in his aggression against Kuwait.

When the Kurds became a nuisance, he attacked them with mustard gas — in gross violation of international law.

Lessons of Jefferson's Embargo

I come from the University of Virginia — Thomas Jefferson country. Jefferson was a man of extraordinary brilliance, and as such he recognized "war" to be a great evil. In a May 1, 1794, letter to Tench Coxe, Jefferson proposed an alternative to warfare against Great Britain, arguing: "[A]s to myself, I love peace, and I am anxious that we should give the world still another useful lesson, by showing to them other modes of punishing injuries than by war, *which is as much a punishment to the punisher as to the sufferer.*"[12]

Thirteen years later, as his second term as President was approaching its final year, he attempted yet another alternative to war by persuading Congress to enact the Embargo Act of 1807.[13]

In response to the impressments of American seamen and other indigni-

---

11. I am not suggesting that the U.N. embargo ought to be terminated. I would have preferred a more carefully focused sanctions policy designed to maximize discomfort for Saddam Hussein and his key lieutenants, but the long-term consequences, if he is permitted to succeed in developing nuclear and other Weapons of Mass Destruction, will dwarf the current suffering and must be avoided if at all possible.

12. Letter from Thomas Jefferson, to Tench Coxe (May 1, 1794) *available at* www.yale.edu/lawweb/avalon/jefflett/more/let105.htm (last visited Feb. 19, 2002) (emphasis added).

13. *See* Embargo Act of 1807, ch. 5, 2 Stat. 451 (1807).

ties by Great Britain and France alike, Jefferson decided to use economic coercion as a substitute for arms — prohibiting virtually all exports from the United States. Unfortunately, the people in Europe who were most immediately affected by the economic pressure (such as laid-off Liverpool textile workers) were not in a position to exert much influence on the policies of their governments. Jefferson's noble experiment failed — and we soon fought the War of 1812. I am not suggesting that economic sanctions can never work. I prefer to exhaust peaceful options before resorting to armed force.

But it is critically important, especially when trying to influence the conduct of a radical regime like those in Iraq, Iran, Libya, North Korea, and perhaps a few other places, that the world community apply its sanctions against targets that are valued by the regime elites. Saddam Hussein did not *care* if he lost a hundred thousand soldiers in his campaign against Kuwait, and thus, threatening the lives of those soldiers did not succeed in deterring his aggression.

Ask yourself how he might have reacted in September 1990, if the Security Council had given him twenty-four hours to either cease his aggression or present himself before an international tribunal to answer charges as a war criminal.

And if that would not do it, what if they told him that — if he failed to either cease his aggression or to show up for trial — the Coalition Forces would be authorized to use necessary force to apprehend him for trial; and, if that proved impossible, to kill him.

LEGITIMATE SELF-DEFENSE

Now, keep in mind that international law recognizes a right to use lethal force — to take human life — when necessary as an act of "self-defense" or "collective self-defense" to armed international aggression. This principle of customary law was recognized in Article 51 of the U.N. Charter.[14]

Keep in mind also that the Nuremberg Principles — which were resolved to be part of customary international law by a unanimous vote of the U.N. General Assembly in 1949 — establish that high government officials who engage in war crimes are not immune from punishment.[15]

And finally, recall that the principle of "Proportionality" prohibits excessive acts of force beyond those necessary to achieve the permissible objective of ending the acts of aggression.

---

14. *See* U.N. CHARTER art. 51.

15. *See* Nuremberg Trial Proceedings, London Agreement, Aug. 8, 1945, *available at* www.yale.edu/lawweb/avalon/imt/proc/imtconst.htm (last visited Feb. 19, 2002).

This leads to a moral proposition: When forced to choose — that is to say, when peaceful responses to aggression clearly have not or will not be successful — is it better to respond to aggression by taking the lives of a few apparent "criminals" whose wrongful decisions have produced the wrongful act of aggression? Or is it better to give them the protected status accorded Red Cross and hospital workers while directing the defensive blow against hundreds of thousands of clearly more "innocent" soldiers?

I submit that — for both reasons of *morality* and *efficiency* — the time has long since come to start directing the sanctions against aggression against the Aggressors themselves, the governing elite in the Aggressor State, when the only alternative is sending our children out to slaughter the children of the Aggressor State.

OBVIOUS PROBLEMS

Time will not permit me to discuss all of the ramifications of such a proposal — and there are many obvious problems, such as "who decides" when it is permissible to target an alleged "aggressor" and whether such an approach will provoke terrorist attacks against leaders in law-abiding States.

But I would like in the few minutes remaining to address two issues that are important:

- What about the "Procedural Due Process" Rights of the Targeted Regime Elite?; and
- How does all of this fit into the prohibition against "assassination" contained in Executive Order and cherished by most Americans?

PROCEDURAL DUE PROCESS

To be sure, any fair system of justice will follow a process respectful of the rights of any accused individual. We do not allow policemen who witness murders to summarily "execute" the wrongdoer — even though such an approach might save hundreds of thousands of dollars in court costs and prison expenses. We understand that the cop might misunderstand what took place, and we cherish the right of all people to a fair trial by their peers.

But there are exceptions. Sometimes society has to choose between lesser evils. There may be no alternative to the loss of human life. The issue becomes: do we allow the apparent criminal to keep murdering hostages every five minutes or allow the SWAT Team sharpshooter to, as they say, "take the sucker out." Every civilized society allows its police to use deadly force against criminals in defense of the lives of the innocent.

And that, I submit, is the situation the world community faces when we

must choose between intentionally targeting a regime elite or sending our kids out to slaughter thousands of another country's soldiers.

To be sure, it is better to negotiate a settlement or to have a fair trial before anyone is punished. But I am talking about situations where those and other peaceful options simply do not exist. The choice is between the life or lives of a few international criminals who have launched a war of aggression, and the lives of the young men and women those elites have sent into the field to confront the world community.

Two moral principles seem important here:

• It is better to kill an obvious criminal than an apparently innocent person; and

• It is better to take few lives rather than many.

Both of these argue in favor of having the option of targeting regime elites.

ASSASSINATION ISSUE

Now let me say a quick word about assassination.

Put simply, by definition, the intentional killing of a wrongdoer as an act of necessary and proportional self-defense is not "assassination."

I have spent some time tracing the etymology and studying the modern meaning of the word "assassination," and it is clear that the term does not apply to the situation I am describing.

Black's Law Dictionary says:

> *Assassination.* Act of deliberately killing someone, esp. a public figure, usu. for hire or for political reasons.[16]

The 1980 *Oxford Companion to Law* adds similarly:

> *Assassination.* The *murder* of a person by lying in wait for him and then killing him, particularly the murder of prominent people from political motives, e.g., the assassination of President Kennedy.[17]

Insight into the word's meaning can be found from its etymology, as provided in Joseph T. Shipley's *Dictionary of Word Origins*:

> *Assassin.* This murderer of one in high place was originally one of a band (our word comes from the plural form) of *Assassins* (Arab. *hashshashin*, hashish-eaters) who were organized as a secret society in Persia, in the 11th c. Their leader was supposedly the Old Man of the Mountain, who sent them forth to destroy the Christian leaders. . . .[18]

Similarly, *Webster's Word Histories* traces the etymology of "assassin"

---

16. BLACK'S LAW DICTIONARY 109 (7th ed. 1999).

17. DAVID M. WALKER OXFORD COMPANION TO LAW 84 (1980).

18. JOSEFPH T. SHIPLEY DICTIONARY OF WORD ORIGINS 29 (1986) (emphasis added).

through a variety of languages, and concludes: "In all of these languages, the resulting term came to be applied to *murderers* irrespective of religion or Eastern connotations."[19]

In contrast to murder, what is being proposed here would be a lawful use of necessary and proportional lethal force in self-defense. The only difference is that the force would be directed against a few clear wrongdoers rather than many relatively innocent people.

It is both legally and morally justified precisely because the wrongful conduct of the target has left no other alternative but a greater loss of innocent lives. The war criminal is not killed for his political, religious, or other beliefs; he is killed to prevent his wrongful conduct from claiming a far greater number of human lives. This is a critically important distinction because the prohibition contained in Executive Order 12333 prohibiting "assassination" is commendable and ought not be placed in jeopardy by the proposition set forth herein.[20]

If the case for personally targeting Saddam Hussein during the Gulf War was a strong one, the case for going after Osama bin Laden to stop his ongoing campaign of lethal violence against Americans is even stronger. As a non-State actor, bin Laden is not even arguably protected by rules of sovereign immunity. As a terrorist, he is the common enemy of mankind and entitled to no greater protection than one would give a pirate.

CONCLUSION

To be sure, this approach might force the world to live with fewer Saddam Husseins, Adolf Hitlers, or Osama bin Ladens. But in my view, that is a price worth paying if it will help bring an end to unlawful violence in the world and promote lasting peace.

19. WEBSTER'S WORD HISTORIES 24 (1989) (emphasis added).
20. *See* Exec. Order No. 12,333, 46 Fed. Reg. 59941 (Dec. 4, 1981).

National Missile Defense: A Retreat from Dr. Strangelove or How I Learn to Stop Worrying and Love MAD

Dr. Willie Curtis*

The Cold War mentality that clings to the premise that the mere threat of retaliation is enough to deter countries possessing missiles capable of carrying nuclear, chemical, or biological warheads is strategically outdated. As Keith Payne argues, "[g]iven the proliferation of missiles and weapons of mass destruction (WMD), both the need to deny regional challengers their deterrent . . . , and the recognition that deterrence policies may fail, suggest important new roles for missile defense."[1]

The primary threat to the United States is no longer a calculated strategic nuclear attack by the Soviet Union. Among the new threats are accidental or unauthorized missile attacks by rogue states, terrorists, or irresponsible leaders. Indeed, clinging to the notion that Cold War instruments and policies can solve the security challenges of the twenty-first century is illogical.

In testimony before the Senate Committee, Henry Kissinger, former National Security Advisor, stated, "I cannot imagine what an American president would say to the American public if there should be an attack and if he would have to explain that he did nothing to prevent or defeat the resulting catastrophe."[2]

The acquisition of WMD by states of concern (rogue) or potentially other non-state actors have changed the nature of the strategic game. In the new strategic game, rogue challengers possessing WMD and delivery systems complicate U.S. deterrence goals.

The issue of National Missile Defense (NMD) must be placed in the proper context. The new post-Cold War strategic environment is different, therefore a new context and new perspectives must guide United States

---

* Associate Professor, Political Science Department, United States Naval Academy; Ph.D, University of Delaware.

1. KEITH B. PAYNE, DETERRENCE IN THE SECOND NUCLEAR AGE 143 (1996).

2. *See* Baker Spring, *Missile Defense: Back to Reality*, HERITAGE FOUNDATION VIEWS 2001 (Feb. 14, 2001) *at* http://www.heritage.org/views/2001/ed021401.html.

policy makers in developing a proper response to the new threats in a world of strategic multipolarity. "The term '*strategic multipolarity*' refers to an international system in which an increasing number of regional powers acquire nuclear weapons and other weapons of mass destruction."[3]

Critics argue that deployment of a National Missile Defense system would: (1) undermine the strategic logic of our deterrence policy; (2) negate the provisions of the Anti-Ballistic Missile (ABM) Treaty; and (3) generate an arms race with Russia and China. In this article I will advance the following arguments for deploying a National Missile Defense system. First, the current reliance on Mutual Assured Destruction (MAD) as a basis for deterring the emerging threats posed by development and spread of WMD and the means of delivery by ballistic missiles is based on illogical assumptions left over from the Cold War era. Second, the ABM Treaty perpetuated MAD, and while it was the best arrangement during the Cold War, its utility in the decade of the 21st century may be limited. Third, deployment of a Limited National Missile Defense system would enhance deterrence and provide a basis for further reductions in strategic nuclear weapons, while shifting the focus of our deterrence from MAD to an enhanced defensive oriented deterrence strategy.

MUTUAL ASSURED DESTRUCTION (MAD): A STRANGELOVEAN CONCEPT OF DETERRENCE

While Americans may take some comfort that the threat of instantaneous nuclear annihilation has receded, the fact remains that the international security environment in which policymakers seek to advance U.S. interests and provide national security can be characterized as quasi-anarchic at best, for the world is still a dangerous place. "While Russia and perhaps one or two other countries will retain . . . the ability to devastate the United States, Moscow's behavior will be conditioned by the same cold calculus of deterrence that kept the peace during the years of East-West confrontation."[4]

The cold calculus of deterrence of the Cold War decades reflected a bipolar bias for strategic deterrence that was conceived and evolved in an international environment dominated by two opposing ideological oriented superpowers. This calculus mandated a reliance on strategic forces that reinforced a strategic balance between the two nuclear armed adversaries,

---

3. Willie Curtis, *The Assured Vulnerability Paradigm: Can it Provide a Useful Basis for Deterrence in a World of Strategic Multi-Polarity*, 16 DEFENSE ANALYSIS 239 (2000).

4. ZALMAY ZHALILZAD & IAN O. LESSER, SOURCES OF CONFLICT IN THE 21ST CENTURY: REGIONAL FUTURES AND U.S. STRATEGY 18 (1998).

thus emphasizing mutual assured destruction; a rather Strangelovean concept of deterrence based on the capability of both superpowers to insure the destruction of the other's society.

Indeed, in this strategic environment, the two "'superpowers' condition closely resembles that of two scorpions in a bottle, each poised to strike, yet prevented from doing so by the knowledge that if either struck first, the other would respond and the demise of both were assured."[5] Today's strategic environment is characterized by a shift from strategic bipolarity to one of strategic multipolarity in which the strategic assumptions underpinning the deterrence strategy of the Cold War decades must be reexamined.

What is clear is that the bipolar international system that produced an international environment conducive to an aversion to risk-taking on the part of the two nuclear armed superpowers and the strategy of MAD will not suffice to provide a credible strategic deterrence in the twenty-first century world of strategic multipolarity.

The transformed international strategic landscape clearly suggests that the current reliance on MAD as a basis for deterring the emerging threats posed by the development and spread of WMD and the means of delivery by ballistic missiles is based on illogical assumptions left over from the Cold War era. The major assumptions underpinning this Strangelovean concept of deterrence include a policy described by Keith Payne as the Assured Vulnerability Paradigm.[6]

Payne explains that in the strategic bipolar environment of the Cold War "[t]he superpowers, calculating rationally and sensibly, would refrain from extreme provocation because of the ultimate possibility of nuclear retaliation."[7] Thus, the implicit assumptions of this Strangelovean approach to deterrence were based on (1) the capability of both superpowers to inflict unacceptable levels of destruction on each other and (2) that Soviet and U.S. leaders were rational decision makers willing to make policy decisions using dispassionate, rational, cost-benefit calculations.

Indeed, the absence of a nuclear war between the two superpowers

---

5. Curtis, *supra* note 3, at 239.

6. *See* PAYNE, *supra* note 1, at 70-78 (discussing the Assured Vulnerability Paradigm). Payne argues that

> [t]his theory, the Assured Vulnerability theory [MAD] of deterrence was popularized in the 1950s and 1960s, finding tremendous receptivity on Capitol Hill, among Academic Commentators, and think tank specialists, journalists, government officials, and eventually in the military. It became the prevailing paradigm ... [a]s applied during the Cold War, its basic precept was that threats of nuclear retaliation could provide a reliable basis for deterrence.

Id. at 60-61.

7. *Id.*

lends credibility to these assumptions; therefore, many strategic analysts and policymakers are convinced that the reliability and predictability of the MAD paradigm remains a desirable framework for strategic deterrence in the twenty-first century. However, Payne challenges this assumption and argues that, "[t]he Assured Vulnerability [MAD] paradigm assumes the character of the opponent and the context to be amenable to deterrence policies, and therefore it largely sets aside important contextual factors to focus on the character of the strategic balance . . . ."[8]

The question becomes, will this Strangelovean concept of MAD provide an effective deterrence, despite the dramatic changes that have and are still occurring in the contemporary international environment? If one proceeds from the premise that proliferation of weapons of mass destruction and ballistic missile technologies is inevitable in a world of strategic multipolarity, one must posit that several regional powers, with hostile intent to the interests and goals of the United States will become a part of the strategic landscape.

Within the context of strategic bipolarity, the logic of the principle assumption of the MAD paradigm of rational decision-making on the part of the two superpowers was valid, for in a bipolar world in which both the Soviet Union and the United States held similar views, experiences, and knowledge of the destructive potential of nuclear weapons, one might argue that they shared similar western perspectives on the use of force during the nuclear induced stability of the Cold War decades. However, can we infer that a similar condition will exist in a world of strategic multipolarity, with several regional nuclear powers having dissimilar cultural and decision making processes?

Michael O. Wheeler explains the implications of culture for rational decision making in his *White Paper on Deterrence.* He suggests that "[t]raditional deterrence theory assumes a human actor who can reason, communicate, anticipate and assess consequences, weigh them against interests and values, form intentions, assess risks, balance gains and losses, and act accordingly — an actor who knows his own mind, keeps emotions out of his decisions . . . ."[9]

In a world of strategic multipolarity, one essential task is to be aware of the impact of culture on perceptions of rationality from one culture to another. While this element of policy awareness was not a critical factor in deterrence calculations between the two superpowers during the Cold War decades, it is essential in a world of strategic multipolarity. Therefore,

---

8. *Id.* at 70-71.

9. Michael O. Wheeler, *White Paper on Deterrence, Weapons of Mass Destruction in Regional Equation: The New Deterrence Equation* 18-19 (SAIC, 1995).

"instead of thinking in terms of a general theory of deterrence," Wheeler argues, "we need to develop culture and society specific theories of deterrence."[10]

Another factor that must be given critical analysis is one of the risk-taking propensities of regional leaders armed with WMD and the means of delivery. In a world of strategic multipolarity, an understanding of the impact of culture on the propensity for regional leaders to take risk is critical.

In his study of risk-taking in decision-making, Yaacov Y.I. Vertzberger notes that

> [c]ulture plays a role in molding perceptions of risk and in defining risks as acceptable or not. 'Culture would seem to be the coding principle by which hazards are recognized. The cultural standards of what constitute appropriate and improper risks emerge as a part of this assignment of responsibility.'[11]

The assumption of rationality as a key factor in the MAD theory of deterrence is also questionable because it is clear that a leader's propensity for risk-taking is linked to culture. Understanding this will be a key factor in responding to the new strategic environment. Indeed, Vertzberger concludes that "[n]ot only is risk assessment culture-bound, but the preferred strategy of coping with risk sometimes depends on attitudes nested in cultural norms about the desirability of risk-acceptance and risk aversion."[12]

As an example, Vertzberger contends that "[i]n cultures that put a high premium on life, decisions involving risk to lives produce a risk-averse propensity, whereas in cultures that put a low premium on life, similar decisions do not preclude a risk-acceptant propensity."[13]

It is not the intent of this article to provide an extensive analysis of the role that risk-taking plays in deterrence. Further research is necessary to develop the data for understanding the many factors that played a minor role in the calculus of the MAD theory of deterrence during the Cold War. Given the new strategic environment, the assumption of a generically ra-

---

10. *Id.* at 37. Wheeler cites as an example: "It is worth recalling that Saddam Hussein's closest and most culturally similar neighbors didn't believe until the last minute that he was serious in the gulf crisis of 1990." *Id.* at 21. Indeed, Wheeler cautions that "what is rational behavior may be quite complicated to figure out, even within one's own culture much less across culture." *Id.*

11. YAACOV Y.I. VERTZBERGER, RISK TAKING AND DECISIONMAKING: FOREIGN MILITARY INTERVENTION DECISIONS 61 (1998)

12. *Id.* at 62. Vertzberger elaborates on the influence of culture in shaping the leader's perception of risk. He explains that "[b]y proving filters through which people look at the world, culture may affect the assessment of all three components of risk." *Id.* at 61. He cites as an example the fact that, "the value of human life varies across cultures." *Id.*

13. *Id.*

tional opponent model is not feasible. Therefore, consideration of the cultural implications for risk-taking is essential.

While an understanding of the motivations for the proliferation of nuclear and other weapons of mass destruction will depend on factors in addition to culture, it is important to remember the words of Albert Einstein: "the atomic bomb has changed everything except the nature of man."[14] Within the contexts of the new strategic environment, the assumptions of the MAD strategy are at best of questionable utility in deterring future regional leaders armed with nuclear and other weapons of mass destruction and the proliferation of ballistic missile technologies.

Whereas, the analogy of the two scorpions in a bottle depicted the situation in a world of strategic bipolarity during the decades of the Cold War, in a world of strategic multipolarity, the image may best be represented by two large scorpions surrounded by several smaller, yet just as deadly scorpions. [15]

MULTIPLE SCORPIONS IN THE BOTTLE: THE NEW CHALLENGE FOR DETERRENCE AND THE RELEVANCE OF THE ABM TREATY

Critics of the National Missile Defense (NMD) system also argue that deployment would undermine the Anti-Ballistic Missile (ABM) Treaty, which they see as the cornerstone of the nuclear relationship between the U.S. and Russia. They further insist that the ABM Treaty is critical to preserving strategic stability. The fact is that in the new strategic game, the ABM Treaty is rapidly becoming an anachronism that only permits the U.S. and Russia to mutually annihilate each other, thereby mutually deterring each other through the strategy of MAD.

Indeed, the ABM Treaty perpetuates MAD because it prevents both the U.S. and Russia from protecting themselves against nuclear retaliation. In the new strategic environment where proliferation of WMD and ballistic missile technology is spreading, it is questionable if the MAD strategy is appropriate, and thus whether the ABM Treaty, which was devised to cope with a bipolar, rather than multipolar world, is appropriate as well. As Henry Kissinger suggests:

Whatever, tenuous plausibility the MAD theory may have had in a two-power world evaporates when eight nations have tested nuclear weapons and many rogue regimes are working feverishly on development of nu-

---

14. MARGOT A. HENRIKSEN, DR. STRANGELOVE'S AMERICA: SOCIETY AND CULTURE IN THE ATOMIC AGE xvi. (1996).

15. Willie Curtis, *The Assured Vulnerability Paradigm: Can it Provide a Useful Basis for Deterrence in a World of Strategic Multi-Polarity, in* SEARCHING FOR NATIONAL SECURITY IN AN NBC WORLD 239 (James M. Smith ed., 2000).

clear, chemical, and biological weapons of mass destruction, as well as on the ballistic missiles with which to deliver them.[16]

He further argues,

> [t]he contrast between the security situation of 1972, when the ABM Treaty was signed, and today is stark. One signatory, the Soviet Union, has disappeared as a legal entity. Missile technologies have evolved in sophistication and proliferation [in] nations (North Korea, Iran, Iraq, India, Pakistan) not considered likely candidates for advance military technology when the agreement was concluded."[17]

Kissinger concludes, "[a]s for the argument that national missile defense runs counter to the long-standing strategic concept of Mutual Assured Destruction, a reassessment of that doctrine is long overdue, whatever the view regarding missile defense.[18]

While not a common practice in nation-state relations during the twentieth century, it remains legal under international law to withdraw from a treaty when it no longer serves the interests of the contracting parties.[19] The ABM Treaty is a bilateral treaty between the Soviet Union (now its successor state, Russia) and the United States. Its provisions provided a framework for a measure of security in the bipolar era of the Cold War decades.

However, the collapse of the Soviet Union, and the demise of the nuclear threat that characterized that period, suggest that the 1972 ABM Treaty is rapidly becoming an anachronism in the multipolar world of the twenty-first century. The fact is that one consequence of reliance on MAD and the ABM Treaty constraints is that "efforts of so-called rogue states to acquire long-range missiles are encouraged by banning significant missile defenses."[20]

While the ABM Treaty affected strategic relationships across the international system, altering or withdrawing from the Treaty has the potential

---

16. HENRY KISSINGER, DOES AMERICA NEED A FOREIGN POLICY? TOWARD A DIPLOMACY FOR THE 21ST CENTURY 66 (2001).

17. *Id.*

18. *Id.* at 65.

19. *See* David E. Sanger & Elisabeth Bumiller, *U.S. to Pull Out of Arms Treaty, Clearing the Path for Antimissile Tests*, N.Y. TIMES, Dec. 12, 2001, at A14. The argument may be a moot point because President Bush announced that the United States would exercise its options under Article XV of the Anti-Ballistic Missile Treaty. *See id.* Article XV provides that either party, with six months notice, may withdraw from the treaty "if it decides that extraordinary events related to the subject matter of [the] Treaty have jeopardized its supreme interests." *Id.*

20. JAMES J. WIRTZ & JEFFERY A. LARSEN, ROCKETS' RED GLARE: MISSILE DEFENSES AND THE FUTURE OF WORLD POLITICS 6 (2001).

not only of "replacing deterrence with offense -- defensive arms races, [but more importantly,] it also has the potential to initiate a transition away from deterrence based on mutual vulnerability to missile attacks and towards a global strategic defense regime on other cooperative national security policies."[21]

Finally, it defies strategic logic to continue reliance on MAD and the ABM Treaty to provide for deterrence in a world of strategic multipolarity. These bilateral agreements and perceptions of security needs were based on a paradigm whose assumptions are questionable in the new international security environment of the twenty-first century. Therefore, the critics of NMD should at least concede that while the Anti-Ballistic Missile Treaty that perpetuates MAD might have been the best arrangement during the bipolar decades of the Cold War, its utility in the present world of strategic multipolarity is questionable at best.

ESCAPING THE NUCLEAR SPECTRE: DEPLOYING A LIMITED NATIONAL MISSILE DEFENSE

The notion proposed by the rare heretic, that a proliferation of deterrents would magnify the influence of deterrence, was treated only as further proof that it takes all kinds to fill an asylum. But any real Strangelovean—one who loved the bomb in the purity of his heart—could only see proliferators as potential converts to the noble goal of demilitarized warfare in a disarmed world.[22]

In his book, *Dr. Strangelove and the Hideous Epoch: Deterrence in the Nuclear Age*,[23] John Renaker cites the case of Theodore B. Taylor, a nuclear scientist working at Los Alamos weapons laboratory in the 1950s. Taylor composed a letter home as a kind of apologia for his choice of profession:

> I think that there is only one realistic way to avoid war, and that is to make the world really afraid of it .... If A-bombs in their present form will make another war something which mankind cannot bear, and if most people don't realize this, then, I say, therefore is only one thing to do: develop a bomb which will leave no doubt in anyone's mind.[24]

MAD, as a condition and strategy sought to accomplish Taylor's vision of creating an aversion to war for it is based on the threat of assured destruction of the two superpower's societies. However, the question be-

---

21. *Id.* at 7.

22. JOHN RENAKER, DR. STRANGELOVE AND THE HIDEOUS EPOCH: DETERRENCE IN THE NUCLEAR AGE, 292 (2000).

23. *Id.*

24. *Id.* at 318.

comes: is this feasible in a world of strategic multipolarity, or is it unthinkable to assume that policymakers in the United States and Russia will grant that same status to a nuclear armed Iraq or Iran or other potentially hostile regimes?

What then are the requirements of deterrence in a world of strategic multipolarity? Should the United States continue to rely on MAD and the ABM Treaty to provide for strategic stability or should National Missile Defense become a part of the strategic equation? To answer these questions, let us revisit the scorpions in the bottle analogy. The requirements of deterrence have become increasingly more complicated and specifically the military requirements of deterrence under the Strangelovean concept of MAD "would require that all scorpions in the bottle maintain an ability to retaliate against all of the remaining scorpions, including the largest ones."[25]

In a world of strategic multipolarity with the inevitable proliferation of WMD and ballistic missile technologies, it is prudent for both the United States and Russia to move beyond reliance on MAD for deterrence. Leaders of regional states with small intercontinental ballistic missile forces may not be risk-aversive to the extent that they find credible a threat of nuclear retaliation and may be quite willing to risk massive destruction to obtain regional objectives and regime survival.

RETREATING FROM DR. STRANGELOVE OR HOW I LEARN TO STOP WORRYING AND LOVE MAD

In scene six of that 1960s classic film, *Dr. Strangelove: Or How I Learned to Stop Worrying and Love the Bomb*,[26] General Ripper explained to Group Captain Lionel Mandrake the answer to the question—"Do you recall what Clemenceau once said about war?"[27]

He said war was too important to be left to the generals. When he said that fifty years ago, he might have been right. Now war is too important to be left to the politicians. They have neither the time, the training nor the inclination for strategic thought.[28]

The fact that we are entering a world of strategic multipolarity where regional states are acquiring WMD and the means of delivery suggests that the inclination to think strategically about deterrence in the new stra-

---

25. Curtis, *supra* note 3, at 245.

26. DR. STRANGELOVE: OR HOW I LEARNED TO STOP WORRYING AND LOVE THE BOMB (Columbia Pictures 1963).

27. RENAKER, *supra* note 22, at 394.

28. *Id.*

tegic game is becoming essential despite the end of the Cold War. The new strategic game must not be assessed within the context and concepts of the Cold War paradigms and assumptions, for during the first decade of the 21st century nearly thirty countries have ballistic missiles. Moreover, fourteen of those countries produce and export those missiles to some of those regional states that are considered rogue or states of concern.

The new concepts of deterrence should include a multi-track approach that will address the new threats and consist of the following: (1) continued testing and eventual deployment of a Limited National Missile Defense System; and (2) shifting to a Multilateral rather than a bilateral treaty approach in limiting the size of the arsenals of the United States, Russia, Peoples Republic of China and potential regional nuclear states.

Continued testing and eventual deployment of a limited NMD system has the potential to enhance deterrence and provide a basis for further reductions in strategic nuclear weapons, while shifting the focus of our deterrence from a reliance on MAD to an enhanced defensive oriented deterrence strategy.

"Deployment of even a limited national missile defense system consisting of between 20 to 100 interceptors may discourage countries from acquiring ballistic missiles and help to persuade potential adversary that blackmailing or coercing the United States will not succeed."[29] A limited national missile defense system of this nature would not present a threat to Russia or China. As Wirtz argues:

> this deployment option poses no realistic threat to Russian or Chinese ability to strike the United States with warheads carried by intercontinental ballistic missiles (ICBM). This option would provide the United States with significant denial capability, however, against an emerging North Korean missile threat to American territory.[30]

Wirtz discusses three scenarios for consideration: Threshold Capability 1, Limited Defense in a Cooperative Setting; Capability 3 Plus, Enhanced Defenses and the Limits of Cooperation; and Unlimited Defenses, Unconstrained by Treaty.[31] The Threshold Capability Option "provide[s] the United States with a capability to protect itself against an accidental missile launch or very small, deliberate attacks that approach U.S. territory

---

29. WIRTZ & LARSEN, *supra* note 20, at 9.

30. *Id.* For an extended discussion of the various NMD deployment options, see Charles L. Glaser & Steve Fetter, *National Missile Defense and the Future of U.S. Nuclear Weapons Policy*, 26 INT'L SECURITY 40 (2001).

31. The Threshold Capability 1 of NMD "consider[s] a 'threshold' deployment of between 20 and 100 interceptors in a new base located in central Alaska .... The lowest level that the threshold system could engage is a maximum of somewhere between five and ten warheads flying towards the United States over the North Pacific Ocean." WIRTZ & LARSEN, *supra* note 20, at 9-10.

from the northwest."[32]

The deployment of this version of a limited missile defense system should be undertaken with an equal emphasis on diplomatic means to acquire multilateral agreements with Russia and the Peoples Republic of China (PRC) to limit their long-ranged strategic missiles to a level that is not threatened by a limited U.S. national missile defense system.

The first step of this multilateral diplomatic approach is for the United States and Russia to agree to further reductions of strategic missile forces to a level of 1,000 single warhead missiles. The second step is to engage the PRC in multilateral arms limitation discussions[33] and then seek further reductions once they are confident that they are not threatened by NMD deployment. This should provide each state with protection from potential threats from regional nuclear states in a world of strategic multipolarity (the multiple scorpions in the bottle dilemma). A third diplomatic approach would be to engage other regional nuclear states and potential states in strategic arms limitation talks to further reduce the size of their strategic nuclear forces.

These steps should be considered a start on the long path to shifting the security of the United States from a reliance on MAD. While it would be overly optimistic to assume that this will be accomplished in this decade, the fact that we are entering a world of strategic multipolarity -- regional powers will acquire WMD and the means of delivery -- suggests that prudence demands this two-track approach to providing security in the 21st century. By continuing the testing and eventual deployment of a limited NMD system, and starting the process of engaging Russia and the PRC in multilateral negotiations to limit their strategic nuclear weapon systems, diplomatic engagements will enhance the possibility of shifting from MAD to a defense oriented deterrence strategy.

In the meantime, given the nature of the strategic environment, the advice of that wise American, Will Rogers, should be heeded: "[d]iplomacy is the art of saying 'nice doggie' until you can find a rock."[34] In a world of strategic multipolarity, the rock may indeed be a limited national missile defense system.

---

32. *Id.*

33. Initially, in the case of the PRC, consideration must be given to increasing their strategic forces to 1000 single-warhead missiles.

34. TOM CLANCY, OP CENTER: ACTS OF WAR 149 (Berkeley Books, 1997).

National Missile Defense and the 1972 ABM Treaty

Robert F. Turner*

My mission on this panel is to discuss the legal implications of the 1972 ABM Treaty[1] to the proposed U.S. national ballistic-missile defense program. This involves some rather complex issues of both International and Constitutional Law, which are addressed at much greater length in my 1999 book, *The ABM Treaty and the Senate*.[2] The International Law of State Succession is both complex and unsettled. The view, however, with respect to non-dispositive treaties like the 1972 ABM Treaty is that such agreements would not continue in force upon the demise of one of the Parties — in this case, the Soviet Union — *unless* the United States and one or more of the emerging former Soviet Republics agreed to keep it in force. The process by which States agree to continue such an agreement is governed largely by their individual domestic laws — in our case by the U.S. Constitution.[3] On this point there is so much confusion that a brief technical digression at this point may be useful.

When the President was granted the new nation's "Executive Power" by Article II, section 1, of the Constitution, a key component of this power was the general control of American foreign relations.[4] As Thomas Jefferson put it in April of 1790, "[t]he transaction of business with foreign nations is executive altogether; it belongs, then, to the head of that department, except as to such portions of it as are specially submitted to the

---

* Professor, General Faculty, and Associate Director, Center for National Security Law, University of Virginia School of Law. B.A., 1968, Indiana University; J.D., 1981, S.J.D., 1996, University of Virginia. President, 1986-87, U.S. Institute of Peace; Principal Deputy Assistant Secretary for Legislative and Intergovernmental Affairs, 1984-85, U.S. Department of State; Special Assistant to the Under Secretary of Defense for Policy, 1981-82; National Security Advisor to U.S. Senator Robert P. Griffin, 1974-79.

1. Limitation of Anti-Ballistic Missile Systems, July 3, 1974, U.S.-U.S.S.R, 27 U.S.T. 1645 [hereinafter ABM Treaty].

2. ROBERT F. TURNER, THE ABM TREATY AND THE SENATE: ISSUES OF INTERNATIONAL AND CONSTITUTIONAL LAW (1999).

3. *See* U.S. CONST. ART. II § 2, cl. 2.

4. *See* U.S. CONST. ART. II § 1.

Senate. Exceptions are to be construed strictly . . . ."[5] President Washington endorsed this view, as did James Madison and Chief Justice John Jay. Madison had the previous year used identical reasoning to argue that the Senate's "negative" over diplomatic appointments—being an exception to the President's general grant of "executive Power"—did not extend to such matters as passing judgment on a presidential decision to dismiss an officer of the Government who had been appointed with the advice and consent of the Senate. This view was endorsed by both the House and Senate at the time.

Three years later, Alexander Hamilton relied upon the "Executive Power" clause[6] in defending Washington's Neutrality Proclamation.[7] He wrote:

> It deserves to be remarked, that as the participation of the Senate in the making of treaties [was an] exception out of the general 'Executive Power' vested in the President, [it was] to be construed strictly—and ought to be extended no further than is essential to [its] execution.[8]

While intended to be "narrowly construed," the checks or "negatives" given to the Senate and to Congress in this area were nevertheless important. The President may not make a treaty without the approval of two-thirds of the Senate,[9] and it is firmly established in U.S. practice that when an existing treaty is altered in any substantive manner it must be resubmitted to the Senate before it can be ratified by the President. I do not understand this point to be in serious dispute.

To mention but one early example I give in my book; following the War of 1812 a joint commission was established by the United States and Great Britain to determine the U.S.-Canada boundary as established by the 1783 treaty that ended the American Revolutionary War.[10] But the commissioners were simply unable to determine which wilderness body of water had been described as the "long lake" in the earlier agreement, so they did their best to survey a more precise boundary which both sides found acceptable and which probably did not alter the boundary of either country by more than a tiny fraction of one percent.[11] However, when the arrangement was presented to Secretary of State Henry Clay, he quite prop-

---

5. Thomas Jefferson, *Opinion on the Powers of the Senate* (April 24, 1790), *in* 5 THE WRITINGS OF THOMAS JEFFERSON 1788-1792 at 161 (Paul Leicester Ford. ed., 1895).

6. U.S. CONST. ART. II, § 1.

7. 15 PAPERS OF ALEXANDER HAMILTON, JUNE 1793 – JANUARY 1794 at 42 (Harold C. Syrett ed., 1969).

8. *Id.*

9. *See* U.S. CONST. ART. II, cl. 2.

10. *See* TURNER, *supra* note 2, at 162-63.

11. *See id.*

erly responded that the new boundary line was different from the one described in the text of the 1783 Treaty, and therefore it could not be implemented without first being submitted to the Senate for approval as a new treaty.

In my view, it is beyond question that any effort to continue the 1972 ABM Treaty with Russia, or with other former Soviet Republics, would unavoidably involve several substantive changes in the original terms — changes, which, as a matter of clear constitutional law, would require resubmission to the Senate for approval before the Treaty could enter into force.

Indeed, I sense that this was understood by all sides. On September 26, 1997, a new treaty — in the form of a "Memorandum of Understanding" (hereinafter MOU) — was signed in New York by the United States, Russia, Belarus, Kazakhstan, and Ukraine, designed expressly to continue the terms of the 1972 ABM Treaty in force as a new treaty once ratified by the signing States in accordance with their constitutional processes.[12]

Let me mention just three examples of the changes in the Treaty that require resubmission to the Senate:

> • To keep the terms of the 1972 Treaty in force with just Russia would exclude approximately 5.5 million square kilometers of territory that were covered by the original Treaty. That is almost exactly fifty times the area of the Commonwealth of Virginia, and enough territory to theoretically deploy more than a million ABM sites that would have been illegal under the original treaty. And had the Soviets insisted upon excluding even 100 square kilometers of their territory from coverage in 1972, neither President Nixon nor the Senate would have approved it.
>
> • The decision to multilateralize the ABM Treaty through the MOU involves still more substantive changes. In the Treaty ratified with the advice and consent of the Senate in 1972, the United States had half of the vote on the Standing Consultative Commission (hereinafter SCC), which administers the day-to-day operation of the Treaty. Under the MOU, three new countries are added, reducing our voting power from fifty percent to twenty percent and allowing four foreign countries to veto SCC

12. *See* Memorandum of Understanding Relating to the Treaty Between the United States of America and the Union of Soviet Socialist Republics on the limitation of Anti-Ballistic Missile Systems of May 26, 1972, *available at* http://www.state.gov/www/global/arms/factsheets/missile/abm_mov.html (September 26, 1997) [hereinafter MOU].

> actions. There is no precedent in American history for multilateralizing a bilateral treaty without the consent of the Senate.
>
> • Finally, to keep the Treaty in force without the consent of the Senate, one would have to conclude that the Parties were essentially the same. They might have a different name, but in all respects material to the Treaty, they would be the same. And the reality is that a critical consideration in the Senate's consent to the 1972 ratification of the ABM Treaty was that the Soviet Union was an aggressive Leninist dictatorship that might well find it in its interest to destroy the United States as we know it by the use of nuclear-armed ballistic missiles.

I do not believe the Senate would have given up America's right to defend itself against future ballistic-missile threats in response to a similar commitment from Great Britain, France, or Canada. Furthermore, the reality is that modern Russia is not the same country as the old Soviet Union — a point President Boris Yeltsin repeatedly emphasized.

Thus, I believe it is clear that the 1972 ABM Treaty has not been continued in force as a treaty between the United States and either Russia or the MOU signatories. The American President lacked the constitutional authority to do that, and I believe this was understood by all of the other countries involved.

However, this does not resolve our problem. The five signatories to the September 1997 MOU clearly understood the need for ratification of a new treaty, embodying the basic terms of the 1972 ABM Treaty with the necessary changes, and Article 9 of the MOU expressly provides that it shall not enter into force until ratified by the signatory states "in accordance with the constitutional procedures of those States."[13]

Presumably, because of the clearly expressed opposition of the Senate to the ABM Treaty, President Clinton did not submit the MOU for Senate consideration. It has not been ratified, and it is thus not binding upon the United States as conventional international law.

However, there is a "good faith" requirement under International Law — reflected in Article 18 of the Vienna Convention on the Law of Treaties[14] — which prohibits a State from defeating the object and purpose of a treaty it has signed until it gives notice to the other signatories that it does not intend to proceed to ratification. In my opinion, *this* is the primary legal constraint on the development of a national ballistic missile defense system.

---

13. MOU, *supra* note 12.

14. *See* Vienna Convention on the Law of Treaties, May 23, 1969, U.N. Doc. Conf. A/39/27 at art. 18.

The United States has always interpreted this obligation not as a duty to comply with every provision of a treaty, but merely not to take any irreversible steps contrary to the treaty.

Both the Russians and the United States — beginning with the Bush Administration in 1992 and reaffirmed repeatedly during the Clinton Administration — took the view that the ABM Treaty remains in force. These diplomatic statements did not have the effect of actually transforming the 1972 Treaty into an agreement between the United States and Russia (or any of the other MOU signatories) because to have done so would clearly have exceeded the President's authority under our Constitution.

Pursuant to Article 46 of the Vienna Convention on the Law of Treaties[15] — the United States can not be held to an agreement reached in manifest violation of a fundamental provision of its internal law regarding competence to make treaties.

The behavior of both the United States and Russia strongly suggests that neither side was trying to deny the U.S. Senate its proper constitutional role. They elected to implement their desire to keep the ABM Treaty in force by signing a new treaty (the 1997 MOU) which expressly provided for ratification in accordance with our constitutional procedures. Until that is done, though, their expressed desire to keep the Treaty in force has not been legally realized.

So I believe the better view is that the pronouncements of Presidents Bush, Clinton, and Yeltsin — and their diplomatic representatives — could not, and thus did not, have the effect of continuing the ABM Treaty in force. But they probably did create a moral obligation similar to Article 18's duty not to defeat the object and purpose of a signed treaty until notice is given of an intention not to proceed to ratification.[16]

Thus, from a purely legal standpoint, I would argue that if President Bush or Secretary of State Colin Powell this afternoon were to give formal diplomatic notice to the other MOU signatory states that the United States has elected not to ratify the MOU we would immediately be free of all legal constraints related to the 1972 ABM Treaty that might limit our development and deployment of an effective ballistic-missile defense system.

But let me emphasize the word *legal* in that statement. I submit that this is not merely a legal issue. In part because of the statements of our own leaders, most of our citizens and much of the world believes that the ABM Treaty remains in force. Even if my legal analysis is correct, and by giving notice that we will not proceed with ratification of the MOU, we would be technically free of all obligations under the 1972 Treaty, but there would

---

15. *See id.* at art. 46.
16. *See id.* at art. 18.

be serious political costs. If we do that, we will pay a high price both among our own people and people of good will around the world. It is not a price that we should willingly pay if there is an alternative—and there is one.

To avoid even the perception that the United States is violating its treaty commitments and considers itself "above the law," all the President would have to do is include in his notification to the other MOU signatories a six-month delay and an explanation of the "extraordinary events"[17] which have led us to conclude we must build a limited national ballistic-missile defense system.

Article 15 of the ABM Treaty expressly permits either side to withdraw on six month's notice "if it decides that extraordinary events related to the subject matter of this Treaty have jeopardized its supreme interests."[18]

While this test is subjective, I do not believe that anyone could read the Rumsfeld Report[19] and not agree that the United States faces extraordinary new threats to its security that were not anticipated when the 1972 Treaty was concluded.

I do not think it is necessary for us to take a position on whether the 1972 Treaty remains in force or has elapsed with the expiration of the Soviet Union. All we would have to do is to give notice to the MOU signatories saying something like: "consistent with the terms of Article 15 of the 1972 ABM Treaty, upon the expiration of a period of six months from this notice, unless a new agreement is reached regulating this matter, the United States will consider itself free to act as it deems necessary, consistent with its other obligations of international law, to protect its people and territory from limited ballistic-missile attack."

As a matter of international law, however, I do not believe that such a delay is essential. But because I believe it is extremely important for the United States at this juncture not to be perceived as a lawbreaker, were I advising the President, I would strongly urge the inclusion of such a delay.

With this approach, anyone who believes the ABM Treaty remains in

---

17. Article 15 of the ABM Treaty provides:

> Each Party shall, in exercising its national sovereignty, have the right to withdraw from this Treaty if it decides that extraordinary events related to the subject matter of this Treaty have jeopardized its supreme interests. It shall give notice of its decision to the other Party six months prior to withdrawal from the Treaty. Such notice shall include a statement of the extraordinary events the notifying Party regards as having jeopardized its supreme interests.

ABM Treaty, *supra* note 1, at art. 15.

18. *Id.*

19. DONALD RUMSFELD, REPORT TO CONGRESS, COMMISSION TO ASSESS THE BALLISTIC MISSILE THREAT TO THE UNITED STATES (July 15, 1998) *available at* http://fas.org/irp/threat/bm.threat.htm (visited February 22, 2002).

force will receive the six-month's notice under Article 15. Those who share my view would accept another six-months of delay in order to avoid even the appearance that we were acting contrary to our international legal obligations.

Let me conclude with just a couple of observations about Constitutional law. The decision to give diplomatic notice to the MOU signatories that the United States will not ratify the MOU is entirely executive in character, and thus is vested exclusively in the President of the United States by the Constitution. Neither Congress nor the Senate have any role in this process beyond the right of every American to provide the benefit of his thinking to the Chief Executive.

Similarly, the decision to give notice under Article 15 of the ABM Treaty is exclusively executive. This is a direct parallel to the 1789 debate on the role of the Senate in removing diplomatic appointees. Treaties and appointments are covered in Article II of the Constitution because they are Executive business, and the limited checks vested in the Senate were to be narrowly construed. This issue is, in my view, absolutely clear.

Luncheon Address

Rogue Regimes and the Individualization of International Law

Anne-Marie Slaughter*

In considering various specific legal and policy proposals in connection with rogue regimes, it may be helpful to step back and see them within a larger theoretical framework that is itself linked to broader trends in international law. The principal trend I wish to highlight in these remarks is the individualization of international law -- as an ideal and increasingly as a reality. I will describe the different dimensions of this trend and discuss two very broad implications for legal responses to September 11th and to the Taliban as a rogue regime. I will then turn to three more specific implications for legal responses to rogue regimes generally.

I. A VERY SHORT HISTORY OF INTERNATIONAL LAW

So what do I mean by the individualization of international law? It is alliterative, provocative, and enticingly vague. But how can international law be "individualized"?

I mean the process by which we have taken the black box of the state and made it gradually transparent to focus on individuals rather than states as unitary political entities.[1] That process has proceeded through a number of different areas of law over certainly the last half century and more likely the last century. My story is highly stylized, but the oversimplification does not undermine the basic point.

---

* J. Sinclair Armstrong Professor of International, Foreign, and Comparative Law, Harvard Law School.

1. Louis Henkin, among others, has captured this trend in a different but powerful formulation, the gradual shift from "state values" to "human values." *See generally* LOUIS HENKIN, AGE OF RIGHTS (1990); LOUIS HENKIN, THE CONSTITUTION AND RIGHTS (1989); *see also* Charlotte Ku, Global Governance and the Changing Face of International Law, Address at the Annual Meeting of the Academic Council on the United Nations System (June 16-18, 2001) *in* ACUNS REPORTS LECTURE, *available at* http://www.yale.edu/acuns/publications (last visited May 8, 2002).

A. From the Treaty of Westphalia to the Universal Declaration of Human Rights

Begin with the Westphalian cliché. Public international law regulates relations between independent sovereigns. Although we know better, or claim that we do, the Treaty of Westphalia, or our anachronistic interpretation of it, is invariably our starting point as teachers, as case book writers, and even as media pundits trying to explain the basics of international law to an impatient audience.

The Westphalian formulation of relations between states assumes that states are unitary political entities -- described either as black boxes or billiard balls. Indeed, the famous principle of *cuius principio, eijus religio* that is supposed to have permitted an end to the Thirty Years War essentially announces that what sovereign governments did within their own borders was of no concern to their neighbors. States were the subjects of international law; international law regulated only political and economic relations *between* states, not within them.

Exceptions abounded, of course. Blackstone cautioned that public international law required the provision of diplomatic immunity and safe-conducts, and prohibited piracy.[2] Nevertheless, the overall premise could be maintained without too much difficulty until the 20th century, with the advent of human rights law. By its nature, as a body of law regulating relations between a government and its citizens, human rights law rejects any notion of a state as a unitary entity. It is an entity composed of individuals and groups within society, on the one hand, and the government officials who may or may not represent them, on the other.

At first glance, human rights law imposes obligations on states, just as public international law does generally. However, those duties run to individuals, allowing them to claim direct protection under international law. Whether they thereby qualify as "subjects" or "objects" of international law is a debate more over semantics than substance. The point here is that individuals have independent status separate from but in relation to their governments.

B. From the Law of War to International Humanitarian Law

If human rights law breaks apart the black box of the state into governments and citizens, the rise of international humanitarian law, and its cousin international criminal law, takes a further step by focusing on individual government officials. The traditional law of war, *jus in bello*,

---

2. *See* 4 WILLIAM BLACKSTONE, COMMENTARIES ON THE LAWS OF ENGLAND 68-73 (1765); *see also* Anne-Marie Burley, *The Alien Tort Statute and the Judiciary Act of 1789: A Badge of Honor*, 83 AM. J. INT'L L. 461, 469 (1989).

largely addressed the conduct of legal combatants. It laid down the rules of the game for those who had a right to fight.

The move to international humanitarian law has been a move that redefines the purpose of the law of war in terms of the protection of humanity -- particularly the civilians who are evermore the victims of armed conflict of any kind. The threat to these individuals comes not from abstract "governments" making and implementing government policy, but from individual government officials committing crimes of enormous magnitude. These officials are liable for their actions as generals, as politicians, as policy makers -- whether before a war crimes tribunal or an international criminal court.

The advent of the International Criminal Court will create permanent individual liability for crimes against humanity, genocide, and grave breaches of the law of war. Government officials can no longer hide behind *raison d'etat*, shielded by the comfortable anonymity of collective policy-making. On the civil side as well, at least in the United States, plaintiffs have been able to hold individual government officials liable for human rights violations in tort. These individuals are being found liable for torture, crimes against humanity, and genocide., for war crimes.

C. From the Individualization of Governments to the Individualization of Societies

To recap, human rights law broke apart the concept of the unitary state and reconceived it in terms of a set of relations between governments and societies. International humanitarian law and international criminal law have displaced the image of "governments" as unitary entities, seeing them instead as collections of individual officials capable of bearing rights and committing wrongs.

We are now entering a third stage -- the individualization of society. We have seen society as the counterpart to government in the government-society equation. Societies are composed of citizens who are individually protected against torture, disappearance, and murder by their governments. But these citizens have been undifferentiated with regard to one another.

No longer. Society is becoming, in international law as in domestic law, a mass of individuals. To put it in its most simplistic form, the international law governing terrorism and the increased willingness to apply international criminal law to individuals who are not necessarily officials of recognized governments, such as Radovan Karadzic, the self-proclaimed President of the Republika Sprska, is beginning to distinguish between good individuals and bad individuals in domestic and international society. Prior to September 11th, the U.S. was all too inclined to see terrorism as a problem between a government and its dissidents, with an exception for terrorism in the Middle East. After September 11th, it has

become both a global crime and a global security threat. We are now sending armies after individuals.

In sum, international law has moved from seeing states as unitary entities, to recognizing and regulating the specific relations between governments and their citizens, to recognizing and regulating the actions of individual government officials, to recognizing and regulating the actions of individual citizens -- not only against their governments and other governments, but also against their fellow human beings. That is the progressive individualization of international law.

II. IMPLICATIONS OF THE INDIVIDUALIZATION OF INTERNATIONAL LAW

Before turning to the more specific implications of the individualization of international law, let me be clear: I do not mean for a minute to suggest that the state should be read out of the equation. State to state relations remain an enormous part of international law. Indeed, September 11th and its aftermath have largely exposed the myth that the state is somehow fading away, to be replaced by sub-national, transnational, and supranational nodes of power. It is the state that retains the monopoly on the *legitimate* use of force; it is to governments that frightened citizens still look first for protection. Indeed, in the U.S. the government has moved quickly to expand its powers in ways that are alarming to many citizens. And the political branches that might normally be expected to check such expansion now typically fall in behind the Executive, on the theory that war demands unity.

Individualized international law thus co-exists with the more traditional international law governing inter-state relations. It provides policymakers with additional legal tools at their disposal. At the same time, however, it has implications for reshaping policies that states have traditionally targeted at fellow states. I will begin by sketching two broad implications concerning how best to characterize terrorists -- as criminals or combatants -- and how to supplement existing rules against the inter-state use of force. I will then turn to a more specific set of implications for confronting rogue regimes.

A. Criminals or Combatants?

The first implication of this trend of individualization is that we should not be treating terrorists as combatants -- whether lawful or unlawful -- but as criminals. It means that individuals have responsibilities to their fellow citizens and their fellow human beings around the world. When they violate those obligations they are outlaws, under both national and international law. They are not soldiers, except in their own imagined ar-

mies.

The U.S. response to September 11th initially wavered between framing the attacks as terrorism or an act of war, between responding with the criminal justice system or the military. Following the campaign in Afghanistan, the war frame gained ascendance, with Secretary of Defense Donald Rumsfeld comparing the war against terrorism to a new Cold War.[3] War, however, still requires a state enemy, at least under the laws of war -- unless the U.S. is systematically engaging itself in other countries' civil wars. Further, prisoners in a "war" are prisoners of war, entitled to the full protections of the Geneva Conventions. By declaring the Taliban and Afghan detainees at Guantanamo Bay "unlawful combatants," the Bush Administration is denying their right to fight in the first place. But if they are not entitled to kill as soldiers, then their killing is a particularly heinous form of murder.

I do not mean to suggest that the use of military force has no place in combating terrorism. A government that offers open support for terrorists is violating its obligations under international law -- state to state. When it refuses to rectify that violation, force may be a lawful response, particularly when the terrorists in question have threatened a fellow nation's vital security interests. At some level, as with the attacks on the twin towers and the Pentagon, the scale and scope of a terrorist attack amounts to an "armed attack" under Article 51 of the UN Charter, entitling the victim state to use force in self-defense.

Over the longer term, however, the best strategy would bill itself not as a global war but as a global criminal justice campaign. From this perspective, the "war against terrorism" is an all-out fight against a particularly frightening and deadly form of global organized crime. Fighting more traditional organized crime poses many of the same difficulties we are now confronting with terrorists: the tension between securing convictions and jeopardizing informants, security risks, the difficulty of actually getting sufficient evidence to convict. But we have developed laws and procedures that make it possible to hunt down and prosecute drug lords, traffickers in women and children, illicit arms traders and money launderers -- all operating through global networks. We can fight global terrorist networks the same way, by consciously constructing a global criminal justice system.

A global criminal justice system requires courts. National courts around the world will be the linchpin of such a system -- the tribunals of first resort. Military tribunals may also be necessary where trials must be con-

---

3. *Terror Detainees* (National Public Radio broadcast, Jan. 21, 2002) *available at* http://search.npr.org/cf/cmn/cmnpd01fm.cfm?PrgDate=01%2F21%2F2002&PrgID=3 (last visited May 8, 2002).

ducted in a theater of military operations. Ideally, an international option would also exist, either as a backstop or an alternative to national courts. It could be a standing criminal tribunal or an ad hoc tribunal dedicated to trying terrorists. But it would symbolize global justice for global crimes.

This is not the place to debate the precise elements of a global criminal justice system. My point here is simply that the individualization of international law makes it much harder to put the "war on terrorism" into the framework of traditional war -- letting states and their armies fight a new enemy. The terrorists sought have obligations to all of us; we have a collective responsibility to bring them to justice.

B. Protecting Civilians As a Grundnorm of International Law

A second set of implications flowing from the individualization of international law concerns the need to supplement Art. 2(4) of the U.N. Charter. In 1945 the nations of the world came together after two horrific world wars to create a set of rules and institutions that would help guarantee their collective security. The foundation of the system they designed was a fundamental prohibition on the use of force against the political independence or territorial integrity of any state, except in self-defense or as authorized by the Security Council. If international law primarily regulates relations between states, the best international law can do to try to assure peace is to prohibit the use of force between them.

Today, however, individuals as well as states can threaten international peace and security. Individuals can now wreak damage on a scale and of the severity that was once the prerogative only of armies. And they can wreak such damage directly against the people of another state, without ever encountering a member of the national military.

Thus the legal question for states seeking to assure their collective security in the 21st century is how to fashion the equivalent of Art. 2(4) regarding individuals. My co-author, William Burke-White, and I have proposed an answer to that question: the adoption of a fundamental principle prohibiting the deliberate targeting or killing of civilians by any actors and for any purpose.[4] Regardless of the ends to be secured, by governments or non-state actors, the means used must respect the principle of civilian inviolability.

The principle itself already exists, of course. It has long been a cornerstone of the law of war, or international humanitarian law. It has played a central role in the jurisprudence of the International Criminal Tribunal for the former Yugoslavia. And it underlies many of the prohibitions on vari-

4. Anne-Marie Slaughter & William Burke-White, *An International Constitutional Moment*, 43 HARV. INT'L L. J. 1 (2002).

ous more specific conventions against terrorist acts, such as hijacking or bombing.

What is necessary today, as a direct corollary of the overall individualization of international law, is to raise the principle of civilian inviolability to constitutional status in international law -- one of the basic pillars of an international system designed to guarantee peace and security. We propose it as a *grundnorm*, equal to Art. 2(4) in prominence and weight. The term "civilian" itself must be redefined to mean not simply a "non-combatant" in an armed conflict, but rather a non-violent participant in civil society – an individual entitled to guarantees of life, liberty and dignity from both national and international law.

III. THE INDIVIDUALIZATION OF INTERNATIONAL LAW AND ROGUE REGIMES

So what has all this to do with rogue regimes? Interestingly, the very concept of "rogue regimes" reflects one step away from a purely statist view of international law. The standard term, not so long ago, used to be "rogue states." Here the problem of rogue behavior was attributed to an entire state, without differentiating even between a government and its people.

The move from "rogue state" to "rogue regime" is thus a shift consistent with the individualization of international law. A "regime" is not a permanent fixture; moreover, and more subtly, it has a somewhat less savory connotation than "government." But the trends outlined above would take the terminology a step further, attaching the appellation "rogue" to officials themselves. In addition, the individualization of international law has implications for the application of traditional doctrines of state responsibility. Finally, it strongly underlines the desirability of smart sanctions.

A. Ridding the World of Rogue Officials

The label "rogue regimes" is unhelpful. Indeed, it is counter-productive, precisely because it focuses attention on a regime or government as a collective entity and not on the individual officials responsible for systematic violation of international norms. From the perspective of the individualization of international law, it makes far more sense, as a matter both of rhetoric and of international legal language, to refer to rogue officials. The immediate counter-argument is that 100 rogue officials equal a rogue regime. True enough, but it is also possible that even within a government populated by 100 rogue officials, an additional number of officials harbor different views and would behave very differently if they could. It is important that international law be able to distinguish between these two categories.

A shift to the terminology of "rogue officials" does more than simply lose the alliteration. It also highlights the imprecision of the term "rogue," which may be accurate enough to describe the conduct of an entire regime, but is unhelpful, at least in legal terms, when applied to individuals. In the language either of the law of war or the criminal law we are concerned not with rogues but with criminals -- officials who have committed specific crimes.

It is worth noting here that this focus can be broader than crimes against other individuals. It may also come to encompass economic crimes, such as embezzling public funds. Some human rights activists have already coined the term "patrimonicide," which is simply a way of individualizing the official corruption that often lies at the heart of rogue regimes. An individual leader who treats his country as his personal treasure chest and his citizens as disposable cannon fodder belongs on a global "most wanted list."

B. Who's Harboring Whom?

Thinking about governments or regimes in terms of aggregations of individual officials who exercise power over differentiated populations of individual citizens also opens new perspectives on issues of state responsibility. If a government does in fact exercise control over its population, then it can be held responsible for activities taking place in its territory that injure other states. Conversely, however, if the officials in a particular government have in effect ceded control over their populations to an identifiable group of other individuals, then it is not clear that the officials are responsible for their "state's" actions.

This conjecture becomes concrete in the debate over whether the Taliban "harbored" al-Qaida or al-Qaida "hijacked" the Taliban. Both stories have been told by a range of expert observers. This is not the place to resolve the issue, but rather to explore how its resolution one way or the other could make a difference in legal terms -- at least *de lege ferenda*. If the Taliban did in fact govern within Afghanistan, then it is reasonable to hold them responsible for the acts of terrorists operating in their territory. On the other hand, if we were to conclude that the government had in effect ceded control over the political and military instruments of coercive state power to others, then those individuals should be held responsible for their own acts, not as state wrongs but as individual crimes.

More specifically still, the U.S. administration apparently concluded that the Taliban exercised sufficient control over the government of Afghanistan to hold them responsible for harboring al-Qaida and hence to use force against the Taliban government once it refused to surrender al-Qaida members implicated in the September 11th attacks. Had they concluded otherwise, at least within the framework of analysis I am propos-

ing, then our war should have been with al-Qaida alone, although it would have been impossible to pursue al-Qaida members without using military force on Afghan territory.

These may seem fine or highly formal distinctions. Certainly from the perspective of most Afghans, they are likely to be meaningless. But as international lawyers, whom we designate as the target or the enemy or simply the responsible entity has substantial implications in terms of the application of both *jus ad bellum* and *jus in bello*, not to mention the subsequent sorting out of civil and criminal liability. In a world in which corporations can be richer than countries, in which states can collapse, and in which individuals can threaten states, we need more flexible legal frameworks.

C. Sanctioning Rogues, Not Regimes

Finally, the individualization of international law has important implications for the question of sanctions. The legal differentiation first between state and society and then between governments and populations each composed of identifiable individuals should lead us to reject the imposition of sanctions that act indiscriminately on populations. If we mean to target government, we must design sanctions that will in fact affect the lives of individual government officials.

These are "smart sanctions," which are becoming increasingly attractive. Yet this endorsement of smart sanctions may also entail a condemnation of the far blunter instruments that governments all too often use -- sanctions that result in the death of millions of civilians even though the proximate cause is the intervening action of their own governments. Where such action is foreseeable, as it is almost bound to be when dealing with individuals as ruthless as those who typically succeed in leading "rogue regimes," the sanctions themselves may no longer be justifiable. Certainly that is the ultimate logical implication of formally differentiating between a government and its people, as well as among government officials and among individual citizens. We are no longer sanctioning "a state," but rather individuals within it. We can no longer harm the innocent in an effort to punish the guilty.

These propositions are far easier to assert than to implement. And policymakers must often still do the best they can with the instruments at their disposal. At present, the imperative is less to stop applying sanctions *per se* than to hasten the development of the most precise sanctions possible, carefully aimed at specific individuals. The devil, of course, is always in the details.

* * * *

Questions about specifics are always a valid response to a broad theoretical framework. Yet my purpose here has been deliberately general, to think about rogue regimes in a much larger context. In part, that context is a very changed security environment, or at least a security environment in which we are far more aware of significant changes than we were before September 11th. More broadly still, that context is the larger trends in international law that help provide the legal and political tools with which to respond to these changes.

The overarching trend I have identified is the individualization of international law, which has proceeded in three stages. Consistent with those stages, we should think of rogue regimes first in terms of their crimes against their own populations. Second, we should identify specific rogue officials responsible for those crimes. And third, where relevant, we should identify individuals within those populations that engage in global criminal activity, either on behalf of their governments or for themselves. Rogue regimes become rogue officials; rogue officials become criminals; and criminals, even in the eyes of international law, may be individual citizens.

Coercive Appeasement: The Flawed International Response to the Serbian Rogue Regime

Paul R. Williams*

Karina M. Waller**

INTRODUCTION

In April 1987, Slobodan Milosevic addressed a crowd of Kosovo Serbs outside the Kosovo parliamentary building who had gathered to protest the treatment of the Serb minority by the Kosovar Albanians. Milosevic proclaimed to the crowd that "[n]obody has the right to beat Serbs." With this simple phrase, Milosevic began a long campaign characterized by the use of ethno-nationalism and ethnic aggression to accomplish his objective of a mono-ethnic greater Serbia.

During the course of his war of ethnic aggression, Milosevic was predictably aided in his efforts by radical Serbian intellectuals, nationalist paramilitary organizations, the Yugoslav National Army (JNA), Croatian Serb and Bosnian Serb protégées such as Radovan Karadzic and Ratko Mladic, and a generally passive Serbian population.

More surprisingly, however, was that Milosevic was substantially aided by an American and European policy response of coercive appeasement. This response, born out of a diplomatic deficit and based in part upon the approach of myopic accommodation and moral duplicity, coupled with a attempt to minimize the use of force and the norm of justice, enabled Mil-

---

* Rebecca Grazier Professor of Law and International Relations, American University. Ph.D., Cambridge 1998; J.D., Stanford Law School 1990; B.A., UC Davis 1987. From 1991-1993 Professor Williams served as an Attorney-Advisor in the U.S. Department of State's Office of the Legal Advisor for European Affairs. Professor Williams has also served as the legal counsel to the Bosnian delegation during the Dayton Peace Negotiations, the Kosovar Albanian delegation during the Rambouillet/Paris negotiations, the Macedonian Albanian delegation during the Lake Ohrid/Skopje negotiations, and the Montenegrin Government during the Belgrade/Podgorica negotiations on the future of the Federal Republic of Yugoslavia.

** Research Fellow, Public International Law and Policy Group. B.A., Colorado State University 1997; J.D., American University's Washington College of Law 2001; M.A., American University's School of International Service 2002.

osevic and forces under his control to initially capture and ethnically cleanse thirty percent of Croatia and seventy percent of Bosnia. The reliance on the approach of coercive appeasement during the negotiation of the Dayton Accords secured for Milosevic de facto control of half of Bosnia, created institutional mechanisms that enabled him to retain the ethnic purity of his share of Bosnia, and enabled him to structure a politically strategic withdrawal from indefensible areas in Croatia. This success emboldened him to take subsequent action in Kosovo designed to expel up to two million Kosovars from the province.

But for an 11th hour change in American and European policy brought about by British Prime Minister Tony Blair, U.S. Secretary of State Madeleine Albright and NATO Commander Wesley Clark, Milosevic would not now be in The Hague, but likely would have succeeded in his plan to ethnically cleanse Kosovo, and potentially move on to cleanse Vojvodina and other non-Serb parts of Serbia.

The consequences of the approach of coercive appeasement for the people of the former Yugoslavia have been over 200,000 murdered, tens of thousands raped and nearly four million Bosniacs, Croats, Serbs and Kosovars displaced from their homes, many of whom will never be able to return. With a devastated infrastructure and a largely destroyed housing stock, the region has been plunged into an economic abyss in which it will remain for a number of decades to come.

The consequences for America and Europe are wide-ranging and include a costly diversion and degradation of U.S. and NATO military forces resulting in reduced preparedness to confront direct threats to national security, a steady erosion of the credibility of NATO as a deterrent force, the near institutionalization of dissension between America and its European allies, and a direct financial cost of nearly $10 billion associated with the deployment of NATO forces and the reconstruction of Bosnia and Kosovo. This in addition to the billions allocated in social aid for Yugoslav refugees.

The constant diplomatic focus on Yugoslavia for the past decade has diverted both diplomatic attention and scarce resources from other regions of strategic importance. Milosevic has set the tone of international behavior in the post cold war era by effectively hindering American and European efforts to craft a new paradigm of relations based on mutual security and shared political responsibility. Most importantly, the constant attention on the Yugoslav crisis by top level American and European diplomats and key foreign policy staff likely diverted attention and resources from a comprehensive assessment and reaction to the emerging threat posed by al Qaeda, the consequences of which were demonstrated in the September 11 attacks against the World Trade Center and the Pentagon.

The near constant ethnic conflict and economic sanctions have under-

mined efforts to stabilize conflicts in the region and efforts to promote a rapid economic transformation in the former Yugoslavia and in neighboring states such as Romania and Bulgaria.

Many of these consequences could have been avoided had the American and European foreign policy institutions not so dogmatically embraced an approach of coercive appeasement. The objective of this article is to explain how and why the Americans and Europeans, followed by the United Nations and other important international organizations, initially crafted and then employed the approach of coercive appeasement from the spring of 1991 through the winter of 1999.

DEFINING THE CONCEPT OF COERCIVE APPEASEMENT AND ITS ROLE IN PEACEBUILDING

Coercive appeasement begins when politically and militarily powerful third-party states or peacebuilders, such as the United States and the European Union, seek to resolve a conflict by accommodating the primary interests of a rogue regime despite the regime's use of force and commission of atrocities to achieve its objectives. Coercive appeasement more fully develops when third-party states employ very limited force, or pseudo force to relieve public and/or international pressure to take action to curb the behavior of the rogue regime. This limited force is then accompanied by actions which grant that regime many of its primary objectives. Finally, coercive appeasement often entails coercing the victim of the aggression into accepting arrangements that enable the rogue regime to accomplish its objectives. Thus, by its very nature, coercive appeasement encourages further aggression by signaling to the rogue regime that it will suffer few political consequences for its actions.

As the foreign policy of national governments or coalitions is seldom monolithic, it is important to acknowledge that at all times during the Yugoslav peace process a variety of approaches were utilized. These included coercive appeasement, the meaningful use of force, economic inducement, justice based approaches, and coercive diplomacy. Although coercive appeasement remained the dominant approach throughout, at times other approaches significantly influenced the development of the peace process. Importantly, the approach of coercive appeasement was often adapted in order to co-opt the institutions or mechanisms created to further other approaches, including the International Criminal Tribunal for Yugoslavia (Yugoslav Tribunal), United Nations Protection Force (UNPROFOR), and the post-Dayton Implementation Force (IFOR).

It should also be noted that the peacebuilders did not necessarily intentionally set out to implement a strategic approach of coercive appeasement. In fact all of the initial peacebuilding efforts began along the traditional path of accommodation coupled with moderate political and eco-

nomic sanctions and the deployment of peacekeepers. While this approach was the standard operating procedure of diplomats and highly appropriate for many conflicts, it was an utterly inappropriate means by which to confront and contain a Serbian nationalist regime determined to create an ethnically homogeneous greater Serbia.

In the face of constant Serbian aggression, the peacebuilders made tactical decisions and undertook ad hoc actions designed to achieve short-term objectives. These decisions and actions began to cumulatively frame and perpetuate what became over time an approach of coercive appeasement. Because many of the key policy makers were categorically committed to the approach of accommodation they often undertook actions designed to minimize the influence of other approaches such as the meaningful use of force, or justice based initiatives such as the Yugoslav Tribunal, which might have served to divert the approach of accommodation from its slide into coercive appeasement. By minimizing the influence of these other approaches, the sponsors of the approach of accommodation soon found that they had become the sponsors of the approach of coercive appeasement. It was not until after hundreds of thousands of civilians had been slaughtered and the credibility of the U.N. and NATO seriously undermined that those opposed to the approach of coercive appeasement were able to change the American and European approach to one of diplomacy backed by force. Under the new approach, Milosevic was defeated, the atrocities in the former Yugoslavia came to an end, and Milosevic was eventually delivered by the people of Serbia to the Hague to stand trial for his crimes against humanity.

There are five core elements that characterize the approach of coercive appeasement. The first core element is a *diplomatic deficit*, which entails a failure to understand the motivations and objectives of the parties to the conflict coupled with the failure to create the conditions for effective leadership and the articulation of a clear policy objective. This initial deficit is frequently augmented by the inability to structure a coordinated or capable diplomatic process for peace building. Often the diplomatic deficit encompasses the unintentional misuse of diplomatic signaling and the readily transparent articulation of intentions and available means by the peacebuilders. This diplomatic deficit also produces deficient institutional learning whereby peacebuilders are unable to adequately undergo institutional and personal "learning" during the peace building process.

The second core element may be termed *myopic accommodation*. This element entails the pursuit of actions designed to meet the needs and interests of the aggressor, and is often accompanied by the intentional or unintentional obfuscation of the aggressor's objectives.

The third core element, *moral duplicity*, involves declarations and actions designed to create the perception of moral equivalence among the

parties, thereby eroding the distinction between aggressor and victim and spreading culpability among all parties. These official pronouncements are often designed to actively erode the moral and strategic imperative to adopt approaches other than accommodation. Moral duplicity also entails the application of political, economic and sometimes military pressure on the victims in order to compel their acquiescence to the primary demands of the aggressor. Moral duplicity is often seen as a necessary behavior by those interested in pursuing accommodation or appeasement as it enables the peacebuilders to create divisions within the coalitions representing the victims and thus deflates their ability to prevent the peacebuilders from accommodating the interests of the aggressor.

The fourth core element is *constrained force* and entails those activities designed to constrain and minimize the use of legitimate force. Constrained force is often utilized as a means of coercion calculated to set the stage for further accommodation and appeasement once the aggressor has been brought to submission. Constrained force is often the result of an unwillingness on the part of third-party states or peacebuilders to articulate and implement strong policy initiatives which are not politically viable, or are inconsistent with the policy approach previously adopted.

The final core element is the *marginalization of justice*, which requires actions designed to minimize the role of justice, including the political resurrection of culpable partners in peace. While the role of justice is essential to bringing about a long-term peace, it is often deemed expendable by peacebuilders who utilize it as a bargaining chip in order to negotiate with erstwhile aggressors in the name of peace.

CRAFTING A POLICY RESPONSE TO THE YUGOSLAV CRISIS: SEARCHING FOR AN APPROACH

From its inception, the international community's effort to resolve the Yugoslav crisis incorporated many of the core elements of coercive appeasement - moral equivalence, erosion of the moral or strategic imperative to get involved, and the lack of a coordinated or able diplomatic approach. As the crisis in Yugoslavia unfolded, the British, French and German governments, working formally through the European Community institutions and through more traditional bilateral relationships, took the lead in crafting a response to the conflict. The initial approach relied on a formula designed to accommodate the interests of Milosevic and Croatian President Franjo Tudjman while trying to prevent the widespread commission of atrocities by deploying E.U. and U.N. monitors.

A. Planting the Seeds of Coercive Appeasement

Diplomatic Deficit

Many factors contributed to the diplomatic deficit that plagued the American and European approach to the conflict. At the time the Yugoslav crisis began, Europe experienced monumental changes with the Maastricht Treaty and the reunification of Germany as it sought to create a united European community.[1] This geopolitical occurrence created a sense of euro-nationalism, which led the European states to possess an over-inflated view of the power of Europe and their ability to prevent atrocities through persuasion.[2] Moreover, the European peacebuilders also failed to adequately understand Milosevic's intention of creating an ethnically homogeneous greater Serbia and thus began with the fundamental miscalculation that the crisis could be resolved with minimal effort in a matter of weeks.

This European position proved an ill fated mistake as the European Union's preoccupation with its internal operations left very little time and resources available for an effective strategy regarding the Yugoslav crisis.[3] According to one European scholar, Europe "sacrificed the interests of peace in the region to their own desire to forge a semblance of foreign policy competence."[4] This institutional egotism contributed to its shortsightedness, as evidenced by the E.U.'s continual misperception of

---

1. *See* Samantha Power, *The Reluctant Superpower*, *in* WITH NO PEACE TO KEEP: UN PEACEKEEPING AND THE WAR IN FORMER YUGOSLAVIA 148, 149 (Ben Cohen & George Stamkoski eds.,1999) (recounting the European Union's belief that it was the only power capable of restoring peace to Yugoslavia). Some of the member countries to the European Union believed that Yugoslavia would provide the needed incentive to create a common foreign and security policy. *See id.*

2. *See* Age Eknes, *The United Nations' Predicament in the Former Yugoslavia*, *in* THE UNITED NATIONS AND CIVIL WARS 109 (Thomas G. Weiss ed., 1995) (discussing the European Union's control of the Yugoslav situation during the early stages of the crisis).

3. *See* Philip Towle, *The British Debate About Intervention in European Conflicts*, *in* MILITARY INTERVENTION IN EUROPEAN CONFLICTS 94, 99 (Lawrence Freedman ed., 1994) (discussing Britain's internal debate over intervention in Yugoslavia); *see also* Lawrence Freedman, *Introduction*, *in* MILITARY INTERVENTION IN EUROPEAN CONFLICTS 1, 5 (Lawrence Freedman ed., 1994) (analyzing the dichotomy between the European Union's perceived role as the preeminent force in European conflicts and the reality of the European Union as a fledgling power). The European Union was unprepared for the task of intervening in a crisis that had the potential of escalating into an ethnic war because of its inability to create a security system capable of handling such conflicts. *See id.*

4. Jane M. O. Sharp, *Appeasement, Intervention and the Future of Europe*, *in* MILITARY INTERVENTION IN EUROPEAN CONFLICTS 49 (Lawrence Freedman ed., 1994).

the nature of the conflict, its adoption of a euro-centric view of the parties which assumed they would think and behave like their counterparts in the E.U., a historic bias toward the interests of Serbia, and simple anti-Muslim racism.[5] These factors led the E.U. to become entrenched in a policy of coercive appeasement.[6]

Despite the apparent failings of the European approach during the early stages of the evolving Yugoslav crisis, the United States made a strategic decision to defer to the Europeans while supporting the efforts of the peace process undertaken by the E.U. and eschewing the use of force. The United States' reluctance to become involved during the initial stages of the Yugoslav crisis was the product of its focus on the breakup of the Soviet Union, its concern over its increasing military engagements,[7] and the belief that the crisis could be resolved in a relatively quick fashion. As such, the United States felt that the United Nations and the European Union would project its sentiments that it would not allow the policies of ethnic cleansing to take place in Yugoslavia,[8] and were capable of bringing peace to the region through whatever means necessary.[9] This sentiment was consistent with United States' foreign policy that it would not continue to be the world's policeman. The United States was unwilling to "intervene and pay disproportionately to protect these principles when they seemed logically more salient in this case to her European allies."[10] While in theory, such a policy of allowing those closest to the conflict to organize and intervene seems the best use of limited resources, in reality the international community was ill-equipped to take on such a task, as evidenced by the European Union's repeated mishandling of the crisis.

To leverage their capabilities and to minimize Germany's influence,

---

5. For a concise review of early European political efforts to resolve the crisis, see generally Stanley Hoffmann, *Yugoslavia: Implications for Europe and for European Institutions*, *in* THE WORLD AND YUGOSLAVIA'S WARS 1 (Richard H. Ullman, ed., 1996); *see also* Sharp, *supra* note 4, at 34, 47 (discussing the less than competent response of the European Union to the Yugoslav crisis).

6. For a rare insight into the extent to which the approach of accommodation influenced by a perception of moral equivalence influenced the day-to-day operation of the E.U. Monitoring Mission and UNPROFOR, see generally BRENDAN O'SHEA, CRISIS AT BIHAC: BOSNIA'S BLOODY BATTLEFIELD 1 (1998).

7. *See* Power, *supra* note 1, at 149 (examining the reasoning behind the United States' reluctance to become involved in Yugoslavia).

8. *See* Thomas Halverson, *Disengagement by Stealth: The Emerging Gap Between America's Rhetoric and the Reality of Future European Conflicts*, *in* MILITARY INTERVENTION IN EUROPEAN CONFLICTS 76, 90 (Lawrence Freedman ed., 1994) (discussing America's involvement at the initial stages of the Yugoslav conflict).

9. *See id.* (observing the United States' position regarding involvement in the Yugoslav crisis by the United Nations and European Union).

10. *Id.* at 91.

France and the United Kingdom brought the crisis within the purview of the United Nations. The United States readily consented to this approach as it initially desired to minimize its role in the peace process, while retaining the ability to exercise control over developments which might affect its national interest. Given the initial passivity of the Americans, the British and French were able to structure an approach framed by an arms embargo, which prevented the weaker parties from using force to protect their citizens, the imposition of limited economic sanctions to prod the aggressors, the deployment of lightly armed U.N. peacekeepers who represented an international presence, but who were generally prohibited from reporting on the atrocities they witnessed and who were insufficiently armed to protect the civilians in their areas of operation, a mediation initiative designed to mollify the aggressors, and a dual key requirement of NATO and U.N. approval for the use of force. The British[11] and French[12] were also able to deploy key commanders on the ground with instructions to carry out the central elements of what became the approach of coercive appeasement. The Secretary-General of the United Nations further entrenched the approach of coercive appeasement by narrowing the Security Council dictated mandate to the U.N. peacekeepers such that they were able to use force only in instances of their own self defense – and not to protect civilians from atrocities and crimes of genocide. This regime was then used successfully to blunt subsequent American efforts to change the approach from coercive appeasement to the legitimate use of force.[13]

This diplomatic deficit at the beginning of the Yugoslav conflict created

---

11. *See generally* JANE M. O. SHARP, BANKRUPT IN THE BALKANS: BRITISH POLICY IN BOSNIA 1 (1993); JANE SHARP, HONEST BROKER OR PERFIDIOUS ALBION? – BRITISH POLICY IN FORMER YUGOSLAVIA 1 (1997); MARK ALMOND, EUROPE'S BACKYARD WAR: THE WAR IN THE BALKANS 1 (1994); BRENDAN SIMMS, UNFINEST HOUR: BRITAIN AND THE BOSNIAN WAR 1 (2001).

12. For a review of French policy, see Olivier Lepick, *French Perspectives*, *in* INTERNATIONAL PERSPECTIVES ON THE YUGOSLAV CONFLICT 76 (Alex Danchev & Thomas Halverson eds., 1996). *See generally* DANIEL VERNET & JEAN-MARC GONIN, GUERRE DANS LES BALKANS – LE MIROIR BRISÉ YOUGOSLAVE 1 (1994); MARC BENDA & FRANCOIS CRÉMIEUX, PARIS-BIHAC, LES TEMPS MODERNES 1 (1995) (discussing an account by two French soldiers of the passivity of the French forces in the face of Serbian defiance); ALAIN FINKEILDRAUT, THE CRIME OF BEING BORN 1 (Graham McMaster trans., 1997); Jolyon Howorth, *The Debate in France over Military Intervention in Europe*, *in* MILITARY INTERVENTION IN EUROPEAN CONFLICTS 106 (Lawrence Freedman ed., 1994) (reviewing the debate in France concerning intervention in Bosnia).

13. For a candid interview with the Secretary-General concerning his preference for a negotiated outcome and his aversion to the use of force, see Georgie Anne Geyer, *How the Conscience of the West Was Lost*, *in* THE CONCEIT OF INNOCENCE LOSING THE CONSCIENCE OF THE WEST IN THE WAR AGAINST BOSNIA 107-08 (Stjepan G. Meštovic ed., 1997).

conditions ripe for the adoption of coercive appeasement as a policy approach. Rather than formulating a cohesive, long-term strategy for the resolution of the situation in Yugoslavia, the international community, with the E.U. and the U.N. at the forefront, reacted to the growing conflict on an ad hoc basis, whereby decision making was promulgated from a reactive rather than proactive stance. This created an environment that lacked any meaningful dialogue among the members of the international community, culminating in an inability to structure a coordinated diplomatic process for peace building. The consequences of this ad hoc approach to resolving the conflict resulted in an unintentional misuse of diplomatic signaling which Milosevic used to his advantage.

The misuse of diplomatic signaling began in 1991 when then U.S. Secretary of State James Baker visited Belgrade and warned of the "dangers of disintegration," urged that Yugoslavia maintain "territorial integrity," and stated that "the United States would not recognize unilateral declarations of independence."[14] President Milosevic took this as a green light to use force to halt secession and to protect the Serbs living in Croatia and Slovenia. "What they read between the lines of the Baker visit," writes Ambassador Zimmermann, "was that the United States had no intention of stopping them by force."[15]

The effects of U.S. posturing were further exacerbated by the inconsistency by which the European Union dealt with Yugoslavia. By continually qualifying its statements and failing to project clear goals and a unified stance on the crisis, the E.U.'s actions may have contributed to the continued atrocities.[16] According to one European commentator, "[t]he uncertainties that western policies created led each of the contending Yugoslav parties to believe that its actions would not be punished, but rather, that after a brief interval of time, they would be accepted and perhaps supported by the western governments."[17] The European Union

---

14. LAURA SILBER & ALLAN LITTLE, YUGOSLAVIA: DEATH OF A NATION 150 (1996); Anthony Lewis, *War Crimes*, *in* THE BLACK BOOK OF BOSNIA: THE CONSEQUENCES OF BASIC APPEASEMENT 61 (Nader Mousavizadeh ed., 1996).

15. WARREN ZIMMERMANN, ORIGINS OF A CATASTROPHE: YUGOSLAVIA AND ITS DESTROYERS -- AMERICA'S LAST AMBASSADOR TELLS WHAT HAPPENED AND WHY 137 (1996).

16. *See* James Gow, *Nervous Bunnies: The International Community and the Yugoslav War of Dissolution, The Politics of Military Intervention in a Time of Change*, *in* MILITARY INTERVENTION IN EUROPEAN CONFLICTS 14, 22 (Lawrence Freedman ed., 1994) (stating that the actions of the European Union were ineffective because of its inability to portray a clear stance on the conflict).

17. Saadia Touval, *Lessons of Preventive Diplomacy in Yugoslavia*, *in* MANAGING GLOBAL CHAOS: SOURCES OF AND RESPONSES TO INTERNATIONAL CONFLICT 407 (Chester Crocker, et al. eds., 1996) (discussing the tools utilized by the European Union to keep Yugoslavia together); *see* BOGDAN DENITCH, ETHNIC NATIONALISM: THE TRAGIC DEATH OF YUGOSLAVIA 7 (1994) ("Thus, all those who support the creation of

also failed to portray a clear stance on the fate of Yugoslavia. Throughout the crisis, the E.U. continually qualified its statements, adding more confusion as to its ultimate position. It insisted that Yugoslavia remain a unified state, while at the same time sympathized with the plight of those wishing to secede. The conflict within the E.U. regarding recognition, especially with regards to Germany's and Denmark's vocal support for an independent Croatia and Slovenia, also created confusion among the Yugoslav actors.

As explained by European historian Mark Almond, Europe was rife with national divisions and "remain[ed] bedeviled by national rivalries despite all the talk of common policies and a new united identity" and by a reliance on outdated political and historical perspectives.[18] As a result,

> The breakdown of Yugoslavia brought out deep-seated Anglo-French suspicions of Germany which were partly shared and partly played on by the Serbian regime. During this time, the British and French politicians revealed that in their heart of hearts they could not see Germany as anything other than a domineering and aggressive nation, thereby trapping them with an inherited vision of 'plucky little Serbia,' the ancient ally against the Boche.[19]

Instead of projecting clear goals and a well-defined stance on the crisis, the often conflicting messages portrayed by the E.U. only strengthened the resolve of the Croat, Slovenian, and Serb governments to push towards their individual ends.

Myopic Accommodation

Throughout this period, the E.U. and the U.N. held a perpetual international summit in Geneva to seek a negotiated resolution of the conflict.[20] In Geneva, recurrent efforts by the E.U. and U.N. mediators to broker a lasting cease-fire and a framework for peace met with little success.[21] From 1991 to 1995 the U.N./E.U. peace conference proposed a series of peace plans for the de facto and de jure partition of Bosnia. Although occasionally a party would tentatively accept a plan as a tactical move to curry favor with the international community or to forestall international sanction, no proposed peace agreement was ever agreed to by the parties.

Throughout this time, the U.N./E.U. Co-Chairmen of the Peace

---

ethnic or national states, instead of states that will embrace all their citizens, support, consciously or unconsciously, policies of ethnic cleansing.").

18. ALMOND *supra* note 11, at xiii.

19. *Id.*

20. *See* SILBER & LITTLE, *supra* note 14, at 258.

21. On June 27, 1992, the European Community issued a declaration stating that while all parties were responsible for the continuing violence, the greatest share of responsibility for the crisis fell on the Serbian leadership and the JNA controlled by it. *See* Keesings, vol. 38, No. 6., p. 38943 (June 1992).

Conference squarely adopted the approach of accommodation, which quickly became one of appeasement. According to Ed Vulliamy, a journalist who followed the conflict for The Guardian,

> The history of appeasement of the Serbs is the history of the entire war. There were countless moments when the Serbs were told not to cross a line, and that they faced dire consequences if they did; and every time the bluff was called, the West climbed down, and the handshakes resumed. Some of the Serbs' more infamous 'last chances' may be briefly recalled: the fall of Jajce in 1992; the revelations of systematic mass rape in December; successive water- and bread-queue massacres in Sarajevo; the shelling of orphanages and hospitals; the first debacle at Srebrenica and the farcical establishment of the 'safe havens' that came out of it; the Serbs' pretense of accepting the Vance-Owen plan; two bloody crises in Bihac and two even bloodier ones in Gorazde.[22]

In response to growing public criticism, the E.U. Council of Foreign Ministers during the London Conference adopted the principle that it would not propose peace plans which ratified the gains of ethnic cleansing. The United States endorsed the same view.[23] Despite numerous similar public commitments, the Co-Chairmen of the peace process admittedly put forward proposals which legitimized territorial conquest involving ethnic cleansing and genocide. In fact, one of the Co-Chairmen, David Owen, expressly notes in his memoirs how he was irritated by the Dutch Foreign Minister's objection to a peace plan because it rewarded ethnic cleansing, but chose not to make an issue of it because it could be dismissed as purely moralistic and because he "personally liked" the Dutch representative.[24]

Despite the myopic focus on accommodating the interests of Slobodan Milosevic, the U.N./E.U. peace process did not produce peace, and in fact actually encouraged the parties to commit ethnic cleansing and continue the conflict.[25] The most notable example of this was when the U.N./E.U. proposed the Vance-Owen Peace Plan, which divided Bosnia into ethnic cantons, the Serb and Croat parties began to ethnically cleanse those cantons to which they had been assigned authority.[26] In fact, the U.N./E.U.

---

22. ED VULLIAMY, BOSNIA: THE CRIME OF APPEASEMENT 80 (1998).

23. *See generally* Assistant Secretary of State John R. Bolton, Remarks at the U.N. Human Rights Commission Session on the Situation in the Former Yugoslavia (Aug. 13, 1993), *available at* http://foia.state.gov/Documents/foiadocs/2af8.PDF.

24. DAVID OWEN, BALKAN ODYSSEY 236 (1995).

25. For a comprehensive assessment of the effect of these proposals, see generally LEE BRYANT, THE BETRAYAL OF BOSNIA (1993). For a review of the numerous proposals for the territorial division of Bosnia, see generally MLADEN KLEMEN, TERRITORIAL PROPOSALS FOR THE SETTLEMENT OF THE WAR IN BOSNIA-HERCEGOVINA (1994).

26. For a pointed critique of the Vance-Owen peace plan, see generally RENEO

accommodation approach permissively led to the commission of so many atrocities that even the U.N. Special Rapporteur for Human Rights Abuses in the Former Yugoslavia, Tadeusz Mazowiecke, felt obliged to resign in July 1995, charging that the failure of the U.N. and the international community to make any serious effort to stop atrocities in Bosnia made it impossible for him to continue.[27]

Moral Duplicity

While the Europeans relied upon moral duplicity as a means for justifying its accommodation of Serbian interests, the United States relied upon moral duplicity as a means for justifying its continued disengagement from the conflict. These justifications resulted in the adoption of a number of platitudes utilized by the Americans and Europeans to defend their actions, including statements that the war was caused by the bubbling over of "ancient ethnic hatreds," that all the parties were in effect "warring factions" equally responsible for the commission of atrocities, that the conflict was a "civil war" not involving Serbia, and that the Bosnian government was prone to killing its own civilians in order to garner international sympathy and intervention. In particular, E.U. negotiator David Owen readily adopted the notion of warring factions equally responsible for atrocities as it promoted his objective of a negotiated settlement of the conflict without the complicated involvement of the norm of justice. Similarly, Secretary of State Warren Christopher adopted Milosevic's notion of ancient ethnic hatreds along with the notion of warring factions[28] to create the

---

LUKIC, THE WARS OF SOUTH SLAVIC SUCCESSION YUGOSLAVIA 1991-1993 1 (1993).

27. *See* Balkan Action Council, *Balkan Watch Week in Review*, July 17-23, Vol. 2.28 *at* http://www.bosnet.org/archive/bosnet.w3archive/9507/msg00398.html.

28. *See, e.g.*, Secretary of State Warren Christopher, *New Steps Toward Conflict Resolution in the Former Yugoslavia* (Opening Statement at News Conference Feb. 10, 1993). Secretary Christopher proclaimed,

> [t]hose circumstances have deep roots. The death of [Yugoslav] President Tito and the end of communist domination of the former Yugoslavia raised the lid on the cauldron of ancient ethnic hatreds. This is a land where at least three religions and a half-dozen ethnic groups have vied across the centuries. It was the birthplace of World War I. It has long been a cradle of European conflict, [and] it remains so today.

Id. See also Remarks by Secretary of State Warren Christopher at the Conference on Security and Cooperation in Europe, (Plenary Session Nov. 30, 1993):

> [w]e call upon all warring parties to stop their unconscionable conduct that blocks the delivery of critically needed supplies through [Tuzla airport]. We also call upon the warring parties to live up to their recently signed agreements to permit secure land access for relief convoys. The warring parties must see that this is in their best interests. Full access will serve the vital needs of all Bosnia's factions.

Id. See also Presidential Exchange with Reporters, I PUB PAPERS 122 (Jan. 24, 1994) (adopting the notions of warring factions and civil war). President Clinton declared that "the killing is a function of a political fight between three factions. Until they

impression that the conflict was inevitable and that the American government could therefore not be faulted for failing to prevent the conflict or the continuing atrocities.[29] And the propensity for UNPROFOR Commander General Janvier to "believe Serb propaganda" according to his aides was in part responsible for his rejection of Close Air Support to defend the U.N.-declared safe area of Srebrenica, and the subsequent massacre of 7,000 civilians.[30] Not all those involved in seeking a resolution of the conflict fell victim to Milosevic's propaganda ploys, as illustrated by General Clark's assessment that "[a]bove all, I recognized that fundamentally, quarrels in the region were not really about age old religious differences but rather the result of many unscrupulous and manipulative leaders seeking their own power and wealth at the expense of ordinary people in their countries."[31]

Beginning in 1992, the United States publicly disavowed the fact that atrocities were taking place in Yugoslavia as a strategy for waylaying growing public concern and to relieve the moral imperative for US involvement. To further this end, President Bill Clinton permitted Secretary Christopher to embark on a public campaign of declaring that all sides were responsible for the atrocities and that the conflict did not amount to attempted genocide against the people of Bosnia.

An illustration of this approach can be found in Secretary Christopher's May 18, 1993 testimony before the House Foreign Affairs Committee. In response to pressing questions from Congressman Frank McCloskey, Christopher refused to acknowledge that the Serbian forces were committing genocide in Bosnia, asserting instead that "all sides" were responsible for the atrocities there -- thus removing the imperative for action.[32] In fact, according to the

---

agree to quit doing it, it's going to continue. I don't think that the international community has the capacity to stop people within the nation from their civil war until they decide to do it." *Id.*

29. *See* Geyer, *supra* note 13, at 91.

30. *See* DAVID ROHDE, ENDGAME THE BETRAYAL AND FALL OF SREBRENICA, EUROPE'S WORST MASSACRE SINCE WORLD WAR II 367 (1997). According to Mr. Rohde,

> [a] key element in Janvier's thinking was an apparent belief that he could do business with the Bosnian Serbs. Janvier may have turned down the crucial request for Close Air Support on the night before the town fell because he sincerely believed General Tolimir's promise the Serb attack had stopped. Janvier was quick to believe Serb propaganda and Mladic's complaints about Muslim provocations, according to aides. Janvier argued in the 9 June meeting in Split that the Serbs would no longer defy the U.N. if they were treated with respect.

Id.

31. WESLEY K. CLARK, WAGING MODERN WAR: BOSNIA, KOSOVO, AND THE FUTURE OF COMBAT 68 (2001).

32. *See* Editorial, *Foreign Policy Focus*, CHRISTIAN SCIENCE MONITOR, May 26, 1993, at 18; *see also* Saul Friedman, *Christopher Assailed Official: U.S. Downplayed Bosnia Genocide*, NEWSDAY, Feb. 4, 1994, at 4 (describing the continued obfuscation

former Yugoslav desk officer, Richard Johnson, the evening before his testimony, Secretary Christopher's office "sought urgent information from the [State Department's] Human Rights Bureau on Bosnian Muslim atrocities only," and during the testimony "insinuat[ed] that Bosnian Muslims [were] suspected of genocide themselves."[33] Presumably, in the mind of Christopher, if all parties were equally culpable then the Clinton administration would not be morally at fault for failing to take adequate action to stop the atrocities.[34]

Secretary Christopher's testimony was remarkable in that five months earlier, on January 11, 1993, a classified memorandum was prepared by the State Department's bureau of Intelligence and Research, the first sentence of which read, "[o]ver the past year Bosnian Serbs have engaged in a range of deliberate actions contributing to the attempted genocide of Bosnian Muslims."[35] The memorandum further stated that the Bureau of Intelligence and Research "believes there is substantial evidence indicating that Bosnian Serb efforts to eliminate Bosnian Muslim communities have been widespread, systematically planned, and ruthlessly implemented," and that "the results of well-organized genocidal activities are evident throughout Bosnia."[36] The memorandum also directly implicated the political leadership of Serbia in the planning and conducting of the campaign of attempted genocide.[37]

Secretary Christopher's testimony was all the more detrimental to the peace process as it lended important credibility to the war rhetoric of Radovan Karadzic. For instance, in March 1993, two months before his testimony, a classified Department of State cable from Belgrade argued that unconfirmed

---

of the genocide issue by the Department of State).

33. Richard Johnson, *The Pinstripe Approach to Genocide*, *in* THE CONCEIT OF INNOCENCE: LOSING THE CONSCIENCE OF THE WEST IN THE WAR AGAINST BOSNIA 72, n. 13 (Stjepan G. Mestovic ed., 1997).

34. *See* Geyer, *supra* note 13, at 89 (noting that the position of moral equivalence adopted by the British government served the position that "everybody's guilty, so we're not guilty for doing nothing").

35. Information Memorandum Regarding Yugoslavia, *Bosnia: Actions Contributing to Genocide*, DEP'T ST. DECLASSIFIED DOC. (Jan. 11, 1993), *at* http://www.foia.state.gov/Documents/foiadocs.PDF [hereinafter Information Memorandum]. For a well documented account of the early stages of the campaign of genocide in Bosnia and the origins of this policy in the political and academic institutions of Serbia, see NORMAN CIGAR, GENOCIDE IN BOSNIA: THE POLICY OF 'ETHNIC CLEANSING' 48-61 (1995). For more information on the role of religion and myth in the Serbian nationalist movement, see MICHAEL SELLS, THE BRIDGE BETRAYED: RELIGION AND GENOCIDE IN BOSNIA ch. 2 (1996).

36. Information Memorandum, *supra* note 35.

37. *See id.* Secretary Christopher subsequently undertook similar efforts to downplay the genocidal nature of crimes being committed in Rwanda despite assessments from the Office of the Legal Advisor and the Bureau of Intelligence and Research finding that genocide was in fact occurring in Rwanda. *See* Neil A. Lewis, *Papers Show U.S. Knew of Genocide in Rwanda*, N.Y. TIMES, Aug. 22, 2001, at A5.

accounts of atrocities against Serbs "have been widely publicized and embellished as part of a propaganda campaign in support of Serbian war aims in Eastern Bosnia,"[38] which were characterized by crimes of genocide. The cable then pointed out that Radovan Karadzic himself was using the "all sides are responsible" argument to justify the Serbian offensive against Srebrenica.

The element of moral duplicity was also embraced when the peacebuilders sought to politically legitimize those responsible for atrocities in order to secure their participation in the peace process. For instance, David Owen repeatedly legitimized Radovan Karadzic by embracing him as a legitimate partner in peace during the Geneva negotiations, and lauding him as a "gracious host," with "excellent English," while failing to mention his clear culpability for attempted genocide. It was only when Judge Richard Goldstone, the Yugoslav Prosecutor for the International Tribunal, indicted Karadzic for genocide that he was effectively delegitimized. Similarly, Ambassador Richard Holbrooke's now famous quote just before the negotiation of the Dayton Accords, in which he stated that "you can't make peace without President Milosevic,"[39] re-established Milosevic as a legitimate partner in peace despite his orchestration of genocide against non-Serbs. Slobodan Milosevic (subsequently indicted for genocide) was also characterized by Secretary Christopher as "though unscrupulous and suspected of war crimes, Milosevic has a rough charm and he appealed to some Western European leaders as a bulwark against an Islamic tide."[40] Most damaging was the legitimization of war criminals by those who were sent to protect civilians from atrocities. For instance, one British Army officer characterized General Mladic (also subsequently indicted for genocide) in the following terms, "he has presence, and when he had power he wielded it ruthlessly. That brought him some grudging respect, if not admiration."[41]

Constrained Force

After fighting broke out in Slovenia and Croatia in 1991, Belgrade requested that the Security Council impose an arms embargo on Yugoslavia

---

38. Information Memorandum Regarding Yugoslavia, *American Embassy in Belgrade Dispatch to Secretary of State*, DEP'T ST. DECLASSIFIED DOC. (Mar. 1993), *at* http://www.foia.state.gov/documents/foiadocs/54ec.PDF.

39. Jurek Martin, *Holbrooke Sees "Tough Slog" to Peace in Bosnia*, FIN. TIMES, Nov. 2, 1995, at 3.

40. WARREN CHRISTOPHER, IN THE STREAM OF HISTORY: SHAPING FOREIGN POLICY FOR A NEW ERA 352 (1998).

41. Unnamed British UNPROFOR officer, REUTERS, Nov. 10, 1996. *See* CLARK, *supra* note 31, at 40 ("[Mladic] carried a reputation among the U.N. forces for cunning and forcefulness, I found him coarse and boastful. He knows far less than he thought about NATO, airpower, and the capabilities of the United States.").

to prevent an escalation of the conflict.[42] Initially, such a measure was consistent with the dictates of international law due to the fact that Milosevic's brutality and the unknown nature of the conflict itself had yet to be realized. However, once the methods employed by Milosevic's forces were being executed against the Bosnian Muslim populace, the Security Council refused to reassess its earlier position and reaffirmed that its arms embargo would continue to apply to all parts of the former Yugoslavia, "any decisions on the question of the recognition of the independence of certain republics notwithstanding."[43] The only state truly affected by the arms embargo was Bosnia, which was left with no means to defend itself, while Serbia had all it needed in terms of military equipment and supplies. President Clinton had campaigned on a pledge to lift the arms embargo on Bosnia, but backed down under pressure from the United Kingdom and Russia.

Despite the failure of the Geneva talks to produce a peace agreement or even a lasting cease-fire, the negotiators did manage to obtain permission from the conflicting parties to send in UNPROFOR to secure the Sarajevo airport in order to open a humanitarian aid pipeline into the besieged city, an important concession given that the main thrust of early Security Council action in Bosnia was to provide humanitarian aid. UNPROFOR, however, proved unable to keep the pipeline open on a sustained basis due to frequent attacks on aircraft bringing in humanitarian aid.

In response to frequent Serb attacks on United Nations humanitarian aid convoys, on August 13, 1992, the Security Council adopted Resolution 770, which authorized governments to take "all measures necessary" to ensure the safe delivery of relief aid in Bosnia. This was the same formula contained in Resolution 678, authorizing the use of massive military force to expel Iraq from Kuwait. International expectations were high for a corresponding response in Bosnia. But, unlike Resolution 678, Resolution 770 led to no military intervention. There was no attempt to launch air strikes and no plan to send in coalition forces. As U.S. Secretary of State Lawrence Eagleburger explained two weeks after the adoption of the resolution, such action would not be stomached on either side of the Atlantic.[44]

The Security Council soon found itself faced with a new challenge when Bosnian Serb aircraft began to attack civilian targets.[45] The Bosnians, who had

---

42. *See* S.C. Res. 713, U.N. SCOR, 46th Sess., 3009th mtg., at 42-43, U.N. Doc. S/RES/TB (1991).

43. S.C. Res. 727, U.N. SCOR, 47th Sess., 3023 mtg., at 7, U.N. Doc. S/RES/727 (1992).

44. *See* John M. Goshko, *Eagleburger Debuts on the Balkan Crisis Acting Secretary Gives Major Credit for Pact*, WASH. POST, Aug. 29, 1992, at A15.

45. In 1992, the European Community issued a declaration stating that while all

no air force, were extremely vulnerable to such "ethnic cleansing by air" and the casualties quickly mounted. In response, on October 9, 1992, the Security Council adopted Resolution 781, which imposed a "no-fly zone" over Bosnia. At the urging of the British and French, the clause providing for enforcement of the no-fly zone was omitted from the Resolution. Instead, the Resolution called only for monitors to report on violations. During the next six months, there were over 465 documented violations of the no-fly zone.[46] Yet it was not until March 31, 1993, that the Security Council adopted Resolution 816, authorizing NATO to enforce the no-fly zone, and it was not until over a year later on February 8, 1994, that NATO would finally take action to shoot down Serb aircraft violating the ban.[47]

By far the most controversial of all of the acts of constrained force taken by the Security Council was the creation of and subsequent failure to defend so-called "safe areas" in response to the sustained Serb attacks on the Bosnian population centers at the beginning of 1993. The attacks in the eastern Bosnian town of Srebrenica were particularly ruthless in the spring of 1993 and the city was on the brink of collapse by the beginning of April.[48] On April 16, the Security Council adopted Resolution 819, which demanded that all parties treat the city as a "safe area" free from armed attack. A week later, the Council adopted Resolution 824, designating the predominantly Bosnian cities of Sarajevo, Tuzla, Zepa, Gorazde, and Bihac as additional safe areas. As a quid pro quo for the withdrawal of Serb forces, UNPROFOR was assigned the task of overseeing the demilitarization of the safe areas. Yet the Council provided no real enforcement component to the safe area concept. While the UNPROFOR commander indicated that it would take 35,000 troops to protect the safe areas, the Security Council authorized only a 7,500-troop reinforcement to accomplish the mandate. When the Serbs attacked the safe areas, the UNPROFOR forces retreated, and thousands of defenseless civilians were massacred and carted off to mass graves in the nearby countryside.[49]

---

parties were responsible for the crisis, the greatest share of responsibility for the crisis fell on the Serbian leadership and the JNA controlled by it. *See* Keesings, *supra* note 21.

46. *See* The United Nations and the Situation in the Former Yugoslavia, U.N. DEPT. OF PUBLIC INFO 13 (1993).

47. *See* OWEN, *supra* note 24, at 355.

48. *See generally* ROHDE, *supra* note 30; CHUCK SUDETIC, BLOOD AND VENGEANCE: ONE FAMILY'S STORY OF THE WAR IN BOSNIA 1 (1998) (providing compelling accounts of the Srebrenica massacre and the passive role of the international community).

49. *See generally* LAURENCE DE BARROS-DUCHÊNE, SREBRENICA: HISTOIRE D'UN CRIME INTERNATIONAL 1 (1996); JAN WILLEM HONIG & NORBERT BOTH, SREBRENICA: RECORD OF A WAR CRIME 1 (1996); BOB VAN LAERGOVEN, SREBRENICA: GETUIGEN VAN EEN MASSAMOORD 1 (1996); ROHDE, *supra* note 30; ERIC STOVER & GILLES PERESS, THE GRAVES: SREBRENICA AND VUKOVAR 1 (1998); Report to the Secretary General

"Historians will show," wrote the editors of *The New Republic* shortly after the Srebrenica massacre, "that the most important allies of the Bosnian Serbs have been the peacekeeping forces of the United Nations."[50]

From 1991-1995, the primary rationale for a preference of accommodation over the use of force was that if the international community employed air strikes against the Bosnian Serb forces, the Yugoslav Army would enter the war in support of the Bosnian Serb army. Although this view was publicly stated by the primary peace negotiators, a then classified Department of State cable unequivocally reported on February 17, 1994 that, in light of the February 9, 1994 NATO ultimatum threatening air strikes, "there is, so far, no indication that the Yugoslav Army is planning to initiate a general mobilization in Serbia either in anticipation of or reaction to NATO air strikes."[51] In fact, the cable reported that in light of the perceived seriousness of the threat and the perception of western resolve, the Bosnian Serbs would comply with the NATO demand after a period of brinkmanship.[52]

As confirmed by a complex time-series statistical analysis conducted by political scientists Joshua Goldstein and Jon Pevehouse, the threat or actual use of force by NATO produced compliance and cooperation by Serbian forces. The research further demonstrated that "Serb forces were more responsive to American actions than European ones."[53]According to the authors, "the results thus support the assumptions of the aggressor-victim school of thought that the international use of force could induce Serbian cooperation in this regional conflict."[54] Importantly, the authors also note, "[b]y contrast, the warring-factions school of thought, with its preferred policy of using international cooperation to elicit Serbian cooperation toward Bosnia, receives little support."[55]

Marginalization of Justice

During this time, the primary actors sought to marginalize any role for

---

pursuant to General Assembly resolution 53/35, the Fall of Srebrenica.

50. Lewis, *supra* note 14, at 175.

51. Information Memorandum Regarding Yugoslavia, *Belgrade Press Focused on UN Investigation of Sarajevo Massacre, Mladic*, DEP'T ST. DECLASSIFIED DOC. (Feb. 17, 1994), *at* http://www.foia.state.gov/documents/foiadocs/5557.pdf.

52. *See id.*

53. Joshua S. Goldstein & John C. Pevehouse, *Reciprocity, Bullying, and International Cooperation: Time-Series Analysis of the Bosnia Conflict*, 91 AM. POL. SCI. REV. 515, 527 (1997).

54. *Id.*

55. *Id.*

justice on the basis that it would inhibit their efforts to attain peace.[56] As explained by an anonymous U.N. official, the quest for "justice and retribution" is traditionally believed to hamper the search for peace, which in turn prolongs the conflict, enables the continuation of atrocities, and increases human suffering. The U.N. official also asserted that the intrusion of fact-finding missions seeking to investigate crimes committed by one side may complicate the task of peace negotiations to the point where they become prolonged or impossible.[57] According to Payam Akhvan of the Office of the Prosecutor of the Tribunal,

> From its very inception in 1993, the International Criminal Tribunal for the former Yugoslavia was surrounded by the so-called 'peace versus accountability' controversy . . . . It was argued that indicting political and military leaders such as Radovan Karadzic and Ratko Mladic would undermine the prospects of a peace settlement because they were indispensable to on-going negotiations, and because they would have no incentive to put an end to the fighting without assurances of immunity or amnesty.[58]

In fact, during his tenure as Co-Chairman of the Yugoslav Peace Conference, David Owen expressly opposed the prosecution of Serbian officials engaged in the peace negotiations on the basis that this would undermine his efforts to craft a settlement.[59]

Even after the massacres in Srebrenica and the clear pattern of genocide, policy makers doubted the compatibility of justice and accommodation. As noted by Richard Goldstone, the Yugoslav Prosecutor for the Tribunal, "[p]articularly at the time of the negotiations at Dayton, Ohio, in September 1995, there were many astute politicians and political commentators who suggested that, in fact, peace and justice were in opposition, and that the work of the Yugoslav Tribunal was retarding the peace process in the Balkans."[60] Some commentators even noted that because Radovan Karadzic's alleged approval rating among Bosnian Serbs was seventy-nine percent, any NATO efforts to capture him would undermine the im-

---

56. *See* Jean E. Manas, *The Impossible Trade-off: "Peace" versus "Justice" in Settling Yugoslavia's War*, *in* THE WORLD AND YUGOSLAVIA'S WARS 42 (Richard H. Ullman ed., 1996).

57. *See* Anonymous, *Human Rights in Peace Negotiations*, 18 HUM. RTS. Q. 249, 255-56 (1996).

58. Payam Akhvan, *Justice in the Hague, Peace in the Former Yugoslavia?*, 20 HUM. RTS. Q. 737, 738 (1998).

59. *See* Laurie A. Cohen, Comment, *Application of the Realist and Liberal Perspectives to the Implementation of War Crimes Trials: Case Studies of Nuremberg and Bosnia*, 2 UCLA J. INT'L L. & FOREIGN AFF. 113, 154 (citing Mirko Klarin, *The Moral Case for a War Crimes Tribunal*, WALL ST. J. EUROPE, Mar. 17, 1994, at 8).

60. Richard J. Goldstone, *Justice as a Tool for Peace-Making: Truth Commissions and International Criminal Tribunals*, 28 N.Y.U. J. INT'L L. & POL. 485, 488 (1996).

plementation of the Dayton Peace Accord and foster the Serbian people's belief that they were subject to perpetual injustice and persecution.[61] Goldstone rightly expressed surprise at this view, especially in light of the atrocities, which had been committed over four years, and the clear view of the Security Council when it established the Tribunal in May 1993.[62]

The initial decision to create the Yugoslav Tribunal was borne less out of a conscious decision to shift from the approach of appeasement to one of justice, but was in fact the result of increasing and intense public pressure to take some action to stop the atrocities in Croatia and Bosnia. By the middle of 1993, the Americans and Europeans had tried every other approach, including mediation, the deployment of UNPROFOR, economic inducements and economic sanctions as a means to de-escalate the conflict. As the peacebuilders remained unwilling to use substantial force, the creation of a Tribunal presented itself as the only remaining approach not taken.

The full story of the establishment of the Yugoslav War Crimes Commission in October 1992, and the Yugoslav War Crimes Tribunal seven months later, suggests that the members of the Security Council embraced the norm of justice mainly as a public relations device, while giving breathing room for the other approaches to peace building to succeed.[63] There was little expectation that the approach of accountability would succeed where the other approaches had failed. Indeed, some of the members of the Security Council saw the various approaches as fundamentally incompatible, and would work behind the scenes to undermine the Tribunal. Other members saw the Tribunal as useful for isolating offending leaders and marshaling support for more aggressive responses.

B. Consequences of Coercive Appeasement and a Temporary Course Correction

With the green light from the Americans and Europeans and the increasing support of the Russians, Milosevic and his forces were able to capture and ethnically cleanse 30% of Croatia and 70% of Bosnia early in the con-

---

61. *See* Charles G. Boyd, *Making Bosnia Work*, FOREIGN AFF. Jan.-Feb. 1998, at 50, 51.

62. *See* Goldstone, *supra* note 60, at 488; *see also* Floyd Abrams & Diane F. Orentlicher, *In Cambodia, as in Bosnia, Issue Is Punish or Pardon*, L.A. TIMES, Sept. 15, 1996, at 1M.

63. For a more detailed argument of the hypothesis that the Tribunal was established in part to relieve pressure to take more direct action to prevent atrocities, see generally ARYEH NEIER, WAR CRIMES: BRUTALITY, GENOCIDE, TERROR, AND THE STRUGGLE FOR JUSTICE (1998).

flict. By the time of the Dayton negotiations, less than 1% of Bosnia's pre-war Muslim population remained in the territory controlled by Milosevic's forces. To encourage the non-Serbs to flee Serbian controlled territory, Milosevic's forces established rape and torture camps, laid siege to and shelled towns and villages, and massacred large numbers of unarmed civilians.[64]

Milosevic's plans only became seriously threatened when his protégés, Radovan Karadzic and Ratko Mladic, failed to heed his instructions for moderation in the commission of atrocities and intensified their siege of Sarajevo and massacred over 7,000 unarmed civilians in Srebrenica in the summer of 1995. While these atrocities highlighted the continued failure of the European led efforts to quell the growing public perception of a weak and ineffectual international effort, they also presented an opportunity for greater U.S. involvement in the Yugoslav conflict. Given the heretofore ineffectiveness of the Europeans, NATO credibility was seriously eroded, forcing the United States to demand an increased role in the Yugoslav peace process. Congressional involvement in the foreign affairs prerogative of President Bill Clinton and increasing calls for action by the public represented a crucial juncture in the process and a turning point in the peace process. As a result of the massacres and active United States involvement, the international community was forced to take action in the form of limited air strikes. Milosevic's forces responded by seizing U.N. peacekeepers as hostages and chaining them to potential targets. After negotiating their release and redeploying UNPROFOR, the United States persuaded its allies to accept a more sustained bombardment of Bosnian Serb targets.

This action signaled to Milosevic the heightened willingness of the international community to resort to the use of force, compelling Milosevic to change his strategy. Ever cognizant of the fact that the actions of the

---

64. *See* CIGAR, *supra* note 35, at 56-57; *see generally* ZLATKO DIZDAREVIC, SARAJEVO - A WAR JOURNAL 1 (1993) (describing the siege of Sarajevo); JANINE DI GIOVANNI, THE QUICK AND THE DEAD: UNDER SIEGE IN SARAJEVO 1 (1994); ADRIANO SOFRI, LO SPECCHIO DI SARAJEVO 66 (1997). For additional information on the extent and nature of atrocities committed in the Yugoslav conflict, see generally VAHIDA DEMIROVIC, VISAGES FROM THE WASTELAND: A COLLECTION OF TRUE WAR STORIES FROM BOSNIA 1 (1999); LAURENCE DE BARROS-DUCHÊNE, SREBRENICA: HISTOIRE D'UN CRIME INTERNATIONAL 1 (1996); ILSE BAUMGARTNER & WOLFGANG BAUMGARTNER, DER BALKAN-KRIEG DER 90ER 1 (1997); CLAIRE BOULANGER, ET AL., L'ENFER YOUGOSLAVE: LES VICTIMES DE LA GUERRE TÉMOIGNENT 1 (1994); THE GOLDEN APPLE: WAR AND DEMOCRACY IN CROATIA AND BOSNIA 1 (Nick Ceh & Jeff Harder eds., 1996); SMAIL C KIC, THE AGGRESSION ON BOSNIA AND GENOCIDE AGAINST BOSNIAKS: 1991-1993 1 (1995); FRANCOIS CHASLIN, UNE HAINE MONUMENTALE: ESSAI SUR LA DESTRUCTION DES VILLES EN EX-YOUGOSLAVIE1 (1997); BOŽICA ERCEGOVAC-JAMBROVIC, GENOCIDE: ETHNIC CLEANSING IN NORTH-WESTERN BOSNIA 1 (1993).

Bosnian Serb army was jeopardizing his reputation with the West[65] and threatening his tenuous hold on power, Milosevic abandoned his Greater Serbia idea[66] and began to distance himself from Karadzic and Mladic in the hopes of regaining favor with the international community.[67] This demonstrated willingness to use force, coupled with a Bosnian-Croat offensive that had freed the Krajina region of Croatia and approximately twenty percent of Bosnia from Serb control (and which was poised to retake most of the western Bosnia territory held by the Serbian forces) provided the turning point in the Yugoslav conflict, with Milosevic realizing the necessity for consolidating his gains through negotiations.[68] This shift-

---

65. Information Memorandum Regarding Yugoslavia, *Codel Hall Tells Pale Leadership It Should Improve Humanitarian Access*, DEP'T ST. DECLASSIFIED DOC. (July 31, 1995), *at* http://foia.state.gov/Documents/foiadocs/2af8.PDF [hereinafter Declassified Documents, Hall Memorandum]. The declassified document noted that "[Congressman] Hall urged Jovanovic to 'show the world that the Serbs had hearts and were not monsters.'" *Id.* He went on to state that "Serbian President Milosevic could play a major role in changing the brutal behavior of the Bosnian Serbs." *Id.*

66. *See* Mihailo Markovic Interview, *NIN*, Feb. 24, 1995. Mr. Markovic observed that

> [o]ne Serbian perspective (Milosevic's) thus clashed with another (Karadzic's), while a third (the Krajina Serbs) went along with Milosevic down the road towards a progressive normalization of relations with Croatia. (A comprehensive cease-fire between the Krajina Serbs and Croatia was agreed in March 1994, followed by the re-establishment of certain economic relations in December 1994.) The interface between the three in all directions is complex and makes the search for a solution extremely complicated. To simplify, Milosevic and others originally launched/provoked the 'trans-Drina' Serbs into their 'independence adventure'. Since then Milosevic has attempted to reverse the process, abandoning the project of a 'Greater Serbia' – which was always more of an instrumental goal, to help him to retain power, though many of his nationalist confederates may have seen it differently.

Id.

67. *See* Declassified Documents, Hall Memorandum, *supra* note 65. The document stated that "[name withheld] told us that there was a lack of command and control at every level in the Serbian/Bosnian Serbian relationship. [Name withheld] was convinced that Milosevic wanted desperately to avoid direct conflict with NATO and was twisting Karadzic's arm in order to force compliance with the February 9 decision. Milosevic was pressuring the Bosnian Serbs in 'every way he could' to withdraw their heavy weapons or place them under UN control." *Id.* "A journalist for the semi-independent Belgrade daily Borba told us that Milosevic's efforts to project an image of normality and create distance between himself and the Bosnian Serbs, has succeeded in keeping most Belgraders relatively blasé about the possibility of air strikes on Bosnian Serb positions." *Id.*

68. *See* International Crisis Group, *Serbia - The Milosevic Factor*, *at* http://www.crisisweb.org/projects/sbalkans/reports/yu01rep.htm (last visited Feb. 23, 2002) ("Milosevic has radically altered his professed beliefs and loyalties and abandoned his strategies on several occasions. Once again, the common denominator lying

ing of alliances would guarantee the international community's compliance with his demands throughout the Dayton negotiations, thereby ensuring his continued place as the leader of the now defunct Yugoslavia.[69]

Another turning point in the conflict was the Bosnian government's absolute rejection of continued European leadership. This rejection enabled the United States to assume the leadership role in the Yugoslav peace process, and to initiate a new round of mediation. Importantly, this change in leadership provided a momentary opportunity for the international community to discard the failed coercive appeasement approach and to adopt a more decisive and coherent strategy for resolving the conflict.

Several factors contributed to the decline of coercive appeasement and the reinvigorated international effort in the summer of 1995. These include: the July massacre of over 7,000 civilians in Srebrenica; the displacement of the European architects of coercive appeasement by the U.S., which assumed the mantle of chief negotiator; the demonstrated effectiveness of air strikes after the August Sarajevo market place massacre, which confirmed the applicability and utility of the use of force; the success of the Bosnian-Croat offensive, which validated the approach of lift and strike; and the growing dependence of Russia on international political and financial assistance, which moderated its negative influence on the peace process.

Unfortunately, rather than capitalizing on Milosevic's weakened position and its new found leadership role by employing the use of force to defeat Milosevic and his protégées, the United States opted to continue the approach of coercive appeasement in the hopes of reaching a quick settlement and pressured all parties to commit to the Dayton negotiations. Had the former approach been adopted, it is likely the gains of ethnic cleansing in Bosnia could have been reversed, the war in Kosovo avoided, and possibly the democratic transformation in Serbia accelerated. What ensued, however, was a U.S. peace initiative encompassing a summer of shuttle diplomacy and preliminary negotiations in Geneva and New York, culminating in a three-week negotiation at the Wright-Patterson Air Base in Dayton, Ohio. To create parity among the parties, the U.S. actively undermined the effectiveness of the Bosnian-Croat offensive by telling each

---

beneath each of these shifts has been a hunger for power or an acute fear of losing it.").

69. *See id.* (noting that "[a]fter the international community accepted Milosevic as a 'peacemaker' at the Dayton talks, the Serbian leader immediately began exploiting his new-found international prestige to shore up his political position at home"); *see also* Press Conference, Special Representative Robert S. Gelbard, U.S. Department of State (Feb. 23, 1998) *available at* http://www.state.gov/www/policyremarks-/1998/980223_gelbard.html (holding Milosevic out as a partner in peace). "Milosevic played such an important role in reaching the Dayton Agreement . . . we have been particularly encouraged by the support we received from . . . Milosevic." *Id.*

party they could not trust the other, and by intentionally misrepresenting to the Bosnian and Croatian governments that the Bosnian Serbs were preparing a counter-offensive which would likely roll back their recent gains.

THE DAYTON NEGOTIATIONS: SAVING FACE WHILE INSTITUTIONALIZING COERCIVE APPEASEMENT

In November 1995, the parties to the Yugoslav conflict, including Milosevic, President Franjo Tudjman of Croatia, and President Alija Izetbegovic of Bosnia, and representatives of the Contact Group (which consisted of the United States, France, Germany, Russia and the United Kingdom) gathered in Dayton, Ohio to negotiate a series of accords designed to bring about the end to the war and the peaceful re-integration of Bosnia. Given the calculated effort to legitimize Milosevic as a reliable partner in peace, seasoned journalist David Rieff observed that the only issue for Dayton negotiators was to determine how unjust the agreement between those who instigated the genocide and those who were its intended victims would turn out to be.[70]

A. Changing Horses, but Continuing on the Path of Coercive Appeasement

Despite the success of the air campaign and the successful Croat/Bosnian ground offensive, the United States, like the Europeans before them, calculated that it was necessary to "get to yes" with war criminals, and that the only means for doing so would be to re-deploy the approach of coercive appeasement. Thus, during the Dayton negotiations, the U.S., accompanied by its European allies, continued to actively pursue the approach of coercive appeasement, and relied upon many of the core elements that had played a central role in the earlier efforts to resolve the Yugoslav crisis with disastrous consequences.

The ultimate goal of any peace process is the attainment of a durable, lasting peace. To achieve this, it is usually necessary for the negotiators to bargain with the parties to the dispute in order to reach a compromise settlement. However, this type of negotiation in situations where one party is clearly the aggressor and the other is a devastated victim often lends itself to the approach of coercive appeasement.

In fact, some commentators might even contend that negotiations by their very nature undermine the real chances for peace. As noted by Helen Fein, "whereas intervention is required against acts already deemed crimi-

---

70. David Rieff, *Reconciliation at a High Price*, N.Y. TIMES, Nov. 23, 1995, at A27

nal in international law, 'conflict-resolution or mediation model[s]' may only be appropriate before the commission of crimes against humanity and sends the 'wrong message when invoked during genocide, overlooking or rewarding crime and aggression.'"[71] According to Fein, "if we view genocide as state crime rather than 'ethnic conflict,' we are constrained to prevent, deter and stop it, rather than to negotiate in the middle with its perpetrators."[72]

Herein lies the conundrum of peace building faced by the Dayton negotiators: If you exclude those responsible for war crimes, then you are unlikely to secure a negotiated peace. If you include them in the process, you legitimize their actions as well as their agenda, and likely increase the possibility of continued atrocities or the ratification of a fundamentally flawed peace agreement that encourages additional ethnic aggression. The solution to this conundrum would have been to build the political will to continue with the use of force against those responsible for war crimes and to limit contact with these individuals to the extent necessary to secure the implementation of dictated terms of peace. As evidenced by the Dayton peace initiative, this approach was not taken.

Diplomatic Deficit

The American delegation assembled at Dayton was of the highest caliber and thus did not suffer from a deficit in intellectual capacity or experience. Despite the ability of the American delegation, the circumstances and decisions made early in the planning stages did, however, result in the re-emergence of the diplomatic deficit that plagued earlier European efforts.

From the beginning, the tension between those in support of a peace deal and those in support of a peace package laid the foundation for re-establishing the approach of coercive appeasement. Those in favor of the peace deal, most notably the American and French delegations, clashed with the British and European Union delegations who sought to appease various parties, who in turn clashed with the Bosnian and German delegations who sought a peace package. This peace package consisted of a structured set of proposed commitments that, if fulfilled, could promote the peaceful reintegration of Bosnia.

Some members of the Contact Group recognized that to secure a just peace, it would have been necessary to reverse the gains of ethnic cleansing and attempted genocide, and to weave throughout the agreement vari-

---

71. Helen Fein, *Genocide As State Crime: Examples from Rwanda and Bosnia* (Paper Prepared for the Annual Meeting of the American Sociological Association, Toronto, Ontario Aug. 9-13, 1997), at 10.

72. *Id.* at 11.

various mechanisms of justice that would delegitimize and remove from power those responsible for the war and its atrocities, and those most likely to inhibit the return to a multi-ethnic society and the creation of a democratic state. The main negotiators, however, feared that an attempt to craft a peace agreement based on these principles would not be accepted by Milosevic, and thus were unwilling to further threaten the use of force to accomplish these objectives.

Importantly, many U.S. officials now believe that it would have been possible to pressure Milosevic at Dayton into accepting a more just peace than what was ultimately drafted at Dayton. Historians will likely reflect that the American and European habit of negotiating with Milosevic had become so strong that they were unable to correctly assess that Milosevic was desperate for a peace agreement in order to secure his political situation within Serbia, and that by embracing him as a peacemaker at Dayton and providing him with the attainment of his main objectives they sealed the fate of a perpetually divided Bosnia and set the stage for future crimes against humanity in Kosovo.

The second aspect of the diplomatic deficit was the inability of the Dayton negotiators to understand the relationship between all aspects of the conflict. In an effort to resolve the conflict between the Bosnians and the Croats, the American and German teams negotiated a modification of the existing Federation agreement which devolved significant powers from the central government to the Federation. This back door partition orchestrated under the umbrella of strengthening the Bosnian-Croat Federation created an opportunity for the Serbian delegation to reap the benefits of Bosnian and Croat differences without having to spend any of its own political capital. By permitting the negotiations to unfold in such a manner, the international community demonstrated to Milosevic its inability to grasp both the larger dimensions of his plan for a greater Serbia and to orchestrate its own diplomatic efforts in a manner sufficient to build the necessary coalitions to confront and stymie these plans. Moreover, the Federation negotiations during Dayton laid the foundation for a Bosnian constitution which drastically limited the authority of the central government and enabled the Bosnian Serbs to consolidate their power over the territory already under their control and to prevent the significant return of refugees.

Myopic Accommodation

As a result of the perceived need to accommodate the interests of Milosevic, the text of the proposed agreement minimized the future role of force in promoting refugee return and the reintegration of Bosnia, limited the ability of justice-based mechanisms to cull war criminals from civil society, legitimized an ethnically pure Republika Srpska while creating a

constitutionally dysfunctional Bosnian state, and divided territory in a manner that rewarded rather than reversed ethnic cleansing. In essence, the agreement was merely a vehicle, which provided the international community with an opportunity to affix the delegations' signatures to the document.

Consequently, the international community was able to declare that it had brought peace to Yugoslavia, while Milosevic was able to consolidate many of his objectives and prepare for the next round of ethnic cleansing in Kosovo. Moreover, by diplomatically engaging and, at times, toasting Milosevic, the international community confirmed that so long as Milosevic committed ethnic cleansing in moderation, the Americans and Europeans would not prevent him from accomplishing his objectives, and might even confirm his gains in a peace agreement.

The myopic accommodation of Milosevic during Dayton also undermined the real chance for the long term survival of a multi-ethnic Bosnia and set a precedent for what Milosevic could expect from the international community even after it has used force against him. Rather than rolling back the gains of ethnic cleansing, the international community at Dayton formally ratified these gains. Most importantly, the international community permitted the Republika Srpska to retain control of the villages of Srebrenica and Zepa after it was known that the Serbian forces had massacred over 7,000 unarmed civilians from these villages. In fact, the international community allowed Milosevic to force the Bosnian delegation to negotiate the return to Bosnia of the residents of Zepa who had fled into Serbia and were being held in camps. It also required the Bosnian delegation to turn over to the Serbian forces twenty percent of Bosnian territory that had been freed from Serb occupation in order to maintain the 49/51 territorial split. This split was premised on the principle of equivalence among the parties.

Finally, by completely excluding the issue of Kosovo from the Dayton negotiations, the international community signaled to Milosevic that they were unprepared to constrain his actions against the people of Kosovo and that they would permit him to exercise a free hand in the continuation of his regime of oppression and near-apartheid as long as it did not amount to the scale of atrocities seen in Bosnia. Thus, with the tacit acquiescence of the international community, Milosevic continued his suppression of the basic human rights of the people of Kosovo and accelerated their disenfranchisement from civil society. This laid the foundation for the Kosovo war in 1998 and set the stage for yet another Balkan conflict and confrontation with Slobodan Milosevic.

Moral Duplicity

Whereas a basic moral compass would require the presence of truthful declarations concerning Milosevic's culpability, resulting in his prosecution for war crimes, the perceived necessity of Milosevic's participation at Dayton negated the peacebuilders' willingness to follow such an approach. Instead, the peacebuilders, led by Secretary Christopher, legitimized Milosevic as a partner in peace and obfuscated his culpability by arranging for the now famous "handshake" among the three presidents of Serbia, Croatia and Bosnia.[73] This reinforced the perception of the moral equivalence amongst the parties, and rehabilitated Milosevic. If President Izetbegovic was willing to shake his hand, then there would be no bar to full American and European engagement with Milosevic as a legitimate partner. Christopher, however, characterized the handshake in these terms:

> Before my opening statement, I walked around the table to these three, encouraged them to shake hands. Although they had done so before, I believed that such an act – like the Rabin-Arafat handshake on the White House's South Lawn – would provide an important boost symbolically, not only for the delegates in Dayton, but for the millions watching around the world.[74]

The rehabilitation and legitimization of Milosevic was particularly important in this respect because despite the significantly weakened position of Milosevic, much of the American and European efforts were directed at persuading Milosevic to accept a plan favorable to his interests, while maintaining a minimum threshold of provisions that would allow them to declare they had brought peace to Yugoslavia while maintaining Bosnia's status as an independent state.[75] Similar efforts were directed at Tudjman, whose sole goal seemed to be a fixed timetable for the re-establishment of Croatian government control over all Croatian territory, and the creation of an opportunity for the future erosion of Bosnian sovereignty. Most of the diplomatic squeeze was saved for the Bosnian delegation, which on numerous occasions was threatened with closure of the talks and the withdrawal of international support. In order to pressure the victim, it was necessary to both minimize Milosevic's role as the aggressor and Bosnia's as the aggrieved.

As a direct result, the proposed Bosnian governmental structure was characterized by the institutionalization of ethnicity and plagued by numerous mechanisms capable of being utilized by ethnic nationalists to

---

73. *See* Keith Doubt, *We Had to Jump Over the Moral Bridge: Bosnia and the Pathetic Hegemony of Face-Work*, *in* THE CONCEIT OF INNOCENCE: LOSING THE CONSCIENCE OF THE WEST IN THE WAR AGAINST BOSNIA 123-24 (Stjepan G. Meštovic ed., 1995).

74. CHRISTOPHER, *supra* note 40, at 353 n.6.

75. *See id.* at 366.

create political gridlock. This was done in order to give credence to Milosevic's rejection of the notion of a Yugoslav or Bosnian identity, while also creating a political structure, which allowed his protégées to work to undermine the effective reintegration of Bosnia. The political system also disenfranchised individuals, such as members of the Jewish community, who did not belong to one of the three dominant ethnic groups, or individuals who refused to define themselves in ethnic terms. The institutionalized gridlock was subsequently compounded by the failure of Carl Bildt, the former Swedish Prime Minister who served as Bosnia's High Representative — and subsequent High Representatives, to actively utilize their authority to impose legislative initiatives when these mechanisms were abused for the purpose of undermining the Dayton Accords. The willingness of the international community to craft an ethnic-identity based form of government signaled to Milosevic its acceptance of his framework for resolving the Yugoslav conflict and indicated its willingness to sanction arrangements based upon principles of ethnicity as opposed to pluralism and human rights.

Constrained Force

The aversion to the adequate use of force manifested itself numerous times throughout the Dayton negotiations. Most relevant was the refusal of the mediators to create an armed international police force or to provide IFOR with the authority and obligation to create the necessary conditions for the return of refugees and the protection of basic human rights. The failure to include sufficient authority for the police and military forces in the Dayton Accords was compounded by the very limited interpretation given to the authority that was incorporated, and the military's high degree of concern with mission creep. This situation permitted Milosevic and his forces to cement their territorial and political gains accomplished through ethnic cleansing, and signaled that even if force was used against his troops and heavily armed peace enforcers were deployed, they would not take the actions necessary to roll back gains accomplished through ethnic cleansing and widespread and systematic crimes against humanity. Recently the Office of the High Representative (OHR) and the European Union have both called for the deployment of an armed international police force to provide protection for returning refugees and to combat organized crime.

Importantly, during the Dayton negotiations, the American military did not want to be subject to undue constraints on the potential use of force during its deployment in Bosnia, thus ensuring that the Accords provided wide authority for the use of force when IFOR thought it necessary. While providing extensive authority to use force, the Accords actually assigned IFOR forces few obligations other than separating the militaries of the

opposing parties and patrolling the inter-entity boundary line. Thus, while the Accords did not constrain the use of force during the implementation phase of the peace building process, they did not serve to increase the possibility that force or the threat of force would actually be used to ensure the fulfillment of the stated objectives of the Accords.

Marginalized Justice

Immediately prior to the Dayton negotiations, a number of prominent international leaders raised expectations for the chance to craft a lasting peace. Many of these leaders involved in the peace process made a number of public and private declarations concerning the need to incorporate justice into the process -- statements which were missing from earlier attempts to negotiate a peace in the former Yugoslavia. While on their face supportive of the norm of justice, these statements were part of an effort to set the stage whereby justice would be largely (and possibly unnecessarily) sacrificed in order to reach a political settlement with Milosevic. By publicly proclaiming their support for justice in the negotiations, the negotiators satisfied public calls for justice, while creating negotiating space. Sensing the erosion of the influence of the norm of justice and recognizing its potential sacrifice in a negotiated settlement, Justice Goldstone, prior to the Dayton negotiations, indicted the two primary Bosnian Serb culprits, Radovan Karadzic and Ratko Mladic, for genocide on July 24, 1995 and again on November 16, 1995 for the Srebrenica massacre. With these indictments Judge Goldstone effectively prohibited their participation in the Dayton negotiations.

During the actual negotiations, tension developed between the need to establish a variety of mechanisms to cull war criminals from civil society, and the need to accommodate the interests of the aggressors. Unfortunately, as noted by Kenneth Doubt, "at Dayton, the more the Bosnian delegation insists on justice, the less the Bosnian delegation is viewed as being interested in peace."[76]

Out of its desire to ensure that the Accords sufficiently obliged the parties and the international community to arrest, extradite and prosecute war criminals, the Bosnian delegation proposed a number of specific provisions relating to the creation of vetting mechanisms in the police and military and civil authorities as well as the automatic reimposition of sanctions against Serbia and the Republika Srpska for failure to cooperate with the Tribunal.[77] Driven by the need to secure a peace deal, the members of the

76. Doubt, *supra* note 73, at 125.

77. *See* Paul C. Szasz, *The Dayton Accord: The Balkan Peace Agreement*, 30 CORNELL INT'L L.J. 759, 762 (1997).

Contact Group generally objected to the inclusion of such specific obligations as proposed by the Bosnian delegation on the basis that they were unnecessary. More specifically, the Contact Group, led by the British and French, rejected any attempt to require the reimposition of sanctions on parties not cooperating with the International Tribunal, and indicated to the Bosnian delegation that they had personal assurances from Slobodan Milosevic that he would ensure cooperation with the Tribunal -- including compelling the surrender of Ratko Mladic and Radovan Karadzic. In fact, as recounted by General Clark, Milosevic expressly avoided making any commitment to ensure the arrest and extradition of Karadzic and Mladic.[78] The failure to create a mechanism for the automatic reimposition of sanctions proved a crucial strategic mistake as subsequent to Dayton, Slobodan Milosevic expressly refused a personal request by Ambassador Holbrooke to fulfill his "promise" to have Karadzic and Mladic removed to The Hague.[79] Without the meaningful ability to reimpose sanctions, there was simply no incentive for Milosevic to cooperate with the Tribunal. And, as discussed below, the International Crisis Group has concluded that the single most important factor relating to the non-implementation of the Dayton Accords and the continued *de facto* partition of Bosnia was the continued presence of Karadzic in Bosnia.[80]

Concerning the creation of a vetting mechanism, the Contact Group feared this would involve E.U. and International Police Task Force personnel too deeply in the operation of the military and police structures of the sub-state entities and that they might be perceived as biased in the event that they removed a disproportionate number of officials from one of the parties. With respect to the participation of IFOR in the vetting mechanism, the Contact Group military representatives at Dayton steadfastly refused to commit to any involvement in activities associated with the apprehension of war criminals, including the assignment of staff to assist in the identification of those likely responsible for war crimes.

The one concession the Contact Group did make to the Bosnian delegation was to accept the premise that those indicted for war crimes could not stand for elected office or hold other public office after the Dayton Accords were signed.[81] Ambassador Holbrooke would subsequently claim that this second provision was a major success in limiting the ability of

---

78. *See* CLARK, *supra* note 31, at 65.

79. *See* RICHARD HOLBROOKE, TO END A WAR 320 (1997).

80. *See* International Crisis Group, *War Criminals in Bosnia's Republika Srpska, available at* http://intl-crisis-group.org/projects/balkans/bosnia/reports-A400001-02112000.pdf (last visited Feb. 23, 2002) [hereinafter ICG, *War Criminals*].

81. Dayton Peace Agreement, Annex 4, art. IX (Nov. 21, 1995), *available at* http://dosfan.lib.uic.edu/ERC/bureaus/eur/dayton/07Annex4.html.

war criminals to influence the implementation of the Dayton Accords and to poison the development of civil society in Bosnia. Yet this provision did little to inhibit publicly indicted war criminals from exercising significant influence over the electoral process, or from blocking meaningful implementation of the Dayton Accords.

Somewhat ironically, the formal remarks of Secretary Christopher at the initialing ceremony at Dayton were titled, "The Dayton Peace Agreement: Building Peace with Justice[,]" and in his first paragraph, he proclaimed that "the time has come to build peace with justice[,]" arguing that the agreement represented "a victory for all those in the world who believed that with determination, a principled peace is possible."[82]

B. The Consequences of a Faustian Bargain: Empowering War Criminals

Rather than building upon the successful use of force, creating peace through justice and rolling back the gains of ethnic cleansing, the international community at Dayton returned to its preference for accommodation and formally ratified these gains. As a result, the norm and institutions of justice were only minimally present in the agreement, and as a consequence indictable war criminals maintained positions of power and influence and exercised their ability to stall any meaningful implementation of the Dayton Accords.

Compounding this appeasement, the international military forces consistently refused to provide the security necessary for the remaining former residents of Srebrenica to return to their homes, or to discover the whereabouts of the remains of their murdered family members. Such actions clearly signaled to Milosevic that he could generally retain territory conquered through ethnic aggression and that this territory would remain mono-ethnic. Undoubtedly, these perceptions influenced Milosevic's calculation to expel 1.5 million Kosovars during the NATO air campaign in the hopes that the international community would not have the political will to ensure their return. Over time small numbers of mostly elderly refugees began to return to their homes on both sides of the inter-entity boundary line, but there has been no significant return over the past seven years.

The extent of the failure of the Dayton Accords to form the basis for a meaningful peace was illustrated in November 2000, when the International Crisis Group (ICG) documented that over seventy-five individuals indictable for major war crimes have held important positions of power and influence in nearly all the municipalities and political party structures

---

82. CHRISTOPHER, *supra* note 40, at 366-67.

across the Republika Srpska.[83] For instance, paramilitary leader Ljubisa Savic, whose forces were responsible for numerous acts of ethnic cleansing, was supported in his assumption of the position of Deputy Minister of the Interior; Biljana Plavsic, later indicted for genocide,[84] was supported in her position as the President of Srpska, and at one time received the protection of British SFOR troops; Petar Cancar, responsible for the ethnic cleansing of Foca and the establishment of rape camps, served with international support for a time as the Srpska Minister of Justice; and Momcilo Krajisnik, who, with international backing, was elected as the Serbian representative to the Bosnian Presidency, was subsequently indicted for genocide against the people of Bosnia.[85] According to the ICG, these individuals also include the President of the Supreme Court, the Minister of Defense, the second highest ranking police officer,[86] at least seventeen individuals of the Republika Srpska police force, and eleven holding seats on municipal councils – which were acquired in the OSCE-supervised and certified elections.[87] Almost all of these individuals are reported to meet regularly with key members of the international community operating in Bosnia despite the international community's awareness of their culpability for serious crimes committed during the conflict. Interestingly, these indictable war criminals have been allowed and even encouraged by the international community to assume positions of power because they were perceived as important to the implementation of the Dayton Accords.[88]

The widespread presence of indictable war criminals in the institutions of power in the Republika Srpska and the failure of the international community to secure the arrest of those who were indicted emboldened the forces interested in continuing the campaign for ethnic partition during the post-conflict phase of the Yugoslav crisis.[89] In particular, "[t]he failure to arrest Karadzic himself has sent a message to his wartime colleagues and political successors that they can obstruct return, actively work against Dayton implementation, exploit nationalist sentiments, and remain untouchable."[90]

The importance of securing the arrest of Mladic and Karadzic early in the implementation phase of the peace process was essential because, according to

---

83. *See* ICG, *War Criminals*, *supra* note 80, at ii-iii.

84. *See Krajisnik Amended Indictment*, Mar. 21, 2000, IT-00-39; *Plavsic Initial Indictment*, Apr. 7, 2000, IT-00-40.

85. *See* ICG, *War Criminals*, *supra* note 80, at 69.

86. *See id.* at 2.

87. *See id.* at 69, 78.

88. *See id.* at 68.

89. *See* ICG, *War Criminals*, *supra* note 80, at 69.

90. *Id.* at iii.

a Serbian judge interviewed by the ICG, "Karadzic and Mladic represent the single knot holding together a twisted thread of collective and individual guilt that must be untied for Bosnia to begin the reconciliation process."[91] According to the judge, "[f]or the 'small-time' ethnic cleansers of Bosnia, Karadzic demonstrates that ethnic cleansing worked and that they can continue to rule over their fiefdoms, waiting for the weak resolve of the international community to dissipate entirely."[92]

Those interested in securing the ethnic separation of Bosnia actively used their power and influence to undermine the conditions necessary for the fulfillment of the Dayton Accords by obstructing the return of refugees, inhibiting freedom of movement, and denying basic human rights to non-Serb minorities. In many instances, the public institutions themselves have actively prevented the fulfillment of these obligations,[93] as "[o]ften, those individuals who are meant to be protecting and supporting returning minorities are exactly the same individuals who expelled them in 1992-95."[94]

Individuals who were culpable for the atrocities committed during the war have continued the *de facto* partition of Bosnia and the denial of basic human rights to non-Serbs because they are allowed to control essential political and economic institutions. According to the ICG, "[t]hey function as pillars of their respective communities, dispensing political and economic patronage. Many control the local economies, both legal and illegal, while others continue to control illegal paramilitary groups."[95] In fact, the ICG asserts, "[i]n all [Republika Srpska] communities, indicted and suspected war criminals appear to enjoy respected status,"[96] and many receive protection from the Republika Srpska police force.[97]

In May 2001, a British SAS officer made public a report prepared under his command, which called upon the international authorities in Bosnia to arrest twenty-two Croatian individuals who had committed crimes during the war and were engaged in "the violent intimidation of refugees trying to return to Stolac, in south-eastern Bosnia, in 1998 and 1999."[98] According to the officer, while some war-crimes suspects were in hiding, most of

---

91. *Id.* at 77.

92. *Id.*

93. *See* International Crisis Group, *Is Dayton Failing?: Bosnia Four Years After the Peace Agreement*, *at* http://www.intl-crisis-group.org/projects/showreport-cfm?reportid=58 (last visited Apr. 16, 2002).

94. ICG, *War Criminals*, *supra* note 80, at 77.

95. *Id.* at 2.

96. *Id.*

97. *See id.*

98. Nick Thorpe, *Goat Town Now a Criminal Haven*, GUARDIAN (May 2, 2001), *available at* http://www.guardian.co.uk/print/0,3858,4179182.html (last visited Feb. 23, 2002).

those identified as torturers at the Dretelj, Heliodrome and Gabela concentration camps still held power and were often sighted in public by their former victims, who had begun to return in small numbers and who were still subject to gross violations of human rights. Despite the report, none of the individuals identified were ever indicted or arrested.

The manner in which the international community engaged Milosevic during the Dayton negotiations, and the concessions forced upon the Bosnians as a condition for Serbian acquiescence and the deployment of NATO troops, significantly bolstered and encouraged Milosevic's continued influence over the nature of the Yugoslav conflict. By elevating Milosevic to a partner in peace and by permitting him to use the Dayton negotiations as a means for accomplishing many of his objectives, the international community tacitly sanctioned Milosevic's use of ethnic aggression as a means for achieving his political objectives, confirmed to Milosevic and the people of Serbia the unwillingness and inability of the international community to deny Milosevic the means to continue his program for a greater Serbia, and signaled that they would be unlikely to respond in a forceful manner to actions in Sandzak or Kosovo.

Unfortunately for the people of Yugoslavia, as the approach of coercive appeasement was validated during the Dayton negotiations as a method which could ultimately produce perceived results, it became the template for the approach of Dayton implementation, and reconfirmed for Milosevic the utility of this strategy of ethnic aggression. Seeking to capitalize on his gains in Bosnia, Milosevic undertook a modified approach in Kosovo in 1998.

THE KOSOVO CRISIS: FROM BOSNIA REDUX TO DIPLOMACY BACKED BY FORCE

From 1995 to 1998, while the international community turned implementation of the Dayton Accords over to European bureaucracies, Milosevic continued his campaign of ethnic separation and ethnic terror against the majority Kosovar Albanian population of Kosovo. Frustrated with the lack of progress that Ibrahim Rugova's peaceful resistance was having on Milosevic's regime, and acutely aware of his failure to attract the interest of the United States and its European allies, a segment of the Kosovar Albanian population began to engage in armed attacks against Serbian police and perceived Albanian collaborators. These armed groups soon formed themselves into the Kosovo Liberation Army (KLA), which established independence for Kosovo as its primary political objective. The efforts of the KLA had their intended effect as the international community began to take notice of the growing crisis in the region.

A. The Legacy of Dayton: Treating Milosevic as a Peacemaker

The initial approach to the Kosovo conflict mimicked the approach to the Bosnia conflict – that of coercive appeasement. Importantly, however, under the stewardship of Prime Minister Tony Blair, Secretary of State Madeleine Albright and General Wesley Clark, the British and Americans slowly steered the policy to one of diplomacy backed by force, and were able to finally bring an end to Milosevic's reign of terror in the Balkans.

It is important to note that even after the Americans and their European allies adopted the approach of diplomacy backed by force and undertook an air campaign to deter Milosevic's aggression, the proponents of accommodation sought to significantly limit the scope of the air campaign. As will be discussed below, the limited nature of the air campaign created a situation where the Americans and Europeans were forced to commit to the deployment of ground troops in order to protect the Kosovar Albanians. Had the ground troops actually been deployed in combat, there would have been significant casualties among these forces as well as among Serbian forces. If the Clark doctrine of overwhelming force had been adopted at the outset of the air campaign, it is unlikely that the United States and its European allies would have been forced to contemplate the deployment of ground troops. It is thus the case that coercive appeasement had become so ingrained in American and European policymaking that it created a situation where the only option remaining became one which was likely to involve large casualties among allied and Serbian troops.

Diplomatic Deficit

In 1997, the United Nations' General Assembly began calling for an end to Serbian repression in Kosovo.[99] This failed to bring an end to the violence as both the KLA and Serbian forces intensified their efforts, resulting in the deaths of numerous civilian Albanians and Serbs. In response to the ongoing skirmishes, NATO began air exercises in Macedonia and Albania as a show of force to the warring parties, and the United Nations continued to debate possible military action. The opposition to the use of force by Russia and China, however, stymied the U.N.'s efforts and effectively eliminated the possibility of the use of force as a tool of negotiation.[100]

The need for consensus with its allies and within NATO also restrained

99. *See* Ted Baggett, *Human Rights Abuses in Yugoslavia: To Bring an End to Political Oppression, the International Community Should Assist in Establishing an Independent Kosovo*, 27 GA. J. INT'L & COMP. L. 457, 464 (1999).

100. *See id.* at 458, 461-62.

the international community from taking the action promised by the Bush administration in the infamous Christmas warning, which threatened American use of force if Milosevic used force against the Kosovar population.[101] According to Ambassador Holbrooke, the Clinton administration was reluctant to propose threatening the use of force at this stage given its multinational foreign policy approach, which was designed to maintain cooperative relations with its European allies.[102] Given that the British, French and Germans believed it necessary to obtain United Nations Security Council authorization prior to using force, the United States acquiesced and tailored its foreign policy accordingly.[103] Thus, the international community embarked on a two-pronged foreign policy approach of limited economic inducement and accommodation from the autumn of 1997 to the early winter of 1998.

With primary responsibility for resolving the crisis, the Contact Group adopted the position of mediator and facilitator, encouraging the parties to "join in a peaceful dialogue,"[104] challenging the parties to "assume their responsibility to promote stability and a solution,"[105] indicating that it would "support a mutually agreed solution that respects democratic standards,"[106] and offering its support in implementing an Education Agreement.[107] The unintended effect of these statements only served to undermine the Contact Group's credibility and authority, as it was clear to the parties that it would not become politically involved in the negotiation process beyond facilitating dialogue.

The result of the international community's abdication of primary responsibility for the handling of the Kosovo crisis to the Contact Group was the perpetuation of hostilities throughout 1997 and 1998. In response,

---

101. *See* David Binder, *Bush Warns Serbs Not to Widen War*, N.Y. TIMES, Dec. 28, 1992, at A6; Elaine Sciolino & Ethan Bronner, *How a President, Distracted by Scandal, Entered Balkan War*, N.Y. TIMES, Apr. 18, 1999, at A6.

102. *See* CECIL V. CRABB, JR., ET AL., CHARTING A NEW DIPLOMATIC COURSE: ALTERNATIVE APPROACHES TO AMERICA'S POST-COLD WAR FOREIGN POLICY 138 (2001).

103. *See* Ambassador Richard Holbrooke & Ambassador William Walker, On the Record Briefing on the Kosovo Situation (U.S. Dep't of State, Oct. 28, 1998).

104. Contact Group Foreign Ministers, Press Statement on Kosovo (Sept. 24, 1997), *available at* http://www.ohr.int/other-doc/contact-g/default.asp?content-id=3543; *see also* Contact Group Foreign Ministers, Press Statement on Kosovo (Feb. 25, 1998), *available at* http://www.ohr.int/other-doc/contact-g/default.asp?content-id=3549.

105. U.S. Dept. of State Office of the Spokesman, Press Statement on Kosovo (Jan. 8, 1998), *available at* http://secretary.state.gov/www/briefings/statements-/1998/ps980108.html.

106. *Id.*

107. *See id.*

the international community, working through the Contact Group, modified their approach in late February and early March 1998 to incorporate the approaches of economic inducement and enhanced diplomatic engagement. Secretary of State Madeleine Albright also publicly proposed that the Contact Group establish a working group to study how best to "implement an asset freeze and a ban on foreign direct investment, should that prove necessary."[108] The Contact Group and the U.N.[109] then called upon the United Nations High Commissioner for Human Rights (UNHCHR) to send a mission to Kosovo; expressed support for the return of the OSCE long-term mission and the establishment of a Kosovo specific OSCE mission; encouraged OSCE member states to intensify visits to Kosovo; and proposed the creation of an international consortium including NGOs that would promote civil-society building in Kosovo and the distribution of humanitarian assistance.

The problematic aspects of such a stance were identified by the United States' lead member of the Contact Group, Secretary Albright, who expressed concern that the group was "meeting for the sake of meeting, rather than resolving something."[110] These ineffective meetings were the result of allied concerns over the legality of intervention beyond mediation and negotiation. Without definitive leadership, nor a coherent, long-term strategy for resolving the crisis, the crisis continued to escalate. In fact, these diplomatic measures were seen as "diplomacy light," and had the effect of signaling that the Contact Group and the international community, particularly the United States, were not entirely serious about involving themselves in the mediation of the Kosovo conflict. In particular, by calling upon enhanced roles for the UNHCHR, the OSCE and even NGOs, the Contact Group and the UN signaled that they were unwilling to take a leading role in resolving the conflict, setting the stage for the same type of institutional competition and disarray that had hampered efforts to achieve peace in Bosnia.

These concerns were well reflected during the Rambouillet/Paris negotiations. The selection of Rambouillet as the site of the peace talks was an important concession to America's European allies. Despite the fact that the American component of the Contact Group had been responsible for

---

108. Secretary of State Madeleine K. Albright, Statement at the Contact Group Meeting on Kosovo (Mar. 25, 1998), *available at* http://secretary.state.gov-/www/statements/1998/980325.html.

109. *See* U.N. S.C. Res. 1160, U.N. SCOR, 53rd Sess., 3868th mtg., U.N. Doc. S/RES/1160 (1998).

110. *PBS Frontline: The War in Europe*, Interview with Madeleine K. Albright, Secretary of State [hereinafter Albright Interview], *at* http://www.pbs.org/wgbh/pages-/frontline/shows/kosovo/interviews/albright.html (last visited May 13, 2002).

much of the diplomacy prior to the peace talks, the British and French representatives insisted on holding the talks in Europe. This was done in part to compensate for the perpetual political slight of having had the Bosnian peace declared in Dayton, Ohio after the massive investment of European resources and political capital in the peace process, as well as to signal that Europe would take the political lead in implementing any Kosovo peace arrangement.

Just as the Europeans had earlier declared that the United States would manage the dissolution of the Soviet Union while the EU managed the dissolution of Yugoslavia, the Europeans were now declaring that the United States would manage the Bosnian peace process, while the Europeans would take on responsibility for the Kosovo process. The dramatic entrance of President Jacques Chirac in a wheezing and battered Presidential helicopter, which landed on the front lawn terrace of the Rambouillet Chateau on the first evening of the peace talks, was symbolic of the European desire to re-establish their political control over events in the European theater despite their tangible lack of credible resources to accomplish the task of peace building in the Balkans without the overwhelming assistance of the United States.

On February 6, the first round of talks between the Kosovar Albanians and Serbs were held in Rambouillet, France under the auspices of the Contact Group. The Rambouillet/Paris negotiations were similar to the Dayton negotiations in many important procedural and substantive respects. The Rambouillet/Paris negotiations, however, diverged from the Dayton model in a number of important respects that indicated an ability by a limited number of key players in the United States and European Union alliance to recognize the failure of their approach of coercive appeasement and begin to rely upon the credible threat of the use of force. Moreover, the Rambouillet/Paris Accords lacked many of the more gaudy trappings of moral equivalence displayed at Dayton, such as the introductory handshake -- in large part because the Serbs refused to shake hands with the KLA representatives of the Kosovar Albanian delegation.

While at Dayton, Milosevic was primarily interested in securing his territorial gains through the peace process, and genuinely concerned about further NATO use of force. At Rambouillet, however, he had not yet accomplished his strategic objectives on the ground and was unconvinced that NATO would in fact use force to prevent him from accomplishing those objectives. In large part, Milosevic's perspective was likely influenced by publications written by former U.S. and European government officials, like Ambassador Holbrooke, who detailed the American and European reluctance to use force in Bosnia, and who declared that NATO could have only sustained the most minimal form of an air campaign.

The forced absence of NATO from the peace talks by the French gov-

ernment also undermined the perceived seriousness of the negotiations in Rambouillet. This state of affairs required that the Contact Group employ significant pressure on the Kosovar delegation, which opened the door for certain Contact Group members to slip back into their more familiar role of coercive appeasers, and also required the American delegation to make overt threats of the use of force against the Serbian delegation in an attempt to compel their constructive participation. As a result, neither the Kosovar nor Serbian delegations to the peace talks believed it to be absolutely necessary to sign a peace agreement at the conclusion of the talks in Rambouillet.

As the negotiations reached a stalemate, with Milosevic refusing to consider NATO peacekeepers in Kosovo and the KLA's continued call for independence, the international community again relied on diplomatic wrangling to reinvigorate the peace talks. To this end, Christopher Hill was sent to Belgrade in an effort to persuade Milosevic to accept the agreement proposed at Rambouillet. Again, it seems, the international community persisted in its belief that Milosevic was the key to a successful settlement, thereby promoting him once again to the position of "partner in peace."[111] Secretary Albright was also dispatched to pressure the KLA to accept the peace plan, which they had refused to sign, insisting on nothing less than full independence.[112] The refusal of the Kosovar delegation to sign the agreement placed the Clinton administration in a quandary, as their strategy to pressure Milosevic into capitulating by threatening air strikes if the Serb delegation failed to reach an agreement hinged upon the Kosovars' acceptance of the peace deal.[113] While the Kosovar delegation eventually accepted the terms of the peace agreement, the Serbian delegation refused to sign – an act which brought about the collapse of the approach of coercive appeasement and opened the way for the Albright and Clark doctrine of diplomacy backed by force as will be discussed in the section below on the use of force.

Myopic Accommodation

Despite knowing that Kosovo was the point where Milosevic had promised to make his stand for the Serb people, and that he had demonstrated a clear pattern of conduct of moving from one conflict to another, the international community, and the United States in particular, demonstrated no

---

111. *See* Tom Walker, *U.S. Stance on Kosovo Talks and 'Secret Flight' Infuriate Diplomats*, TIMES (LONDON), Feb. 18, 1999, at 17, *available at* 1999 WL 7973918.

112. *See* Marc Weller, *The Rambouillet Conference on Kosovo*, INT'L AFF., Apr. 1999, at 233.

113. *See* Steven Erlanger, *Albanian 'Nay' Puts Milosevic in Role of Yea-Sayer*, N.Y. TIMES, Feb. 22, 1999, at A6.

significant interest in 1997 in taking the lead in pursing preventative diplomacy or in holding Milosevic accountable for his actions in Kosovo. Instead, the international community persisted in its policy of viewing Milosevic as a partner in peace, and was content to allow the U.S. Department of State, working through the Contact Group, to work on negotiating a peaceful settlement to the crisis. As noted by General Clark, even after Milosevic had begun a campaign of village busting in the summer of 1998 and even though "[f]or months it had been clear that a fundamental problem in Kosovo was Milosevic himself, . . . there was still an effort to rely on him as a legitimate negotiating partner."[114] In fact, despite the lack of progress in negotiations, the American and European negotiators would periodically announce progress, and according to Clark this "lulled Western governments into postponing the tough decisions before them."[115]

This decision was the result of several factors. Most importantly, the international community was reluctant to become enmeshed in the Kosovo crises because it continued to rely on Milosevic in implementing the Dayton Peace Accords and viewed him as a necessary actor in successful negotiations with the Kosovar Albanians. Thus, its ability to pressure Milosevic to end the violence was constrained and was a key factor behind the international community's reluctance to adopt a more confrontational position.

This reluctance to take a hard line approach with Milosevic resulted in the continuation of the international community's appeasement of him. The primary indication of international efforts to accommodate Milosevic is evidenced by the continual reaffirmations of Serbia/Montenegro's sovereignty and territorial integrity. Each Contact Group statement from September 24, 1997 to February 25, 1998 declared support for an enhanced status for Kosovo only within Serbia/Montenegro.[116] The Contact Group members also ensured that each U.N. Security Council resolution dealing with the Kosovo crisis contained a reference to the sovereignty and territorial integrity of Serbia/Montenegro, as well as other states in the region.[117]

---

114. CLARK, *supra* note 31, at 128.

115. CLARK, *supra* note 31, at 126.

116. *See* Contact Group Foreign Ministers, Press Statement on Kosovo (Sept. 24, 1997), *available at* http://www.ohr.int/other-doc/contact-g/default.asp?content-id=3543; Contact Group Foreign Ministers, Press Statement on Kosovo (Jan. 8, 1998), *available at* http://www.ohr.int/other-doc/contact-g/default.asp?content_id=3547 [hereinafter Contact Group Statement on Kosovo].

117. *See, e.g.*, U.N. S.C. Res. 1160, U.N. SCOR, 53rd Sess., 3868th mtg., U.N. Doc. S/RES/1160 (1998), U.N. S.C. Res. 1199, U.N. SCOR, 53rd Sess., 3830th mtg., U.N. Doc. S/RES/1199 (1998); U.N. S.C. Res. 1203, U.N. SCOR, 53rd Sess., 3937th mtg., U.N. Doc. S/RES/1203 (1998), U.N. S.C. Res. 1244, U.N. SCOR, 54rd Sess., 4011th mtg., U.N. Doc. S/RES/1244 (1999).

The support for Serbia/Montenegro's sovereignty and territorial integrity, while designed to accommodate the interests of the Serbian regime, was also based on the realpolitik concerns of Contact Group members like Russia and the United Kingdom, which were mired in their own secessionist movements. Furthermore, the principles of sovereignty and territorial integrity are well founded in the normative structure of public international law -- although they do not justify, nor permit the type of actions undertaken by the Serbian regime against the people of Kosovo.

It was not until after his outright refusal to sign the Rambouillet/Paris Accords and the continued massacre of Kosovo civilians that the international community finally realized that Milosevic was not in fact a partner in peace, and ceased its efforts to myopically accommodate his interests.

Moral Duplicity

As a consequence of the failure of the Rambouillet peace talks, the oppression of the Kosovar Albanians became more severe and their commitment to peaceful resistance waned. With the formation of the KLA and its use of violence against Serbian government officials and Kosovar and Serbian civilians, Milosevic was able to more aggressively utilize his Ministry of Interior forces to terrorize the civilian population. In a number of instances, Serbian Ministry of Interior Forces (MUP) forces would kill large numbers of civilians during their missions against the KLA.

The international community's response to the crisis was first to minimize its severity and importance. When the media began to report the possibility of mass killings and mass graves, these reports were rejected by European political institutions. When credible accounts continued to emerge from Kosovo, the international community sought to promote a dialogue between the Kosovars and Milosevic. These efforts delegitimized the pacifist leader of the Kosovars, and aided in the rehabilitation of Milosevic.

The Contact Group, led by Russia and the United States, also sought to create a moral equivalence between Milosevic and the Kosovars by declaring that the KLA was a terrorist organization.[118] The Russian government also successfully persuaded the U.N. Security Council to include a condemnation of Albanian terrorist action in each Kosovo related resolution adopted by the Council.[119] These declarations were initially taken as a

---

118. *See* Contact Group Statement on Kosovo, *supra* note 116; Special Representative Robert S. Gelbard, Press Conference in Serbia and Montenegro (Jan. 15, 1998), *available at* http://www.state.gov/www/policyremarks/1998/980115-gelbard_belgrade.html [hereinafter Gelbard Statement].

119. *See, e.g.*, U.N. S.C. Res. 1160, U.N. SCOR, 53rd Sess., 3868th mtg., U.N. Doc. S/RES/1160 (1998), U.N. S.C. Res. 1199, U.N. SCOR, 53rd Sess., 3830th mtg.,

green light for Milosevic to continue his aggressive actions against the Kosovar population. While the United States possessed credible evidence that the KLA had committed unlawful acts, the use of the term terrorist was legally inaccurate and had the effect of endorsing Milosevic's characterization, thus lending perceived American support to his use of indiscriminate force to eliminate the KLA threat.

Although the Russian Foreign Ministry led the effort to identify the Kosovar armed resistance as terrorists, the U.S. Department of State also adopted an aggressive policy of identifying the KLA as a terrorist organization. Despite the fact that the Central Intelligence Agency had warned the Clinton administration in both January and February of 1998 that Serb forces were mobilizing along the Kosovo border for a planned crackdown on the Kosovar Albanians, U.S. policy makers focused on moral duplicity as a means of deescalating the crisis.[120] This entailed sending Robert Gelbard to Belgrade for the purpose of persuading Milosevic to halt the offensive.[121] While the purpose of the visit was ostensibly to connote the United States' sentiments regarding the situation in Kosovo, Gelbard used this opportunity to declare the U.S. Department of State's position that the KLA was a terrorist organization. Thus, on January 15, 1998, Ambassador Gelbard declared in Belgrade, "[w]e strongly believe that violence is not the answer from either side, and we strongly oppose any kind of terrorist activity from Kosovar Albanians or anybody else, as well as violence fomented by government security forces."[122] Subsequently, on February 22, 1998, he again declared in Prishtina, "[a]s I have said before, I consider these to be terrorists actions, and it is the strong and firm policy of the United States to fully oppose all terrorists actions and all terrorists organizations."[123] In subsequent public statements, Gelbard threatened to take action to interrupt the ability of the KLA "to sustain themselves through the full chain of supply."[124]

The invocation of the designation "terrorists" and the threat to impose sanctions on the organization was particularly important as "terrorists" was the term used by Milosevic to describe the KLA and served as the

---

U.N. Doc. S/RES/1199 (1998); U.N. S.C. Res. 1203, U.N. SCOR, 53rd Sess., 3937th mtg., U.N. Doc. S/RES/1203 (1998)

120. *See* R. Jeffrey Smith, *U.S. Assails Government Crackdown in Kosovo*, WASH. POST, Mar. 5, 1998, at A23.

121. *See* Jane Perlez, *U.S. Warned Serb Leader Not to Crack Down on Kosovo Albanians*, N.Y. TIMES, Mar. 15, 1998, at 6.

122. Gelbard Statement, *supra* note 118.

123. Special Representative Robert S. Gelbard, Press Conference in Prishtina, Kosovo (Feb. 22, 1998), *available at* http://www.fas.org/man/dod-101/ops/docs-/press31.html.

124. *Fighting in Kosovo Intensifies*, ASSOCIATED PRESS, Jul. 8, 1998, at 1.

bedrock of his justification for the use of force against the civilian Kosovar Albanian population. Moreover, the threat to impose sanctions created a moral equivalence between the Kosovars and the Serbs, something Milosevic had successfully manipulated to his advantage in the Bosnian crisis. Importantly, the U.S. government subsequently found that there was no legal basis for the KLA to be designated a terrorist organization and instructed its diplomats to cease using that term. Over time, and in the face of increasing atrocities, the Contact Group lessened its approach of moral equivalism and began to identify the Serbian regime as primarily responsible for the atrocities.[125]

Constrained Force

In the face of familiar atrocities and the increasing radicalization of the Kosovar community, the peace-builders initially responded with humanitarian assistance. While this approach served as a policy Band-Aid, it did not provide a means for achieving a meaningful end to the conflict.[126] During this time, the peacebuilders avoided any meaningful threat of the use of force. According to Ambassador Holbrooke, the British, French and Germans believed it necessary to obtain U.N. Security Council authorization prior to using force.[127] Ambassador Holbrooke also opined that the creation of a common strategy was hampered by the fact that "'democracies take a while to get their act together.'"[128]

In March 1998, Serb forces cracked down on the KLA, resulting in the death of eighty-five Kosovar Albanians. It was at this time that the United States, along with its European allies, united to denounce the violence and sought options for how to proceed with the conflict. From the beginning of the conflict, Secretary Albright believed that diplomacy alone was insufficient to achieve a resolution to the conflict. Instead, Albright believed that the use of force was the only option in securing Milosevic's acquiescence to the demands of the international community. As stated by Secre-

---

125. *See* Contact Group Foreign Ministers, Press Statement on Kosovo (Mar. 9, 1998), *available at* http://www.ohr.int/other-doc/contact-g/default.asp?.content-id=3550.

126. As eventually described by the E.U. Humanitarian Aid Commissioner Emma Bonino in 1998, "[t]he international community must face the reality. There is no way we can avert catastrophe without a political solution." *UN Fears Disaster in Kosovo this Winter*, TORONTO STAR, Aug 20, 1998, at A22. Ms. Bonino later observed that "the situation [in Kosovo] is really very bad. Security has not improved. These people had their homes burned down. Now it's snowing. The real problem is to get compliance from Milosevic. Humanitarian aid can provide a lot, but it cannot provide security." REUTERS, Oct. 26, 1998.

127. *See* Holbrooke & Walker, *supra* note 103.

128. *Id.*

tary Albright, "Slobodan Milosevic has made it clear that he is spurning the incentives that the United States and others have offered him in recent weeks -- unfortunately the only thing he truly understands is decisive and firm action."[129] Albright's main purpose was to use her rhetoric to "push the European allies, American public opinion, and even her own government toward concerted action designed to avert the kind of human tragedy that had happened in Bosnia."[130]

These strong statements and Albright's position on the use of force, however, were obfuscated by the international community, which consistently balked at the notion of utilizing the threat of the use of force as a foreign policy tool. During this time, General Wesley Clark, the NATO Supreme Allied Commander, had developed a strategic attack plan designed for Milosevic's capitulation. Acting on this plan, and under pressure from the United States, the NATO member states developed a "light" version of Clark's proposal and adopted Activation Warning, which is the first stage in a three stage process procedure for the approval of NATO military action. The ACTWARN applied to a Phased Air Operation and a Limited Air Operation.[131] With this threat in hand, Ambassador Holbrooke was dispatched to Belgrade for the purpose of persuading Milosevic to cease the Serb forces' activities against the Kosovar Albanians. Ambassador Holbrooke was fully briefed on the Activation Warning and was permitted to threaten the use of force, but it is unclear whether he relied upon this option.[132] According to General Clark, Holbrooke confided that Secretary Cohen had forbidden him to offer NATO ground troops as peacekeepers as part of any agreement reached with Milosevic.[133]

The result Holbrooke's meeting with Milosevic was the Holbrooke-Milosevic deal of October 1998 (October agreement), which provided for the unarmed deployment of OSCE monitors, and the unarmed NATO over flight of Kosovo. Importantly, to the detriment of the policy objectives devised by the Clinton administration, the deal provided Milosevic with a significant strategic windfall. He was able to position himself yet again as a peacemaker with whom the United States could do business, while at the same time undermining consensus among the NATO member states for the use of force to stop further acts of ethnic cleansing. More importantly, as explained by General Clark, the presence of unarmed observers, "inside

---

129. Secretary of State Madeleine K. Albright, Press Briefing at the Ministry of Foreign Affairs (Mar. 7, 1998), *available at* http://secretary.state.gov/www-/statements/1998/980307.html.

130. IVO H. DAALDER AND MICHAEL. O'HANLON, WINNING UGLY 28 (2000).

131. CLARK, *supra* note 31, at 135.

132. *See id.*

133. *See id.* at 137-38.

Kosovo also vitiated the implicit NATO threat against Milosevic or his forces."[134]

As soon as the October agreement was finalized, it became apparent to the international community that Milosevic was reneging on the agreement. Intelligence reports indicated that Milosevic, in violation of the agreement, was amassing forces between Prishtina and Belgrade.[135] In response, President Clinton dispatched General Clark to pressure Milosevic into abiding by the terms of the agreement.[136] Clinton did not, however, give General Clark sufficient latitude to threaten the use of meaningful force if Milosevic failed to adhere to the agreement. Predictably, the meeting failed to convince Milosevic to implement the provisions of the October agreement. Given the perceived failure of the agreement, Clinton's foreign policy team struggled to ascertain Milosevic's intentions and how best to respond.

In December 1998 and January 1999, the fighting continued between the KLA and Serb forces, resulting in the death of more than 70 KLA rebels and 6 Serb civilians. During this time, public criticism and demand for action spurred Secretary Albright to increase pressure on the White House to push for the use of force. Despite Albright's efforts, the Clinton administration remained unwilling to advocate or plan for the use of force.[137] Secretary of Defense William Cohen and National Security Advisor Samuel Berger raised the most strident objections. Cohen's reluctance stemmed from his concern that U.S. troops would become involved in a militarily conflict without a strategic exit plan. Berger, on the other hand, never seemed to quite comprehend the seriousness of the threat posed by Milosevic and the capacity of the United States and its allies to prevent further atrocities in the Balkans.[138]

Ironically, while the international community was debating its next course of action in light of the failure of the October agreement, Serbian military and paramilitary forces massacred over forty civilians in the Kosovo town of Racak. Subsequent intelligence intercepts reported in the *Washington Post* established a clear link between the front line forces responsible for the massacre, the subsequent attempt to cover it up, and high-level officials in the Serbian government.[139] The massacre was met

---

134. *Id.* at 126.

135. *See PBS Frontline: War in Europe*, Interview with General Wesley Clark, NATO Supreme Allied Commander, *at* http://www.pbs.org/wgbh/pages/frontline-/shows/kosovo/interviews.clark.html (last visited February 23, 2002).

136. *See id.*

137. *See* Albright Interview, *supra* note 110.

138. DAALDER & O'HANLON, *supra* note 130, at 70.

139. *See* R. Jeffrey Smith, *Serbs Tried to Cover Up Massacre; Kosovo Reprisal Plot Bared by Phone Taps*, WASH. POST, Jan. 28, 1999, at A1.

by a rapid series of public denunciations by President Clinton and Secretary Albright.[140]

The Racak massacre, the collapse of the October agreement, and the apparent failure of earlier mediation efforts resulted in increased political pressure, especially from the United States Congress. Before, and immediately after the massacre, a number of prominent U.S. Senators began to call for the use of force to promote a cessation of the hostilities. For instance, Senator Mitch McConnell declared, "[i]t is time for the United States to accept reality, recognize Kosovo's independence and provide Prishtina's leadership with the political and security assistance necessary to halt [Serbia's] genocidal war . . . [and] [d]emand a NATO vote to implement the Activation Order for air strikes."[141] Similarly, Senator Frank Lautenberg declared, "clearly Milosevic thinks he can get away with murder, literally, and NATO needs to send him a strong message. He is the Saddam Hussein of Europe, and force is the only language he seems to understand. NATO must follow through on its threats with air strikes to force the Serbs to respect their commitments."[142] The public also exerted public pressure as they increasingly called for a more aggressive incorporation of the use of force into the peace-building process. One of the most insightful and potentially influential calls for force came from the *Washington Post's* Editorial Board, writing,

> This humanitarian disaster [in Kosovo] cannot be ended without a political solution, and a political solution is impossible without a U.S. resolve to use force, if necessary, against Mr. Milosevic's marauding soldiers. President Clinton and his team have promised again and again to show such resolve, but their threats have proved empty. Instead Mr. Clinton sends his emissaries, again and again, to plead with the war criminal to stop his crimes. Mr. Milosevic has learned he can defy them at no cost. . . . The longer Mr. Clin-

---

140. *See* President Bill Clinton, Statement by the President on the Massacre of Civilians in Racak (Jan. 16, 1999); Secretary of State Madeleine K. Albright, Press Availability on Kosovo Following Volunteer Event at Children's Hospital (Jan. 18, 1999), *available at* http://secretary.state.gov/www/statements/1999/990118.html.

141. Senator Mitch McConnell, *Independence for Kosovo*, WASH. POST, Jan. 22, 1999, at A35.

142. Senator Frank Lautenberg, *Crisis in Kosovo*, N.Y. TIMES, Jan. 20, 1999, at A30. Perhaps recognizing the consequences of equivocating on the use of force, Senator Lautenberg also stated:

> NATO should give Mr. Milosevic a short deadline to comply with his promises, including a pullout of his forces and full cooperation with international monitors, humanitarian agencies and the war crimes tribunal. If he refuses, it should pull out the monitoring force and bomb selected Serbian military targets. If, as has been his habit, he promises compliance but then reneges, NATO must immediately resume the threat.

Id.

ton dithers the greater the cost will be.[143]

Given these strong sentiments and a lack of other viable options, Secretary Albright and Prime Minister Blair pushed the approach to one of "diplomacy backed by force." This approach included increased reliance on the threat of the use of force,[144] and the introduction of direct mediation in the form of proximity peace talks similar to those held at Dayton.[145] This later provision was incorporated to tie the threat of force to political objectives as means of garnering NATO support, whose members wanted another attempt at reaching a peace agreement.[146] With the U.S. strategy clearly in place, on January 28, 1999, NATO issued a declaration (1) indicating its full support for the "conclusion of a political settlement under the mediation of the Contact Group," (2) demanding that the Serbian regime comply with their commitments to withdraw military forces under the October 25, 1998 agreement with NATO and with the obligations set forth in the agreement with OSCE, and (3) further demanding that the Serbian regime cooperate fully with the Yugoslav Tribunal by granting immediate and unrestricted access to its representatives so they could carry out their investigation of the Racak massacre and by ensuring the safety of the investigators.[147]

Subsequently, the Contact Group gave Milosevic and the Kosovar Albanians an ultimatum on January 29, 1999 to attend peace talks in France in February or face NATO air strikes. On January 30, 1999, the North Atlantic Council (NAC) reactivated the October 13, 1998 NATO Activation Order and declared that "[t]he crisis in Kosovo remains a threat to peace and security in the region. NATO's strategy is to halt the violence and support the completion of negotiations on an interim political settlement for Kosovo, thus averting a humanitarian catastrophe."[148] If the parties failed to reach agreement at Rambouillet, the NATO Secretary-General was authorized to order air strikes against Serbia/Montenegro. The NAC also announced that it was prepared to take unspecified "appro-

---

143. *A Massacre Without Knives*, WASH. POST, Sept. 16, 1998, at A16.

144. *See* Press Release, North Atlantic Council (99)12 (Jan. 30, 1999), *at* http://www.nato.int/docu/pr/1999/P99-012e.htm.

145. Secretary of State Madeleine K. Albright, U.S. Dept. of State on NATO Final Warning on Kosovo (Jan. 30, 1999), *at* http://secretary.state.gov/www-/statements/1999/990130.html.

146. *See PBS Frontline: War in Europe*, Interview with Samuel Berger, National Security Advisor, *at* http://www.pbs.org/wgbh/pages/frontline/shows/kosovo-/interviews/berger.html (last visited Feb. 23, 2002).

147. *See* NATO Secretary General Dr. Javier Solana, Press Statement on Kosovo (99)11 (Jan. 28, 1999), *at* http://secretary.state.gov/www/statements/1999/990130.html (last visited Feb. 23, 2002) [hereinafter Solana Press Statement].

148. Press Release, North Atlantic Council, *supra* note 144.

priate measures" in the event of a failure by the Kosovar Albanians to fully cooperate with the peacebuilders.[149]

The motivations for a strategy of "diplomacy backed by the threat of force" included a mix of realpolitik assessment of American and European security interests and the political need to save diplomatic face.[150] As explained by President Clinton, "[i]n this volatile region, violence we fail to oppose leads to even greater violence we will have to oppose later at greater cost . . . . There is a serious risk the hostilities [in Kosovo] would spread to the neighboring new democracies of Albania and Macedonia, and re-ignite the conflict in Bosnia we worked so hard to stop. It could even involve our NATO allies, Greece and Turkey. . . . The time to stop the war is right now."[151] This approach would set the tone for the Rambouillet peace negotiations.[152]

Throughout the unsuccessful Rambouillet talks, the atrocities continued on the ground in Kosovo. In March, seven ethnic Albanians were killed and dozens injured in Kosovska, Mitrovica and Podujevo. Later that month, the peace talks resumed in Paris. Under the threat of abandonment by the international community, the Kosovar Albanian delegation agreed to the peace plan, which included broad autonomy and NATO enforcement troops. The Serb delegation, however, refused to sign the peace agreement.[153]

In a final attempt to settle the conflict, the Clinton administration sent Holbrooke to negotiate with Milosevic. The perceived consequences of such an approach was best articulated by the *Washington Post* editorial board, which declared,

---

149. *See id.*

150. The political backlash against the administration's equivocation began reaching a fevered pitch in the media. *See, e.g.*, Jim Hoagland, Editorial, . . . *Or a NATO Opportunity?* WASH. POST, Feb. 14, 1999, at B07. Mr. Hoagland stated:

> President Clinton and the other leaders can expect to be laughed off the. . . stage if Rambouillet fails and ethnic war is raging on Europe's ragged southern fringe while they solemnly party at the [April 23-25 NATO 50th] anniversary celebration. The road to a Washington summit that reflects glory on the good and great of the Atlantic community now passes through the police stations and city hall of the pitiable Kosovar capital of Prishtina. Stopping the bloodshed there and policing the peace is a mission Americans and Europeans should share equitably.

Id.

151. President William J. Clinton, The President's Radio Address, 1 PUB. PAPERS, 190, 190 (Feb. 13, 1999).

152. *See* Jane Perlez, *Albright Foresees No Raids on Serbia if 2 Sides Bar Pact*, N.Y. TIMES, Feb. 22, 1999, at A1; Thomas W. Lippman, *U.S. Could Join Force to Manage Postwar Kosovo,*, WASH. POST, Jan. 27, 1999, at A15.

153. *See* Carlotta Gall, *U.S. Official Sees 'Collision Course' in Kosovo Dispute*, N.Y. TIMES, Mar. 10, 1999, at A1.

> Now Holbrooke has returned to Belgrade for more talks with Milosevic. Since U.S. and European officials have all but assured the dictator that he no longer need fear a bombing campaign, there's reason to worry that Holbrooke will be offering rewards instead - assurances of immunity from war crimes prosecution for what's taken place in Kosovo so far, a lifting of economic sanctions, perhaps an increase in the already-too-high number of troops Milosevic would be permitted to station in Kosovo.[154]

Believing that the NATO threats were hollow, Milosevic opted to reject the agreement and the peace talks were suspended. With the refusal of Milosevic to accept a NATO peacekeeping force on Serbian soil, NAC authorized the NATO Secretary General to establish an air strike plan.

During this time, the Serb government, in preparation for a massive spring offensive, amassed one-third of all its forces in and around Kosovo, a fact that was well known to the international community.[155] The Serbian army then launched its campaign of terror forcing tens of thousands of ethnic Albanians to flee their homes. This escalation in ethnic aggression gave rise to calls for the Contact Group and for NATO to make good on their threat of force and seek to compel the Serbian regime's compliance with their demands. As described by the editorial board of the *Washington Post*,

> [T]owns are being burned, fathers executed in front of their children, thousands of people force-marched to unknown destinations, men separated en masse from women and children these are all to familiar indicators . . . . Mr. Milosevic has embarked on something close to genocide. NATO and President Clinton must not allow this to continue.[156]

Even more pointed were calls for a military response by several U.S. Senators, with Senator Joseph Lieberman for instance declaring, "we have been threatening [Milosevic] since Christmas of 1992 . . . warning him that if he attacked Kosovo, we would respond with force. Great alliances and great countries don't remain great if they issue threats and don't keep them."[157]

President Clinton responded to the increasing atrocities and the intense pressure with a perceived call to arms among the allies by declaring, "make no mistake, if we and our allies do not have the will to act, there will be more massacres. In dealing with aggressors in the Balkans, hesitation is a license to kill. But action and resolve can stop armies, and save

---

154. Editorial, *Modest Assaults Only, Please*, WASH. POST, Mar. 10, 1999, at A22.

155. *See* Sciolino & Bronner, *supra* note 101, at 12.

156. Editorial, *The Ground War*, WASH. POST, Mar. 28, 1999, at B06.

157. Jane Perlez, *Conflict in the Balkans:The Overview, Milosevic to Get One 'Last Chance' to Avoid Bombing*, N.Y. TIMES, Mar. 22, 1999, at A1.

lives."[158] Again, the strong rhetoric of President Clinton belied the actual policy commitment the United States and its European allies were willing to take in resolving the conflict. At this stage, the Americans and Europeans adopted a policy of air strikes without the possibility of ground troops.[159] In arriving at this strategy, the international community eschewed the use of ground troops as it was deemed a deal breaker with respect to the NATO members as it would likely shatter the allied consensus over the imposition of force.[160] Thus, with the issue of ground troops effectively stifled, the United States, along with its NATO allies, prepared to go to war.

More than any other factor, the public refusal to even consider or plan for the deployment of ground troops undermined the effectiveness of the air campaign and thereby increased the likelihood that ground troops would in fact have to be deployed in order to bring an end to Serbian aggression.

On March 24, 1999, NATO began its strategic air campaign against Serbian forces wherein it launched a series of increasingly intense air strikes against Yugoslav forces in Serbia and Kosovo, as well as Serbia's civil and military infrastructure. The purpose of the air campaign, in Clinton's words was to "demonstrate the seriousness of NATO's purpose so that the Serbian leaders understand the imperative of reversing course, [to] deter an even bloodier offensive against innocent civilians in Kosovo, and, if necessary, to seriously damage the Serb military's capacity to harm the people of Kosovo."[161]

Unfortunately the diplomatic deficit that had become such a common factor in the approach to the Yugoslav conflict substantially constrained the initial effectiveness of the air campaign. From the outset of the bombing campaign, the NATO member states had devised a strategic campaign based on their assessment that the war would only take a few days.[162] This assumption was largely based on the severe miscalculation on the part of the NATO member states that Milosevic would succumb to allied demands immediately following an allied air attack and would compromise just enough to avoid further NATO bombing. Based on this erroneous, the

---

158. *See* President William J. Clinton, News Conference on the Situation in Kosovo (Mar. 19, 1999), *at* http://abcnews.go.com/sections/us/Daily News/Clinton Transcripts.990319.html (last visited Feb. 23, 2002).

159. *See* President William J. Clinton, Address to the Nation on Aistrikes Against Serbian Targets in the Federal Republic of Yugoslavia (Serbia and Montenegro), 1 PUB. PAPERS 451, 452 (Mar. 24, 1999).

160. CLARK, *supra* note 31, at 166.

161. *Id.*

162. *See* Thomas W. Lippman, *Albright Misjudged Milosevic on Kosovo*, WASH. POST, Apr. 7, 1999, at A1.

international community was woefully unprepared for an extensive campaign against Serbia, with NATO committing only 366 planes, and approving only fifty-one targets.[163]

Throughout the air campaign, the effective use of force was hampered by a number of other constraining factors – most of which originated from America's European allies. Within a few days of the initiation of the campaign, the European allies, and some American officials sought a bombing pause so that negotiations could resume with Milosevic.[164] Of course, once the NATO bombing had been halted it might have been near impossible to resume the air campaign.[165] According to General Clark, strategic plans were also influenced by an "exaggerated fear of the Serbs' military capabilities, and the American military's reluctance to risk the Apache helicopters."[166] As a consequence, according to Clark, "The air campaign began with enough forces to punish the Serbs, but it lacked the mass and capabilities needed to halt the ethnic cleansing."[167] Moreover, the crucial element of surprise was often lost as a result of diplomatic maneuvering which informed Milosevic of impending targets. For instance the nature of the air campaign and a number of targets were provided to Milosevic during the October 1998 negotiations to establish the OSCE verification mission, as well as subsequent public discussion among NATO governments during the target approval process.[168] As with the case of Bosnia, espionage was also considered a possible key factor in the ability of the Serbs to anticipate NATO targets.[169]

With the failure of the initial campaign to bring a halt to Serbian aggression, the Americans sought to increase the number and range of targets. The objective being to deliver a series of knock-out blows to Milosevic and force him to the table. This approach was met by stiff resistance from Americas' allies, in particular the French, who believed that it was best to strike only secondary targets and "signal" to Milosevic that he had a lot to lose if the campaign continued.[170] This approach coincided with France's belief that Milosevic should remain in power after the end of the campaign, and that he was an individual with whom Europe could do business – both political and economic.

America's European allies also sought to dissuade NATO from exten-

163. CLARK, *supra* note 31, at 425.
164. *See id.* at 204, 424.
165. *See id.* at 205.
166. *Id.* at 424.
167. *Id.* at 424.
168. *See id.* at 424.
169. *See* CLARK, *supra* note 31, at 424-25.
170. *See id.* at 236-37.

sive strikes in Serbia, rather suggesting that NATO should limit its strikes to Kosovo and attempt to strike those units which were carrying out ethnic cleansing, and should avoid targets that might, "antagonize or damage Serbia further."[171] The French, in fact, argued that, "they wanted only 'reactive' strikes, not 'preventative' ones."[172] Having learned the lessons of Vietnam, America sought a more strategic approach, designed to go "after the heart of Milosevic's power."[173] Moreover, the American and European governments began to subject NATO targeting to political approval, with prime ministers and presidents deeply involved in selecting which targets could be bombed by NATO. The disagreements over strategy coupled with the process of political approval for targets led to an approach of incrementalism and according to Clark, "a growing perception that NATO wasn't committed to winning, a perception that was already undermining [NATO's] efforts."[174]

In an effort to enhance the effectiveness of the air campaign, and possibly avoid the need for a ground invasion, General Clark secured the deployment of US Army Apache helicopters to Albania. Unfortunately, as a result of internal political feuding within the Department of Defense, a miscalculation of casualty estimates, the reluctance of the US Army to become involved in the Kosovo conflict, a lack of understanding of the manner in which the Apaches could be used in support of the air campaign, and a "plain lack of knowledge" about the capabilities of the Apaches,[175] coupled with the absence of political leadership from the White House, the Apaches were never used.

With the perceived failure of the air campaign, and the reluctance to discuss the deployment of ground troops, concern increased that it may be necessary to return to a process of negotiation, thereby undermining NATO's deterrent capability. As recognized by the *Financial Times* editorial board,

> It is becoming painfully clear that bombing alone is not working . . . . NATO and above all President Bill Clinton's administration is guilty of a failure to use all necessary force. Mr. Clinton's prevarication about offering the U.S. troops that are vital to a successful outcome has left time on Mr. Milosevic's side. The window of opportunity is closing fast. . . . The alternative [is] a political settlement. There can be no pretense here that such a deal would be anything other than a terrible blow for the alliance.[176]

---

171. *Id.* at 237.

172. *Id.* at 256.

173. *Id.* at 237.

174. *Id.* at 253.

175. CLARK, *supra* note 31, at 289.

176. Editorial, *Time for Choice in Kosovo*, FIN. TIMES (LONDON), May 17, 1999, at 21.

Even other world leaders began to publicly worry. Jordan's King Abdullah was quoted as saying,

> As a friend of the United States, I think you have to be aware that the whole world is looking to see your commitment on this. If you step out of line and do this wrong, you will have a reaction . . . . There is a standard of values being addressed here, and your success or lack thereof will either encourage radicals or give them the sense that they have to tow the line. There is a lot riding on this. I don't think people fully understand the implications, whatever the final solution is in Kosovo.[177]

Given the failure of the limited air campaign to prevent Serbian atrocities against the Kosovar Albanian population, the United States came under increasing domestic and international pressure to consider the use of ground troops. On April 23, the NATO Summit took place in Washington, D.C. Although Secretary Cohen expressly forbad General Clark from raising the issue of ground troops at the Summit,[178] Prime Minister Tony Blair was more astute and recognized that by ruling out the credible threat of the use of ground troops the air campaign was unlikely to succeed. Prime Minister Blair thus privately pressured President Clinton to accept the inevitability of the need to prepare for the deployment of ground troops.

The immediate result of the Summit was the agreement to impose enhanced sanctions, including the blocking of all property and interests in property of Serbia/Montenegro, and a general ban on all exports to and imports from Serbia/Montenegro, including the export of petroleum and strategic goods.[179] Notably, until this ban France and other NATO member states had continued to sell oil to Serbia – even while their armed forces were bombing Serbia.[180]

As a result of the strong urging by Prime Minister Blair during the NATO meeting, the Clinton administration shifted its policy focus and created a "strategic campaign plan that combined military, economic, diplomatic, and other means to achieve core U.S. objectives."[181] This entailed an intensification of the air campaign and planning for the potential use of ground troops. It was finally becoming obvious to the United States and its NATO allies that the war in Kosovo was not a war that the United States and NATO could afford to lose.

---

177. Nora Boustany, *Jordan's King Says Stakes High for U.S. in Kosovo*, WASH. POST, May 18, 1999, at A19.

178. CLARK, *supra* note 31, at 269.

179. *See* FACT SHEET, NEW SANCTIONS AGAINST THE FEDERAL REPUBLIC OF YUGOSLAVIA (White House, Office of the Press Secretary, May 1, 1999); at http//: clinton6.nara.gov/1999/05/1999-05-01-fact-sheet-on-yugoslavia-sanction.html.

180. CLARK, *supra* note 31, at 259, 267.

181. DAALDER & O'HANLON, *supra* note 130, at 141.

Unfortunately, even though the White House had become increasingly committed to the possible use of ground troops, the Pentagon continued its efforts to undermine NATO consensus for the deployment of ground troops. Although it was becoming widely known that Secretary Cohen was the only remaining principle within the NATO mechanism ardently opposed to the use of ground troops, the Pentagon continued to report to President Clinton and others that there was no consensus in NATO for preparing for the deployment of ground troops. While this was technically correct, it was the United States that was blocking consensus. Similarly, the Pentagon relied on the lack of political guidance for the planning process and upon the natural inertia within the Pentagon for the eventual development of plans which indicated that the ground troops could not be effectively deployed until after the first snowfalls in Kosovo, thus delaying the deployment for a year.[182]

Finally, Prime Minister Blair, Secretary Albright, and General Clark were able to persuade President Clinton to embrace the need to become serious about planning for a ground deployment. After over seventy days of air strikes and an increasing willingness of NATO and the United States to commit ground troops to the conflict, the Russian government became involved in trying to negotiate a peace deal with Milosevic. The United States sent Strobe Talbott to negotiate agreeable terms for the cessation of the bombing campaign. The resulting agreement consisted of ten principles which Milosevic had to accept in their entirety. On June 4, 1999 the government of Serbia agreed to peace.[183]

Sadly, the coercive appeasement approach had become so ingrained within the international community that elements of the approach persisted long after the bombing campaign ended – this time with devastating consequences for the civilian Serb population residing in Kosovo. With a shift toward accommodating the interests of the Kosovars, and adopting the perspective that they were the victims in the conflict, the international community failed to foresee the likelihood that atrocities would be committed against the remaining Serb civilians. Given the atrocities which were committed throughout the conflict, it was clear that once the Serb army was forced to retreat under the NATO mandate, the Serb civilians would be subject to reprisals by certain elements of the Kosovar Albanian population. When this in fact happened, there was no immediate and serious NATO effort to protect the Serbs. Perhaps unconsciously, the international community failed to differentiate between the Serb army, which had perpetrated ethnic crimes and the Kosovar Serb population at large. Hav-

182. *See* CLARK, *supra* note 31, at 330-44.

183. *See* Daniel Williams, *Yugoslavs Yield to NATO Terms; Air Assault to Continue Pending Full Compliance* WASH. POST, June 4, 1999, at A1.

ing just defeated "the Serbs," NATO forces on the ground did not immediately possess the inclination to now "protect the Serbs." The approach of coercive appeasement also resulted in the marginalization of justice, whereby the Tribunal failed to indict any Kosovar Albanians for crimes committed against the Serbs residing in Kosovo, and again openly mulled over the uncertainty of its jurisdiction.

Marginalized Justice

Given its aversion to the use of force, and its recognition that a pure reliance on accommodation would likely be insufficient, the Contact Group, in March 1998 for the first time in the conflict, invoked the norm of justice. It first called upon the Serbian regime to "invite independent forensic experts to investigate the very serious allegations of extra judicial killings," and indicated that if the accusations were true, it expected Serbian authorities to prosecute and punish those responsible.[184] The Contact Group also asserted that its growing involvement in the issue was based on the members' "commitment to human rights values."[185] British Foreign Secretary Robin Cook, speaking on behalf of the United Kingdom and the E.U., explained that "serious violations of human rights, of civil liberties, of the freedom of political expression, are matters of concern to every member of the international community and cannot be regarded simply as an internal matter" for Serbia.[186]

The Contact Group further sought to motivate the Yugoslav Tribunal's Prosecutor, Judge Louise Arbour, by urging her to begin gathering information related to the violence in Kosovo, and reaffirmed that the Federal Republic of Yugoslavia (FRY) authorities have an obligation to cooperate with the Tribunal. Robin Cook further elaborated that the Contact Group had agreed the Prosecutor

> [S]hould consider the prosecution of anyone who may have committed a violation of humanitarian law in Kosovo. We are clear and she is clear that she has the legal authority to do that. We invite her to consider whether indictments might be appropriate in the light of the evidence of the past week.[187]

Cook further declared "there must be no impunity for those who break international law."[188] The Contact Group members also pledged to make

---

184. Contact Group Foreign Ministers, Press Statement on Kosovo (Mar. 9, 1998), *available at* http:// usembassy-australia.state.gov/hyper/wfa80309/cpf103.htm (last visited Feb. 23, 2002).

185. *Id.*

186. *See* Secretary of State Madeleine K. Albright and Contact Group Foreign Ministers, Press Conference (Mar. 9, 1998), *available at* http://secretary.state.gov/-www/statements/1998/980309a.html.

187. *Id.*

188. *Id.* In the question and answer session, British Foreign Minister Robin

available to the Tribunal "substantiated relevant information in their possession.[189]

In addition to the actions of the Contact Group, both the U.N. Security Council and the United States took a number of steps to more fully integrate the norm of justice. The Security Council urged the Prosecutor to "begin gathering information related to the violence in Kosovo that may fall within its jurisdiction,"[190] and created a committee to monitor the work of the Contact Group and to provide regular updates to the Council. In order to provide a constant flow of information to the Council concerning violations of international humanitarian law by the Serbian regime in Kosovo.[191] The U.S. also began distancing itself from the policy of moral equivalence[192] with Secretary Albright aggressively urging the Office of the Prosecutor to immediately begin investigations and for Serbia to allow the International Red Cross and UNHCHR to undertake their own investigations.[193] In order to prod the Prosecutor, the U.S. announced a contribution of $1 million to support the Tribunal's investigations in Kosovo.[194]

The emergence of the norm of justice was in large part a result of the belief that the approach of accommodation was failing, an implicit recognition that the approaches of economic inducement and the policy of diplomatic engagement were not serious, and a reluctance to commit to the use of force. More specifically, its invocation was based on the realization that there was no identified role for the use of force, and that greater involvement of justice based institutions (or their intentional exclusion by Milosevic) might provide a moral basis for the invocation of the use of force approach. However, despite the renewed employment of the norm of

---

Cook even declared, "we are insisting on the right of the international community to police international law and that means that we have a perfect right to express concern as we have done today at extra judicial killings and the death of eighty people without any trial or any judicial process." *Id.*

189. *Id.*

190. U.N. S.C. Res. 1160, U.N. SCOR, 54th Sess., 3868th mtg., para. 17, U.N. Doc. S/RES/1160 (1998).

191. *See id.*

192. *See* Secretary of State Madeleine K. Albright, Statement at the Contact Group Ministerial on Kosovo (Mar. 9, 1998), *available at* http://secretary.state.gov-/www/statements/1998/980309.html. (last visited Feb. 23, 2002) (declaring "we must avoid being paralyzed by the kind of artificial even-handedness that equates aggressors with their victims. We need to say clearly what is so clearly true: that responsibility for the violence lies squarely with Belgrade.")

193. *See id.*

194. *See* State Department Spokesman James Rubin, *Secretary Albright Announces $1.075 Million to Support Yugoslav War Crimes Tribunal Investigations in Kosovo* (Mar. 13, 1998), *at* http:// secretary.state.gov/www/briefings/Statements-/1998/ps980313.html. (last visited Feb. 23, 2002).

justice, the international community once again failed to fully support this approach.

With respect to the implementation of the norm of justice, there were several qualifications which limited its utility as a means for bringing the perpetrators of ethnic violence to justice. First, Secretary Albright expressly acknowledged that the Office of the Prosecutor's investigators were required to obtain visas from Serbia in order to investigate the alleged crimes, and thus provided Slobodan Milosevic with an effective veto over the ability of the Tribunal to have any meaningful role in stemming the atrocities. Second, the U.S. further limited the reach of the Tribunal by allowing the FRY to undertake responsibility for domestic prosecutions. Third, prominent U.S. Ambassadors were vocally reluctant to pursue the justice approach. This was evidenced in December 1998 when Ambassador Holbrooke and Ambassador Hill were asked whether they would comment on the fact that Milosevic and other Serbian officials had declared "that they don't have an intention to deliver any suspected war criminals to The Hague," Ambassador Holbrooke declined to take the question, while Ambassador Hill merely stated that "[t]he Hague Tribunal is the appropriate place for people who are accused of war crimes -- that it is the appropriate place, that it should be judged in The Hague."[195]

Thus, the limitations placed on the Tribunal by the international community and the public qualifications of the pursuit of justice seriously undermined its mandate. The Tribunal was given no authority to issue indictments or hold trials, nor was it viewed as having the full support of the international community. At the same time, the international community granted Milosevic the power to deny access to Tribunal investigators and to internally prosecute acts of war crimes in Kosovo, thereby severely hampering the overall effectiveness of the Tribunal itself and calling into question the resolve of the international community with respect to the application of the norm of justice.

The Contact Group's approach of minimal economic inducements, tepid diplomatic engagement, and calls for the greater application of justice during the winter of 1998 did little to dampen the conflict in Kosovo.[196] As acknowledged in June 1998 by Paddy Ashdown, leader of Britain's

---

195. Ambassadors Richard Holbrooke and Christopher Hill, Round-Table Discussion with Representatives of the Serbian Independent Media (Dec. 15, 1998), *at* http://www.state.gov/www/policy_remarks/1998/981215_holo_hill_rnd.html.

196. Commenting on the softening of U.S. policy toward Belgrade and the squandered opportunity to prevent the Kosovo conflict, a western diplomat stated in May 1998, "Milosevic has agreed only to talk not to strike an agreement on the future of Kosovo. The Americans have used our leverage to get the dialogue going, rather than to get somewhere." Steven Erlanger, *Allies Upset as U.S. Eases Stance on Kosovo*, N.Y. TIMES, May 28, 1998, at A12.

Liberal Democrats,

> [A]ll the blandishments of the international community, . . . all the sanctions, all the veiled threats have achieved absolutely nothing. There is no visible sign that any...actions we have so far taken have encouraged Milosevic to be in the slightest more restrained. He is now using tanks, heavy artillery the maximum weapons of war.[197]

As a result of the failure of the weak policy response and the apparent replay of the accommodation approach, which led to the prolonged Bosnia conflict, a number of prominent public officials called for a change in tactics. Czech President Vaclav Havel declared, "the earlier evil is confronted, the less we have to pay in human lives and suffering. Why for eleven years can one man expose the international community to ridicule? Why, when it was clear to any sensible person that Kosovo was bound to explode, do we have to wait for that explosion to happen?"[198] Similarly, former U.S. Ambassador Morton Abramowitz, declared "the United States and its allies have waited four months while Milosevic cleaned the clock of the Kosovo Liberation Army and [have] taken three weeks to discuss military action, with the result that 500 Albanian villages were destroyed."[199]

As a result of this increasing pressure and a growing understanding of the failure of the approach of coercive appeasement, a key senior State Department official met with Judge Arbour while she was visiting the United Nations for consultations and in very strong terms encouraged her to rapidly indict Milosevic. The Prosecutor was stunned as the United States had apparently changed its stance from one of urging caution in pursuing Milosevic and withholding sensitive material that might further the case against him, to an aggressive position favoring an immediate indictment.

At the meeting, the Prosecutor rebuffed the State Department plea as an infringement upon her impartiality. Shortly thereafter, on May 27, 1999, she indicted Mr. Milosevic, and five other top Yugoslav officials for war crimes and crimes against humanity committed in Kosovo from January 1999.[200] The indictment of Milosevic put a quick end to any remaining efforts at accommodation. As noted by former Senator Bob Dole, "the administration cannot reasonably expect the Kosovo Albanian deportees to return to Kosovo with an indicted war criminal and the very man who

---

197. *Voice of America*, June 5, 1998.

198. Robert Anderson, *Havel Urges Slovak EU Entry*, FIN. TIMES (LONDON), Oct. 19, 1998, at 2.

199. Jane Perlez, *Serb Pullback May Forestall NATO Attack*, N.Y. TIMES, Oct. 5, 1998, at A1.

200. *See Milosevic et al., Indictment*, IT-99-37 (June 29, 2001).

attempted to destroy them and their society as the guarantor of their security."[201]

According to General Clark, while the State Department was encouraged by this development, some Pentagon and White House officials were displeased with the indictment, on the grounds that it would limit or preclude their ability to negotiate with Milosevic, and that it would strengthen his resolve to proceed with the ethnic cleansing of Kosovo.[202] These concerns were reminiscent of those of Secretary Christopher when he reasoned that labeling the Serbian atrocities in Bosnia as genocide would create a moral imperative to use force and would limit the ability of the United States to negotiate with Milosevic. Although in the case of Kosovo, the threshold for the use of airpower had been crossed, some in the Clinton Administration were concerned that the indictment of Milosevic would require NATO to actually defeat Milosevic, rather than wound his regime and then rely upon him to guarantee a negotiated settlement as was the case in the autumn of 1995 and the subsequent Dayton negotiations.

Despite these fears, the indictment had the positive effect of strengthening the resolve of America's European partners in the NATO air campaign.[203] The indictment also provided a legitimate basis for the peace builders to call for a regime change, something they had previously been reluctant to do because it would be perceived as a violation of the FRY's sovereignty and political independence. As detailed by British Prime Minister Tony Blair, "the world cannot help you [Serbs] rebuild your country while Mr. Milosevic is at its head. And nor will the world understand, as the full extent of these atrocities is revealed, if you just turn a blind eye to the truth and pretend it is nothing to do with you. . . . This evil was carried out by your soldiers and by your leaders."[204] Moreover, the indictment provided a basis for calls from within the FRY for a regime change. Serbian Orthodox Bishop Artemije, for example, ventured that "there can be no solution under this regime, at least not a just and peaceful one. In Serbia as it is now, neither Serbs nor Albanians wish to live under this regime."[205]

---

201. Senator Robert Dole, *No Deals with Milosevic*, WASH. POST, June 1, 1999 at A15.

202. CLARK, *supra* note 31, at 325-26.

203. *See id.*

204. Gerald Baker et al., *NATO Leaders Urge Serbs to Oust Milosevic*, FIN. TIMES (LONDON), June 22, 1999, at 3.

205. Serbian Orthodox Bishop Artemije, ASSOCIATED PRESS, June 14, 1998.

B. The End of the Milosevic Reign of Terror: The Triumph of Diplomacy Backed by Force over Coercive Appeasement

In the aftermath of the 1999 NATO bombing campaign, Milosevic's local popularity dropped to just 20 percent, the lowest approval ratings in his thirteen-year rule.[206] Serbia's economic woes, brought on by years of international sanctions and exacerbated by the damage wrought by the NATO bombs, ultimately led to Milosevic's defeat in the Yugoslav presidential election on September 24, 2000.[207] Although he relinquished his official position, Milosevic continued to reside in the Presidential residence that had been his home for over a decade, and continued to serve as leader of the Yugoslav Socialist party, perhaps contemplating a way to return to power. For his part, the new Yugoslav President, Vojislav Kostunica, made clear that he had no intention of turning Milosevic or any other indicted Serbs over to the Yugoslav Tribunal. Despite this announcement, the United States and European Union lifted their economic sanctions on the FRY to signal their support for the new Yugoslav regime.[208]

As a result of Kostunica's defiance of the Tribunal, Senator Mitch McConnell instructed his staff member on the Senate Appropriations Sub-Committee on Foreign Operations, Robin Cleveland, to craft a provision

---

206. *See* Independent Center for Political Studies and Public Opinion Research at the Institute of Social Sciences, Belgrade (1999), *at* http://www.cpijm.org-yurprevious/sep2000.htm (last visited Feb. 23, 2002) (opinion poll).

207. *See* NATIONAL DEMOCRATIC INSTITUTE, SERBIA POST-ELECTION POLL (Oct. 6 2000), *at* http://www.accessdemocracy.org/NDI/library/1067.yu.poll/062000.pdf. When the Federal Election Commission nevertheless declared there would be a run-off between Mr. Kostunica and Milosevic, Mr. Kostunica vowed to boycott a second round, and thousands of Serb civilians took to the streets in Belgrade, Novi Sad, Nis, Kragujevac and Kraljevo in protest of Milosevic's perceived attempt to steal the election. As the throngs of protestors steadily grew outside the government offices in Belgrade, there was concern among the international community that Milosevic would quell the popular uprising by employing force, as he had done earlier. But rather than attack the protestors, the members of the local Serb army garrison joined them as they stormed various government buildings. Realizing that Milosevic's fate was now sealed, his confederates quickly scrambled for political cover. Subsequently, the Yugoslavia Constitutional Court ruled that Milosevic's opponent, Vojislav Kostunica, had in fact captured over fifty percent of the vote. This was followed from action by the Serbian parliament recognizing Kostunica's election as President. On October 7, 2000, Milosevic conceded defeat and turned the government over to Kostunica. *See* Steve Erlanger, *After Yugoslavs Celebrate, Belgrade Orders a Runoff*, N.Y. TIMES, Sept. 28, 2000 at A1.

208. *See* Michael Ignatieff, *The Right Trial for Milosevic*, N.Y. TIMES, Oct. 10, 2000, at A1.

conditioning further U.S. assistance to Serbia on compliance with its obligations under international law to the Tribunal. Together with a small group of NGO representatives including James Hooper of the Public International Law and Policy Group, Nina Bang-Jensen of the Coalition for International Justice, Susan Blaustein of the International Crisis Group, and John Fox, Cleveland crafted section 594(a) of the 2001 Appropriations Act. This paragraph provided that the United States should cease all assistance to Serbia and seek to block International Financial Institution lending if by March 31, 2001, the FRY had not provided sufficient cooperation with the Tribunal, including access for investigators, the provision of documents, and the surrender and transfer of indictees or assistance in their apprehension. The provision also required the FRY to take steps consistent with the Dayton Accords to end Serbian financial, political, security and other support which has served to maintain separate Republika Srpska institutions, and to take steps to implement policies which reflected a respect for minority rights and the rule of law. The Clinton Administration, and in particular the Department of State, aggressively lobbied for the removal of this conditionality provision from the Appropriations bill. But the provision was nevertheless enacted into law.

As March 31, 2001 approached, the new administration, led by President George W. Bush, began to signal to the Kostunica government that it was required to take at least some action to comply with the conditionality requirement. In response to continued attention to this issue by Senators McConnell and Leahy, the Department of State in mid-March prepared a demarche for Ambassador Bill Montgomery to deliver to President Kostunica, which expressly made the "arrest" of Milosevic a condition of continuing assistance. The U.S. did not specify on what grounds Milosevic was to be arrested or whether it was necessary that he be transferred to The Hague. The demarche also included a number of other actions the FRY was to take in order to comply with the legislation. Reportedly, the Ambassador ended up under-emphasizing the requirement of arrest and emphasized other more easily attainable conditions. As a result, President Kostunica sought to comply with the conditions by turning over to the Tribunal a Bosnian Serb, Milomir Stakic, the former mayor of Prijedor, who was wanted under a sealed indictment issued to the FRY in March 1997 for atrocities allegedly committed in 1992 and 1993 at the Omarska, Keraterm and Trnopolje camps in Bosnia.[209]

When on the advice of Ambassador Montgomery, the Department of State began to indicate that it would likely recommend that President Bush certify compliance and release the millions of dollars in aid funds despite

---

209. *See* Jane Perlez, *U.S. Set to Finesse Tribunal Issue and Allow Belgrade Aid*, N.Y. TIMES Mar. 29, 2001, at A3.

the FRY's failure to arrest Milosevic or take any other meaningful steps to abide by the conditions, Senator McConnell, backed by Senator Leahy publicly declared there had not been sufficient observance to warrant certification of compliance with the conditions and further assistance.[210] The Tribunals' Prosecutor, Carla del Ponte, also added her voice to the chorus calling for the withholding of aid unless Milosevic was arrested and transferred to The Hague.

On April 1, 2001 Milosevic was arrested by Serbian security forces and charged with corruption, political assassination, and election fraud. On April 3, Secretary of State Colin L. Powell certified the FRY's compliance with the conditions, but qualified the certification by noting that the United States support for up to $1 billion in additional international assistance to be pledged at a future international donors conference would be conditional on the FRY's full cooperation with the Tribunal. While much of the international community welcomed this development, the French objected to the use of economic conditionality to promote the arrest and the transfer to The Hague of Mr. Milosevic.

Many Serbian government officials and NGOs made the claim that it was necessary first to try Milosevic in Serbia for economic crimes committed against the people of Serbia as they were the most serious victims of his actions, and only then transfer him to The Hague to face charges of crimes against humanity and possibly genocide. The United States led a campaign to condition international assistance on his transfer to The Hague. But when Prosecutor del Ponte met with President Kostunica to work out arrangements for cooperating with the Tribunal, she found "he was absolutely [in denial], denying even the existence of the Tribunal — just accusing the Tribunal, because the Serbs are only victims."[211] During this time, half-hearted efforts of the FRY parliament to pass a law on extradition and cooperation with the Tribunal faltered, due in large part to the efforts of the pro-Milosevic Socialist People's Party from Montenegro -- which although delegitimized by the government of Montenegro, was frequently embraced and legitimized by Ambassador Montgomery and the Department of State.

As the June 29, 2001 deadline neared for the international donors conference in Brussels, all of the major European states agreed to attend despite the failure of the FRY government to act on the transfer of Milosevic to The Hague. In an interview with Newsweek Magazine, Prosecutor del Ponte responded to the question "[w]hat kind of pressure are the U.S. and other governments putting on Belgrade to hand him over?," by stating,

---

210. *See id.*

211. Roy Gutman & Daniel Klaidman, *Tracking War Criminals: Del Ponte Says NATO Troops are 'Doing Nothing'*, NEWSWEEK INT'L, May 21, 2001, at 22.

"the United States is giving very, very good support. I am very happy about that because the European Union support is low-level. The E.U. doesn't want to put conditions [on aid]. I am counting on the Americans. So aid must be conditioned? That is the only, only voice they hear . . . . I am sure it is the only way to obtain what we need to make justice."[212] In the face of this failure of European resolve, and against the advice of career officers at the State Department, but under renewed pressure from McConnell and Leahy and other important senators, Secretary Powell declared the United States would not be attending the conference unless Milosevic was extradited to The Hague.

On June 28, 2001, in the face of Kostunica's failure to comply with the conditions, Serbian Prime Minister Zoran Djindjic ordered the transfer of Milosevic to The Hague. While this act was welcomed by the international community, President Kostunica protested that the transfer of Milosevic was "illegal and unconstitutional," as well as "lawless and hasty."[213] Other members of Kostunica's cabinet resigned in dissent, throwing the fragile Belgrade government into turmoil. But the political crisis quickly faded when, two days later, Yugoslavia was awarded 1.28 billion dollars in aid by the United States and its European allies.

Milosevic, once America's and Europe's partner in peace, is now standing trial before the Yugoslav Tribunal for crimes against humanity and genocide. He has indicated that he intends to call numerous international officials, such as former British Foreign Minister Douglas Hurd, General Michael Rose, David Owen, Carl Bildt and Ambassador Richard Holbrooke to establish that they considered him essential to their efforts to bring peace to the former Yugoslavia and that they embraced him as a peace maker and not as a war criminal. While the testimony of these individuals and their explanations under cross examination would likely do little to actually support Milosevic's legal case, a probing of their diplomatic relationship with Milosevic will undoubtedly highlight the predominance of the approach of coercive appeasement and its devastating consequences for the people of Croatia, Bosnia, Kosovo and Serbia as well as the extent to which its use undermined the credibility of European and American foreign policy.

CONCLUSION

The war in Yugoslavia represented a unique opportunity for the interna-

---

212. *Id.*

213. Marlise Simons with Carlotta Gall, *The Handover of Milosevic: The Overview; Milosevic is Given to UN For Trial in War-Crime Case*, N.Y. TIMES, June 29, 2001, at A1.

tional community to act on its post cold war rhetoric of a new world order and the universal protection of human rights and human dignity. Instead, the story of the international response to the Serbian rogue regime represents a case study in failure and provides a very clear picture of the devastating consequences of failed efforts at peace building. The story is replete with missed opportunities, conflicted governments and institutional egotism, which prevented the international community from learning from its mistakes and taking the necessary action to defeat an evil regime bent on committing genocide against its former citizens. The result was a decade long war filled with countless atrocities, the collapse of institutional, economic and political structures, and a region whose future remains in doubt. Although both the Bosnian and Kosovo conflicts were eventually resolved through the use of force, the policy choices made by the international community that led to the approach of coercive appeasement have many reverberations. The status of Kosovo remains unresolved and continues to serve as a destabilizing factor in the region, the credibility of NATO and the U.N. is tarnished, and the ability of the international community to resolve such conflicts remains in doubt. What remains to be seen is whether the approach of coercive appeasement has become so ingrained as to preclude the adoption of other foreign policy approaches more suitable for solving conflicts. With the current war against terrorism and the ongoing crisis in the Middle East, the international community can ill afford to ignore the lessons of its failed approach to the Yugoslav conflict, and it must ever be on its guard not undertake flawed tactical decisions for short term gain which then commit it to a strategic approach of coercive appeasement.

The Bush Administration View of International Accountability

Ambassador Pierre-Richard Prosper*

Michael A. Newton**

INTRODUCTION

The President's decision to retain the Office of War Crimes Issues laid the foundation for the comprehensive and orchestrated diplomacy that today implements our overall war crimes policy. The Ambassador-at-Large for War Crimes deals every day with the conflicting demands required to address the most serious crimes perpetrated by man in a way that preserves and enhances the rule of law. These are far more than purely theoretical, esoteric legal issues. The crimes we seek to deter and punish are very real, yet almost unimaginable in their unyielding cruelty. The impact on victims is life-changing. Every policy and legal decision the United States makes is implemented against the backdrop of the legal traditions that have developed in pursuit of international peace and justice. The United States has maintained its responsibility for working with our closest allies to lead the charge in establishing mechanisms for addressing man's inhumanity to man. We are proud of the fact that we have helped

---

* U.S. Ambassador-at-Large for War Crimes Issues. Pierre-Richard Prosper was nominated by President Bush on May 16, 2001 to be the U.S. Ambassador-at-Large for War Crimes Issues and was sworn-in on Friday, July 13, 2001. In this position, he advises the Secretary of State directly on U.S. efforts to address serious violations of international humanitarian law, including genocide, crimes against humanity and war crimes, committed in areas of conflict throughout the world. Ambassador Prosper also coordinates U.S. support for the International Criminal Tribunals for the Former Yugoslavia and Rwanda, and assists in the creation and operation of other courts and judicial mechanisms to bring offenders of international humanitarian law to justice.

** Lieutenant Colonel Newton is the Senior Advisor to the United States Ambassador-at-Large for War Crimes Issues. He is involved with a wide range of issues related to U.S. policy regarding the law of armed conflict to include support to accountability mechanisms worldwide. He just returned from Sierra Leone where he served as the U.S. representative on the U.N. Planning Mission for the Sierra Leone Special Court. He previously served as a Special Advisor in the Office of War Crimes Issues from January 1999 to August 2000, and helped negotiate the Elements of Crimes document for the International Criminal Court.

the international legal community bear the intellectual and practical brunt of confronting some of the most difficult and painful problems humankind has faced over the past decades of conflict and turmoil.

Despite the difficulties, it is important to recognize that the world is indeed making some dramatic advances in achieving accountability for grave atrocities and war crimes. Leaders around the world can no longer expect to employ their might ruthlessly and remain above the reach of the law. Citizens worldwide are starting to feel that they are no longer at the mercy of forces of brutality, or that justice is nothing more than an unattainable abstraction. States that protect human rights and guard against war criminals are now becoming the norm. The rule of law is beginning to prevail over unremitting evil.

This, then, is a snapshot of the views of the United States on working together to pursue accountability for war crimes, particularly at a time when the world is experiencing new challenges.

SEPTEMBER 11, 2001

It is clear that September 11th has changed the world and the way we look at transnational threats and crimes. The tragic events of September 11th have forced us all to re-examine our traditional notions of security, our understanding of our attackers, and our approaches to bringing the perpetrators to justice. Moreover, September 11th gave the people of the United States a firsthand introduction into the pain and societal trauma that has unfortunately become so frequent in locations across our globe.

For the first time in modern history, mass atrocities were committed against the United States on American soil. The terrorists opened a new front in an unconventional war designed to destroy a way of life. Their goal was to shatter our fundamental beliefs, and test values we take for granted – freedom, peace, and the rule of law. These unprovoked attacks were more than isolated and sporadic acts of violence or other acts of a similar nature. September 11th truly represented a loss of American innocence. The crimes committed at the World Trade Center in New York, at the headquarters of our Department of Defense, and in Pennsylvania were of a magnitude that defied the imagination and shocked the conscience. These acts are unacceptable and will be punished. We are truly grateful for the many nations that have stood at our side in our common fight and pledged to use the full range of tools at their disposal.

The crimes of September 11th remind us that there must be accountability and the rule of law. The fundamental truth is that the rule of law defends civilization by erecting a wall of reason and humanity against savagery and brutality. To that end, and in the interests of preserving civilization across the world, the United States is committed more than ever to bringing perpetrators of war crimes and violators of international humani-

tarian law to justice both abroad and at home. We remain committed to working within the global community to confront the challenges that would destroy us all. We pledge to use our available resources and influence to promote the cause of freedom and liberty for all people.

Our views and our approaches may differ from some of those held by other nations, but we often seek common goals. For that reason, it is important to articulate our perspectives on the establishment of credible, effective accountability mechanisms to bring violators of international humanitarian law to justice and help build a global community that universally respects and upholds human rights.

THE INTERNATIONAL CRIMINAL COURT

In light of the terrorist threat that civilized nations face today, it may seem ironic to some that the vision for an International Criminal Court originated from the 1937 League of Nations Convention for the Prevention and Punishment of Terrorism.[1] The original Covenant never entered into force, though it began a meandering thread of thought and diplomatic discussion over nearly sixty years that led to the Rome Statute of the International Criminal Court.[2] The International Criminal Court (ICC) has been a point of concern for the United States.

This concern has not changed as a result of September 11th. Since the inception of the debates after World War II, the United States has persistently sought a court that would be neutral. We envisioned a court that could be focused on the pursuit of efficient justice, and most of all immune from the poisonous taint of raw political power. Because of its inherently flawed structure, the Bush Administration, as with the previous administration, opposes the Rome Treaty. Despite the signature by the Clinton Administration, we – like the previous administration – will not send it to the United States Senate for advice and consent to ratification.

We are steadfast in our concerns and committed to our beliefs that the United States cannot be part of a process that lacks the essential safeguards to avoid a politicization of the process. We also firmly believe that the ICC Treaty is just that -- a treaty. Therefore it cannot and should not

---

1. *See* Convention for the Prevention and Punishment of Terrorism, 19 LEAGUE OF NATIONS O.J. 23 (1938), League of Nations Doc. C.546(I).M.383(I).1937.V (1938) (Nov. 16, 1937), *reprinted in* M. CHERIF BASSOUNI, INTERNATIONAL TERRORISM: MULTILATERAL CONVENTIONS (1937-2001) 71, 79 (2001).

2. *See* Rome Statute of the International Criminal Court, U.N. Doc. A/CONF.183/9 (1998) (United Nations Diplomatic Conference of Plenipotentiaries on the Establishment of an International Criminal Court, July 17, 1998), *reprinted in* 37 I.L.M. 1002 (1998).

have jurisdiction over a non-party state absent United Nations Security Council action. The United States has a unique role in the world in helping to defend freedom and advance the cause of humanity. To that end, we will continue to meet our responsibility, but not at the price of our national security.

BUSH ADMINISTRATION PHILOSOPHICAL FRAMEWORK

This does not, however, undermine our steadfast determination to pursue accountability and justice on the world stage. To the contrary, we remain deeply committed. We have not abandoned our commitment to the principle of justice, nor will we shirk our obligation to represent the moral force of law-abiding nations in advancing respect for and adherence to the rule of law. We do have some purely philosophical differences with some of those who share our goal of pursuing accountability and justice for international crimes. Our approach begins on the domestic side of the ledger. As Ambassador Prosper testified to the United States Senate Committee on the Judiciary:

> the international practice should be to support sovereign states seeking justice domestically when it is feasible and would be credible . . . . International tribunals are not and should not be the courts of first redress, but of last resort. When domestic justice is not possible for egregious war crimes due to a failed state or a dysfunctional judicial system, the international community may through the Security Council or by consent step in on an ad hoc basis as in Rwanda and Yugoslavia . . . .[3]

In our view, the community of nations should use its means and influence to encourage states to pursue credible justice rather than simply abdicating domestic responsibility to international institutions. Because justice and the administration of justice are cornerstones of any democracy, states should be encouraged to exercise their sovereign right to pursue accountability for war crimes and other egregious violations of international and domestic law.

That said, the United States will continue to support the existing ad hoc international tribunals created under the authority of the Security Council.

We also see a continuing role for the United Nations in pursuing accountability. The United States remains proud of its leadership in working multilaterally to form the two ad hoc tribunals. They have both provided

---

3. *Review of Terrorism Suspects Policies, Hearing on DOJ Oversight: Preserving Our Freedoms While Defending Against Terrorism, Before the United States Senate Committee on the Judiciary*, 107th Congress (2001) (statement of Pierre-Richard Prosper, United States Ambassador-at-Large for War Crimes Issues), *available at* http://judiciary.senate.gov/hearing.cfm?id=129 (last visited April 18, 2002).

groundbreaking legal decisions and have sent the clear message that architects of genocide, crimes against humanity, and violations of the laws and customs of war will be held responsible for their crimes. While the work to date has greatly contributed to humanity, the long-term legacy of the Tribunals may be greatly enhanced over the next few years. In order to be deemed a success, they must attain a successful conclusion. And, in order to fulfill the spirit of the Security Council, they must begin to aggressively focus on the end-game.

FOUNDATIONAL PRINCIPLES OF THE AD HOC TRIBUNALS

The jurisdiction of the International Criminal Tribunal for the Former Yugoslavia (ICTY) and the International Criminal Tribunal for Rwanda (ICTR) are specifically limited to "serious" and "grave" violations of international humanitarian law. In establishing the Tribunals in U.N. Security Council Resolutions 827[4] and 955,[5] the Security Council noted its determination that they were a necessary response to threats to international peace and security. The two ad hoc tribunals were not established to judge each and every violation of law that occurred during the conflicts. And the tribunals were not designed to completely usurp the authority of sovereign states.

The Security Council clearly envisioned the fundamental responsibility of domestic courts for adjudicating some of these serious violations. The statutes specify that domestic courts have concurrent jurisdiction with the Tribunals for the "serious" violations that the Tribunals are empowered to prosecute. U.N. Security Council Resolution 955 calls for strengthening the courts and domestic judicial systems in Rwanda, "in particular to the necessity for those courts to deal with large numbers of suspects."[6] It is clear that the intention of the Security Council was for the tribunals to focus prosecution on those most culpable within the killing machines that committed the atrocities leaving the balance to the states.

The Tribunals should therefore focus on those most responsible as originally envisioned and, in regards to the ICTY, should conclude work by 2007, a date suggested by President Jorda. A hundred or more additional indictments may not be helpful and may undermine the pursuit of justice that may be better left to the regional states. We encourage ICTY to establish clear guidelines with regional states on which cases will be

---

4. *See* S.C. Res. 827, U.N. SCOR, 48th Sess., 3217th mtg., U.N. Doc. S/RES/827 (1993).

5. *See* S.C. Res. 955, U.N. SCOR, 49th Sess., 3453rd mtg., U.N. Doc. S/RES/955 (1994).

6. *Id.*

pursued by ICTY and which cases should be pursued in domestic courts. We also encourage a focus on management issues and improving oversight.

STATE RESPONSIBILITY

In order to achieve the shared vision for the successful culmination of the ad hoc tribunals, regional states must pick up the balance and not abdicate their responsibility. The abuses cannot go unpunished, and regional and local states cannot pretend that they did not occur. We will continue to press the governments in the former Yugoslavia to accept their responsibility, and are working with the government in Rwanda, to hold accountable the mid and lower level perpetrators. The lower level perpetrators in both of these regions should not get a free pass. We do not want to see an abandonment of state responsibility and therefore are encouraging appropriate domestic judicial and administrative action.

The United States also stands prepared to assist the states in rebuilding their shattered judicial systems to make them capable of dispensing truth-based justice and establishing systematic respect for the rule of law. As part of this commitment, we are exploring creative approaches such as the Rwandan *gacaca* system that is designed to deal with the seemingly untouchable mass of offenders.

For these efforts to be successful, however, cooperation by the states in the regions is essential. In regards to the ICTY, the Federal Republic of Yugoslavia must fulfill its obligations and transfer all at large indictees to The Hague. For our part, we want to be perfectly clear: the United States is committed to bringing Radovan Karadzic and Ratko Mladic into custody. Their trials will be a defining moment in the life of the ICTY and a landmark for history.

ICTR: U.S. POLICY – ENDGAME

We will continue to actively support the International Criminal Tribunal for Rwanda. We are increasing assistance, utilizing diplomacy, information collection and the existing Department of State rewards program, in locating and apprehending the remaining persons indicted for war crimes, including those who have killed, attacked, and threatened Americans. We urge other states to assist the Government of Rwanda in all possible ways, including by providing adequate resources. We are pressing states with indicted persons in their territory to turn those individuals over to the Tribunal for trial.

We also are encouraging the ICTR to establish clear guidelines with the Government of Rwanda relating to which cases will be pursued by the ICTR and which cases should be pursued in Rwandan domestic court. We

are stressing that the Rwandan government should continue to pursue justice against mid and lower ranking individuals through domestic processes, including *gacaca*. As with the ICTY, we are focused on efforts to improve the management of ICTR and recognize the positive steps ICTR has taken to date. We support the addition of *ad litem* judges to the ICTR to help accelerate the pace of trials, provided that the management of the ICTR is done in such a manner as to ensure the efficient employment of the *ad litem* judges.

HYBRID APPROACHES

In other areas of the world, such as Sierra Leone and Cambodia, we are working with the international community to pursue justice in each country. Using a hybrid approach, it is possible to more closely achieve our ultimate goal of every state fulfilling its responsibility to humankind to guard vigilantly against those who would commit war crimes, genocide, and crimes against humanity and to hold perpetrators accountable.

In Sierra Leone, the United States has worked diligently along with the international community to facilitate the establishment of a special court to hold accountable those who bear the greatest responsibility for the atrocities. This treaty-based court, with the significant involvement of the Sierra Leone Government as well as the international community, offers the promise of achieving credible justice. It is part of a strategy that will help build the future of the nation by fully rebuilding its legal structure while respecting sovereign responsibility. The United States supports the establishment of the Special Court, believing that it is one of several essential components necessary to restoring peace and stability to Sierra Leone and the region.

Similarly, in Cambodia, the United States supports efforts to bring to justice leaders of the Khmer Rouge who bear most responsibility for atrocities committed between 1975 and 1979. It is important that there be accountability in Cambodia in order to promote the rule of law and develop democracy. We have been encouraging both the Royal Government of Cambodia and the United Nations to be flexible in their approaches and to expeditiously finalize an agreement to ensure credible justice is achieved in the establishment of the Extraordinary Chambers. We are urging a prompt resolution of the remaining issues so that the long overdue process of justice and accountability can unfold. We look forward to U.N. negotiations with Phnom Penh to conclude this final stage in the process and hope they will take place soon.

AFGHANISTAN

We also face tremendous challenges in Afghanistan. To date, there are captured Taliban in the hands of other Afghan groups. We expect and have repeatedly made clear to opponents of the Taliban that captured fighters must not be harmed and must be treated humanely while they are being held. Afghan leaders, like Hamid Karzai, have said that the Taliban leadership will face justice for any crimes they have committed against the Afghan people. The international community should stand prepared to assist where needed.

PRESIDENT BUSH'S WAR CRIMES TRIBUNAL OPTION

Lastly, we want to address an issue that has generated much international media attention and legal commentary, and that is the Military Order that President Bush promulgated on November 13th. Shortly after September 11th, the President stated that he would use every tool at the disposal of the United States to combat terrorism and to bring the perpetrators of the atrocities to justice. He is indeed using the range of available military, diplomatic, economic, and legal tools. In creating the war crimes commissions, the President is fulfilling his pledge to the American people by creating another legal option to be used in appropriate cases. It is another arrow in the quiver.

One of the fundamental reasons the President invoked the military commission option is because he concluded that we are in a state of armed conflict.[7] It is clear that the attacks against the United States were more than isolated and sporadic acts of violence or other acts of a similar nature. A foreign, private terrorist network is waging war against the United States. It has organized, trained, and repeatedly carried out cowardly, indiscriminate attacks, including the largest attack in history against the territory of the United States and other attacks carried out on the soil of our allies.

An objective review of the organization's actions confirms the existence of a state of armed conflict. It begins with a decade's worth of hostile statements by Osama Bin Laden stating that he and all his followers are committed to waging war against the United States. It includes instruc-

7. "International terrorists, including members of al Qaida, have carried out attacks on United States diplomatic and military personnel and facilities abroad and on citizens and property within the United States on a scale that has created a state of armed conflict that requires the use of the United States Armed Forces." President Issues Military Order (Nov. 13, 2001) *available at* http://www.whitehouse.gov/news/releases/2001/11/20011113-27.html

tions to his followers to kill each and every American civilian. There is further evidence to suggest that al Qaida is connected with the bombing of the World Trade Center in 1993 as well as the attacks against U.S. military service personnel serving in Somalia in the same year. Added to it is the fact that Bin Laden and other members of al Qaida are indicted for attacking and bombing our embassies in Nairobi, Kenya and Dar es Salaam, Tanzania, and are believed to have perpetrated last year's bombing of the U.S.S. Cole. And of course, the evidence culminates with the horrifying and unprovoked air assaults on the Twin Towers in New York, the Pentagon, and the airplane tragedy in Pennsylvania.

The conduct of al Qaida is not merely a violation of criminal law – it also amounts to a violation of the laws and customs of war. Our view that this is a state of armed conflict is confirmed by the international reaction to the events. NATO's North Atlantic Council declared that the attack was directed from abroad and "regarded as an action covered by Article V of the Washington Treaty, which states that an armed attack against one or more of the Allies in Europe or North America shall be considered an attack against them all."[8] The Organization of American States, Australia and New Zealand likewise activated provisions in mutual defense treaties. U.N. Security Council resolutions recognized our inherent right to exercise self-defense and that acts of international terrorism constitute one of the most serious threats to international peace and security in the twenty-first century.[9]

Our Congress responded by passing a joint resolution authorizing "the use of all necessary and appropriate force" in order to prevent any future acts of international terrorism.[10] This, then, is a war, an unconventional war conducted by an evil aggressor, an aggressor that has no legal right to wage war. Under long established legal principles, the right to conduct armed conflict, lawful belligerency, is reserved to states and their armed forces or groups under responsible command that adhere to the laws and customs of war. Private persons lacking the basic indicia of organization and the ability or willingness to conduct operations in accordance with the laws of armed conflict have no legal right to wage warfare against a state. If they do so they are unlawful combatants.

The members of al Qaida do not meet the criteria to be lawful combatants under the law of war. They have no right to engage in armed conflict.

---

8. Statement by the Secretary General of NATO, Lord Robertson (Oct. 2, 2001) *available at* http://www.nato.int/docu/speech/2001/s011002a.htm.

9. *See* S.C. Res. 1368, U.N. SCOR, 56th Sess., 4370th mtg., U.N. Doc. S/RES/1368 (2001).

10. Authorization for Use of Military Force, P.L. 107-40, § 2 115 Stat. 224, 224 (2002).

And because they have intentionally targeted and killed civilians in the conduct of armed conflict they have committed war crimes. As a result, and in keeping with the U.S. legal tradition, war crimes commissions can be convened to adjudicate these horrific violations of the laws of war.

The power to create commissions is within the President's constitutional authority as Commander in Chief and is not new. War crimes commissions have been utilized and legally accepted throughout our history. They were used by General Winfield Scott during his operations in Mexico, in the Civil War by President Lincoln, and in 1942 by President Roosevelt.

And as many here know, war crimes tribunals are also an internationally accepted practice with deep historical roots. The international community has utilized war crimes commissions and tribunals to achieve justice, most notably at Nuremberg and in the Far East. Military commissions were used to try most of the leading perpetrators of Nazi and Japanese war crimes. These commissions were followed by thousands of Allied prosecutions of the lower-level perpetrators under the Control Council Law No. 10.[11]

By the end of 1958, the Western Allies had used war crimes tribunals to sentence over 5,000 Germans for war crimes.[12] In the Far East, 4,200 Japanese were convicted before war crimes tribunals convened by U.S., Australian, British, Chinese, Dutch, and French forces. Military commissions were also convened in Guam, Italy, Austria, and China.[13]

The commissions as envisioned by the President's Order, while different from those found in our civilian courts, conform to these historical precedents. Perhaps even more importantly, they adhere to the international law standards of today, including those embodied in the 1949 Geneva Conventions and the 1977 Additional Protocols to the Geneva Conventions. Although the commissions are conducted under a different body of law, they are also consistent with the International Covenant on Civil and Political Rights. They are also consistent with the process followed in the current ad hoc tribunals. Like its predecessors in Nuremberg, the Far East, and the International Criminal Tribunals for the former Yugoslavia and Rwanda, the commissioners may adjudge both issues of law and of fact. As with these prior institutions, decisions such as judicial orders, judgments, and sentences, will be reached by a majority vote and not unanimity, and evidence of a probative value will be admitted (although a

---

11. *See generally* Control Council Law No. 10 (Dec. 20, 1945) *reprinted in* TRIALS OF WAR CRIMINALS BEFORE THE NUREMBERG MILITARY TRIBUNALS (1949).

12. *See* Marjorie M. Whiteman, 11 DIGEST OF INTERNATIONAL LAW 947-56 (1968).

13. *See id.* at 1005.

death sentence will require a unanimous vote of the members of the commission). Proceedings will be open, but as with the United Nations International Criminal Tribunals for the former Yugoslavia and Rwanda, portions of the proceedings may be closed for reasons of safety or security.

In this campaign against terrorism, it is important that the President have the full range of available fora for seeking criminal accountability against persons for their individual and command responsibility for violations of the law of war. The war crimes commission provides a traditional mechanism to address these unconventional crimes.

There is much debate as to why we are not using our federal courts for all of these terrorists. Federal courts have not been excluded. Rather, the option of a war crimes commission has been included. A specialized tribunal, as is now international practice for the prosecution of war crimes, will not only bring in expertise with respect to the law of war, but will also provide security and protection – against persons who have proven that they are willing to kill thousands of people at a time – to the jurors, bailiffs, court reporters, and other personnel who did not sign up for this type of service, to be potential victims of terror.

We are doing what we believe is necessary to fight terrorist war criminals and to protect our country. We have a duty to future generations to remedy this evil. The American people are truly grateful to those nations that have stood at our side and joined us in this effort. These challenges will only be surmounted by the concerted efforts and unflinching dedication of the many distinguished representatives of the international community who have committed to solving the problem of international terrorism. This is a problem that reaches across borders and into the homes and hearts of every civilized nation. We are truly grateful for the outpouring of support that the world has demonstrated.

In the years ahead, the United States will continue to lead the fight to end impunity for genocide, crimes against humanity, and war crimes. We will be implacable in our opposition to those who feel that they are beyond the reach of the law as they commit horrendous crimes. We will help create the political will. We will continue to seek to bring justice as close as feasibly and credibly possible to the victims in order to create a sense of ownership and involvement. In our work with the Rwandan and Yugoslav Tribunals, the Special Court for Sierra Leone, and elsewhere, we will stress that all parties have a responsibility on the road to justice.

For this noble cause to be successful, for justice to endure, the international community, the existing ad hoc Tribunals, and the sovereign states of the world must coordinate, accept their role and individual responsibility, and go down this arduous road together. With our strong support of these efforts, we will continue to overcome obstacles, achieve accountability for the perpetrators, and secure the rule of law. In passing mile-

stones and creating an environment where there is not a dependency on international mechanisms we will bring justice to the victims and restore confidence in domestic institutions in societies throughout the world.

Playing Games with Terrorists

Abraham D. Sofaer*

It is a pleasure to be here at the New England School of Law to discuss the pros and cons of combating terrorism through criminal law enforcement. Professor Scharf and New England School of Law are to be congratulated on this meeting and on the series of which it is a part. The last conference, on the International Criminal Court, resulted in a superb collection of articles.

The case for using the criminal law to combat terror is stated by Professor Slaughter in her luncheon address.[1] I do not disagree with her regarding the use of criminal prosecutions; indeed I know of no one who would oppose criminal prosecution of terrorists.[2] The issue I address here is the degree to which criminal prosecutions had, until very recently, become virtually the sole method for attempting to punish and deter terrorist attacks.

For many years, between 1988 and September 11, 2001, presidents denounced terrorist attacks on Americans and promised to "pursue" attackers until they were "brought to justice." These turned out to be empty words. Rather than use the military to put an end to the groups we knew were responsible, presidents used the FBI to investigate, to develop evidence, and then to pursue through the criminal process those low-level operatives that we were fortunate enough to arrest. This policy of dealing with the terrorist threat allowed Osama bin Laden and his leadership, and the Taliban Government that gave them sanctuary, time to plan, prepare, and implement new attacks of increasing seriousness, with impunity.

Prosecutors and investigators were deservedly proud when, on May 31, 2001, a jury in New York ruled that four defendants were guilty of conspiring with Osama bin Laden to bomb American embassies in Nairobi, Kenya,

---

* George P. Shultz Distinguished Scholar & Senior Fellow, Hoover Institution, Stanford University; Professor of Law, by Courtesy, Stanford Law School. Office: Hoover Institution (HHMB 234), 434 Galvez Mall, Stanford, CA 94305-6010; TEL (650) 725-3763; FAX (650) 723-8583; email: sofaer@hoover.stanford.edu. Copyright © 2001, Abraham D. Sofaer, all rights reserved.

1. *See* Anne-Marie Slaughter, 36 New Eng. L. Rev. 681 (2002).

2. No international forum currently exists for such prosecutions. It would be helpful to have one, although certainly not the dangerous, planned International Criminal Court. *See* Abraham D. Sofaer & Paul R. Williams, *Doing Justice During Wartime: Why Military Tribunals Make Sense* 111 POLICY REVIEW 3 (Hoover Institution Feb./Mar. 2002).

and Dar es Salaam, Tanzania on August 7, 1998. The bombings killed 224 people, including 12 Americans, injured thousands, and caused property and other damage amounting to several hundred million dollars. Convicting responsible individuals establishes that they are subject to the rule of law; brings a sense of justice to survivors and relatives; lends credibility to the pledge that terrorists who kill Americans will be pursued until captured, prosecuted, and punished; leads to information helpful in capturing and prosecuting others, preventing future attacks, and exposing the aims and methods of the terrorist groups involved.

But the anti-terrorism policy of the Bush and Clinton administrations, based principally on criminal prosecution, created the misleading impression that the U.S. government was providing the American people with meaningful protection. In reality, that policy was uncomfortably like playing the computer game "Where in the World is Carmen San Diego." In that game, player investigators are assigned cases involving spectacular crimes -- for example stealing the Golden Gate Bridge -- committed by members of Carmen's extensive gang. Investigators then dash off to collect evidence -- flying from place to place -- based on clues received from witnesses and documents. Gradually, if successful, an investigator puts together enough evidence about the culprit to establish probable cause for an arrest. The investigator must then secure a warrant, and thereafter may arrest the correct person, and only that person, when he or she appears on the screen. Every successful arrest is a win, even though the boss -- the daring and elusive Carmen San Diego -- remains at large, arranging for other gang members to commit new crimes. Successful investigators are congratulated and promoted. According to an eleven-year-old Senior Investigator who lives in my home, it is possible to arrest Carmen San Diego herself, after all her numerous henchmen have been jailed; but he has yet to achieve that sublime result. Even after her arrest, a reincarnated Carmen San Diego appears, and the game goes on.

The criminal process has objectives, procedures, and results very much like those in the Carmen San Diego game. These contrast dramatically with the objectives, procedures, and results that would characterize an effort to deal with bin Laden and his conspirators as enemies of the U.S., and not merely criminals.

OBJECTIVES

The ultimate aim of criminal prosecution is to capture and deter those responsible for bin Laden's attacks. Necessarily, though, each prosecution is of specific individuals who are known to have perpetrated specific crimes. While bin Laden and some of his top associates were indicted along with the four actually tried for the bombings in Kenya and Tanzania, they could not be reached by the criminal process so long as they stayed in Afghani-

stan, which refused to extradite or prosecute them. By contrast, the effort to deal with bin Laden and his group as enemies has focused on bin Laden himself, and on every other member of Al Qaeda, regardless of their susceptibility to prosecution for particular crimes. The objective has not been to obtain convictions of particular Al Qaeda members, but to end the threat it poses.

PROCEDURES

Reliance on criminal prosecution to combat international terrorism, resulted in, after each bin Laden attack, an intensive criminal investigation, led by the FBI. Given its aim of arresting and prosecuting the actual perpetrators of each offense, FBI investigators sought, wherever they were sent, to find and develop usable evidence to identify individual suspects; to secure search and arrest warrants; and to extradite and prosecute those brought to trial. To use evidence at trial, moreover, prosecutors have to satisfy rules unrelated to reliability. And the need for usable evidence often leads prosecutors to offer deals to participants who are sometimes more culpable than those actually prosecuted.

An approach that treats bin Laden as an enemy of the U.S. also requires investigations, but it seeks primarily to determine who was ultimately responsible for attacks, rather than to build criminal cases against particular perpetrators. The test for evidence is its reliability, not its admissibility; and the strength and value of evidence is judged, not by any specific, legal standard, or by its potential for use in a trial, but by the extent to which the evidence justifies measures the U.S. considers taking to end the threat. Deals may be made with bin Laden associates, but to obtain cooperation in destroying his capacity to attack the US, rather than to develop evidence to successfully prosecute underlings.

RESULTS

Criminal prosecutions punish individuals for specific crimes, thereby achieving some measure of justice. This is an important objective in dealing even with a national security threat, and should not be abandoned. Punishing individuals, however, fails adequately to address the need to prevent acts of terror. Anti-terrorism efforts must ultimately be judged by whether they prevent attacks which can result in horrendous human and economic consequences. In some cases, criminal prosecutions may also deter terrorists from committing attacks; certainly, terrorists in prison are effectively prevented from participating in such conduct. But the deterrent effect of criminal prosecutions is limited, and in the most critical situations nonexistent. Criminal cases are generally possible in international terrorist attacks only against low-level conspirators, leaving leaders (like Carmen and Osama) undeterred. Prosecutions also usually take years to complete, and may take

place after additional attacks have been launched by the same group or even the same defendants. Criminal prosecution against bin Laden and his leadership are in fact no more effective than the fantasy prosecutions of Carmen San Diego's underlings.

Law enforcement activity cannot be expected, moreover, to shut down terrorist organizations operating in hostile and uncooperative states like Afghanistan. Nor are such states likely to be convinced by successful prosecutions of underlings to surrender leaders like bin Laden. The Afghan information ministry condemned the embassy bombing convictions as "unfair," and promised not to hand over bin Laden "under any circumstances," calling him "a great holy warrior of Islam and a great benefactor of the Afghan people." [3] The governments of several states openly, or more subtly, support the objectives and activities of various terrorist groups, just as the Taliban made clear it supported bin Laden's objectives. The only way to deal effectively with such situations is, as President Bush concluded, to treat the governments as responsible for the terrorist acts they allow to occur.

Criminal prosecutions are especially ineffective in deterring fundamentalist terrorist groups able to recruit individuals willing to sacrifice their lives in suicide bombings. These terrorists are not "cowards," as US officials sometimes assert; they are trained killers, as prepared for sacrifice as good soldiers. Some will cooperate rather than face life imprisonment or death; but Al Qaeda and similar organizations limit the damage any individual can inflict, by functioning through loose-knit groups.

The U.S. has at last adopted what former U.S. Secretary of State George P. Shultz called a pro-active approach, in order to prevent attacks, not merely to punish their perpetrators. To treat bin Laden as an enemy of the U.S. required more than the prosecution of perpetrators, or even freezing Al Qaeda's assets and imposing economic sanctions against Afghanistan. It required a shift in attitude from one of responding to terrorist attacks as if they were ordinary crimes to one of attempting directly to prevent terrorist groups from carrying out their threats. This shift requires more than just a single use of force in Afghanistan. It also requires taking security standards much more seriously, urgently developing defensive technologies, and, above all, using force sufficient to deter or prevent specific threats from all known sources. The President must order his staff to do all they can, through necessary and proportionate means, to end every known threat. His staff must be capable of implementing his orders, and should be held accountable for failing to do so.

Recent terrorist attacks have been successful due to a reckless indifference to security standards. The U.S. military housing complex, Khobar

3. *U.S. Embassy Verdicts Unfair, Taliban Says*, WASH. POST (May 31, 2001), at A17.

Towers, near Dhahran, Saudi Arabia, and the U.S. Embassy in Nairobi, Kenya, lacked essential perimeter protection. The *USS Cole* was successfully attacked on October 12, 2000 when inadequate security allowed an unknown vessel to come close enough to blow a hole in its hull and kill seventeen sailors. The World Trade Center bombing in 1993 might have been prevented if Arabic-language documents seized earlier from terrorists had been translated. The terrorists who destroyed the World Trade Center and part of the Pentagon were able to seize four of the four planes they targeted. Much more could have been done to prevent their virtually unmitigated success.

Lapses such as these cannot be tolerated if U.S. officials are truly to treat terrorist attacks as the national security threats they are. The President must do far more to hold individuals accountable for their failures in protecting the American people.[4] Moreover, the U.S. needs not only to prevent lapses, but also urgently needs to develop measures to reduce the risk of human error. The CIA and FBI have confirmed that Al Qaeda was attempting to obtain weapons of mass destruction, and that several Islamic states that tolerate or encourage anti-Western terrorism have already developed capacities that could lead to even more devastating attacks than those of September 11th, 2001. These threats must be taken on and ended when they become substantial and likely to be implemented.

When all else fails, force is the essential method of protection. Every recent U.S. president has properly concluded that necessary and proportionate force may lawfully be used in foreign territory to prevent attacks by terrorists given sanctuary there. Nor is it correct to claim that an attack aimed at Osama bin Laden would be an attempt to "assassinate" him. The concept of assassination as a limit on the use of force precludes acts that amount to murder; it does not limit the use of force in self-defense.

Resort to force, even when lawful, does require great care. Mistakes or excessive collateral damage can undermine the potential effectiveness of force. Furthermore, while the U.S. is free to act unilaterally in its self-defense, it must be prepared to defend its actions, or to admit and pay for its

---

4. It is simply wrong to suggest that the need to use force to prevent further attacks could not have been foreseen prior to September 11, 2001. For an excellent description, see Judith Miller, et al., *Planning for Terror but Failing to Act: Missed Signals*, N.Y. TIMES (Dec. 30, 2001), at 1. In an op-ed piece in 1995, I wrote:

> The criminal justice system is pathetically inadequate as a means for deterring state-sponsored terrorism. Individuals motivated to commit mass murder are fanatics, prepared to die in the act, or professionals, operating as instruments of major criminal groups or of state sponsors. They are easily replaced. To deter such acts requires direct pressure against the groups or governments responsible.

Abraham D. Sofaer, Editorial, *A Terrorist Defense Initiative*, WALL ST. J. (Sept. 11, 1995), at A16(W)/A20(E).

mistakes. President George H. W. Bush apologized for the destruction in 1988 of an Iranian passenger airliner by a U.S. Navy warship, killing all 290 people aboard, and negotiated a settlement with Iran. That is the proper way to deal with mistakes: admit them and accept the proper, legal consequences, so that the continued, potential use of force is not undermined or de-legitimized.

Ultimately, it is not any legal inhibition that accounts for the U.S. having for so long followed a policy of prosecuting rather than destroying international terrorist groups. Rather, it was an aversion by some officials and attorneys toward the use of force, even in self-defense. Additionally, it is the claim that the use of force leads to reprisals against the U.S., creating a cycle of violence, and makes terrorists like bin Laden heroes in the eyes of Muslim militants. People cite in this connection the bombing of Pan Am 103 by Libya allegedly in retaliation for the bombing years earlier of Qaddafi's headquarters. These arguments have done enough damage, and should be set aside once and for all.

First, the use of force against terrorists acting from within a state that gives them refuge is both lawful and moral, as the Security Council has now made clear.[5] The U.S. must insist on its right to defend its people and territory effectively, and when the use of force is necessary to achieve that end, it must use force rather than allow the American people to suffer the consequences. It is true that attacks on militant groups in foreign states may lead to further terrorist attacks. But it is also clear that militants who publicly commit themselves to attack U.S. targets are not going to be deterred by criminal prosecutions, and can only be stopped through effective action. The bombing of Tripoli was, in fact, effective, because it prevented Qadaffi from carrying out several attacks that he was planning at the time.[6] The key to success is not merely to attack, but to attack until the capacity of the group involved in committing acts of terror is destroyed. In no form of real warfare would a combatant expect an enemy to cease its attacks merely because of a single, largely symbolic use of force, such as President Clinton's strike against bin Laden in Afghanistan.

Furthermore, the claim that attacks on terrorists risk violent responses is equally applicable to criminal prosecutions. A criminal-law strategy is as likely to anger terrorists as attacks aimed at destroying them. Public trials equally serve to make them heroes in the eyes of militants.

Only in the world of computer games can we expect to avoid the real

---

5. *See* U.N. S. C. Res. 1368, U.N. SCOR 56th Sess., 4385th mtg., U.N. Doc. S/RES/1368 (2001); U.N. S. C. Res. 1373, U.N. SCOR 56th Sess., 4385th mtg., U.N. Doc. S/RES/1373 (2001); *see generally* Abraham D. Sofaer, *International Law and Kosovo*, 36 STANFORD J. INT'L L. 1 (2000).

6. *See* Abraham D. Sofaer, *Terrorism, the Law, and the National Defense*, 126 MIL. L. REV. 89, 103-105 (Fall 1989).

risks of confronting, or of failing to confront, an enemy. President Bush properly decided after September 11th that it was time to stop playing games and to give bin Laden and Al Qaeda the attention they deserved. As he put it, if we cannot bring them to justice, we will take justice to them. Clearly, that is what justice required.

The Case for an International Trial of the al-Qaeda and Taliban Perpetrators of the 9/11 Attacks

Michael P. Scharf*

INTRODUCTION

In the aftermath of the September 11, 2001 attacks on the World Trade Center and Pentagon, and the U.S. military action against Afghanistan, the U.S. Government has turned its attention to the question of what to do with the surviving al-Qaeda and Taliban leaders who were behind the terrorist operation. The Bush Administration has decided that perpetrators apprehended on American soil will be tried in federal district court,[1] while non-citizens captured abroad will be tried before U.S. military commissions.[2]

The Bush Administration's choice of these judicial fora has been rationalized on grounds of political realism: they would ensure optimum U.S. control over the proceedings, provide the greatest protection for intelligence sources and methods, and have the greatest chance of securing a conviction. But perhaps due to its mistrust of international institutions, the Bush Administration has overlooked the comparative real politick benefits of prosecuting the al-Qaeda and Taliban leaders before the existing Security Council-created International Criminal Tribunal for the Former Yugoslavia (ICTY),[3] as advocated by a growing number of

---

* Professor of Law and Director of the War Crimes Research Office at Case Western Reserve University School of Law, and formally Director of the Center for International Law & Policy, New England School of Law; former Attorney-Adviser for Law Enforcement and Intelligence (1989-1991) and Attorney-Adviser for U.N. Affairs (1991-1993), U.S. Department of State.

1. The first of two such trials involved Zacarius Moussaoui, a French citizen accused of involvement in planning the September 11th attacks, and Richard Reed, a British citizen deemed the "shoe bomb terrorist." *See* Gina Robinson, *Shoe Bomb Defendant Enters Plea of Innocent in U.S. Federal Trial*, FINANCIAL TIMES (London), Jan. 19, 2002 at 5.FN

2. Military Order of November 13, 2001, *Detention, Treatment, & Trial of Certain Non-Citizens in the War Against Terrorism*, 66 Fed. Reg. 57,833 (Nov. 16, 2001).

3. *See* S.C. Res. 827, U.N. SCOR, 48th Sess., 3217th mtg., U.N. Doc.

international law experts.[4] To pursue the ICTY option, the Security Council would merely have to adopt a resolution expanding the jurisdiction of the existing ICTY to cover the attacks of September 11, 2001. Because the September 11th attacks were a series of attacks against a civilian population which resulted in mass fatalities, they constitute "crimes against humanity" -- a crime already within the jurisdiction of the Tribunal.[5] By confining the temporal jurisdiction to a single day,

---

S/Res/827 (1993).

4. The author of this article was one of the first to propose the ICTY option, during a nationally televised debate with Professors Ruth Wedgwood and Anne Marie Slaughter on the Charlie Rose Show on October 26, 2001, and in an Op Ed piece which appeared in the Los Angeles Times on November 17, 2001. *See* Paul R. Williams & Michael P. Scharf, *Prosecute Terrorists on a World Stage*, L.A. TIMES, Nov. 18, 2001, at M5. Subsequently, a Committee of Experts convened by the Carnegie Corporation recommended the ICTY approach. *See* Public International Law and Policy Groups Peace-Building Program, *Report of the Select Committee of Experts on Nation Rebuilding In Afghanistan, available at* http://www.intermediatesovereignty.org/PUBS/afghnnation.htm (last visited Apr. 3, 2002). And a few weeks later, Carla Del Ponte, the Chief Prosecutor of the ICTY, went on record in favor of expanding the ICTY to cover the acts of September 11, 2001. *See* Nicole Winfield, *U.N. Urges International Trial for Bin Laden; U.S. Would Want Jurisdiction, Officials Say*, THE RECORD, Dec. 20, 2001, at A16.

5. *See* S.C. Res. 808, U.N. SCOR 48th Sess., 3175th mtg., U.N.Doc. S/Res/808. at 1 (1993). Crimes Against Humanity are defined as crimes "aimed at any civilian population and are prohibited regardless of whether they are committed in armed conflict, international or internal in character." *Id.* The ICTY Statute states that:

> The International Tribunal shall have the power to prosecute persons responsible for the following crimes when committed in armed conflict, whether international or internal in character, and directed against any civilian population: (a) murder; (b) extermination; (c) enslavement; (d) deportation; (e) imprisonment; (f) torture; (g) rape; (h) persecutions on political, racial and religious grounds; (i) other inhumane acts.

STATUTE OF THE INTERNATIONAL CRIMINAL TRIBUNAL FOR THE FORMER YUGOSLAVIA, art. 5 (annexed to S/RES/827 (1993) and as amended by S/RES/1166 (1998) and S/RES/1329 (2000)).

Note, the ICTY's definition of Crimes Against Humanity includes the requirement that the act be committed during an armed conflict. Customary International Law does not include the armed conflict requirement, which was inserted by the drafters of the ICTY Statute merely to reflect the situation in the former Yugoslavia. *See* Report of the Secretary-General Containing the Statute of the International Tribunal, 3 May 1993, *reproduced in*, VIRGINIA MORRIS & MICHAEL P. SCHARF, AN INSIDER'S GUIDE TO THE INTERNATIONAL CRIMINAL TRIBUNAL FOR THE FORMER YUGOSLAVIA 1 (1993). Whether the September 11th attacks were part of, or preceded, an armed conflict is a matter of contention. To avoid any question of the applicability of the Statute, the Security Council Resolution expanding the ICTY's jurisdiction could also stipulate: "For this purpose, the requirement that

September 11th, subsequent actions by U.S. forces in Afghanistan would not be subject to ICTY scrutiny.[6]

ADVANTAGES OF THE ICTY OPTION

There are six comparative advantages related to the ICTY option. The first of these is that the ICTY option can better promote President Bush's goals of establishing and maintaining an international coalition (including Islamic States) against terrorism. Trial before a U.S. jury or a panel of U.S. military judges is likely to be perceived, especially in the Islamic world, as inherently biased. The ICTY, in contrast, is made up of judges from every region of the world. Its bench includes several jurists from Islamic countries who, although unlikely to be the least bit sympathetic to al-Qaeda and the Taliban, would be perceived as ensuring an impartial bench by many in the Islamic world.[7] In addition, the ICTY Prosecutor, Carla Del Ponte, hails from Switzerland, a historically neutral country.

Yet, the ICTY is not some rogue international institution. Indeed, a majority of the ICTY's prosecutors and investigators are on loan from the U.S. Departments of Justice and Defense. More could be added to the ICTY staff to ensure sufficient resources to undertake the prosecution of al-Qaeda/Taliban leaders. While these staff members do not take their orders from the U.S. government, they certainly add an American flavor and sensibility to the institution.

Moreover, in contrast to the unwritten verdicts of federal juries and the cursory judgments of military commissions, the ICTY issues lengthy and detailed written judgments, comprehensively setting forth findings of fact and conclusions of law. The ICTY, therefore, is uniquely suited to establish a detailed historic record of al-Qaeda and Taliban culpability that is capable of piercing denial in the Islamic world and withstanding future revisionism. The Chief Prosecutor at Nuremberg, Supreme Court Justice Robert Jackson, underscored the logic of this proposition when he reported to President Truman that one of the most important legacies of the

---

there exist an armed conflict shall be inapplicable."

6. This would avoid a repeat of the situation that arose after the NATO bombing campaign in 1999, where the ICTY Prosecutor conducted a preliminary investigation into possible NATO war crimes. *See Final Report to the Prosecutor by the Committee Established to Review the NATO Bombing Campaign Against the Federal Republic of Yugoslavia*, May 8, 2000, at 57.

7. *See* U.N., International Criminal Tribunal for the Former Yugoslavia (ICTY) (April 12, 2002), *available at* http://www.un.org/icty/glance/index.htm (last visited April 19, 2002). The judges of the ICTY include Mohamed Shahabuddeen of Guyana, Mehmet Güney of Turkey and Amin El Mahdi of Egypt. *See id.*

international Nuremberg trial was that it documented the Nazi atrocities "with such authenticity and in such detail that there can be no responsible denial of these crimes in the future and no tradition of martyrdom of the Nazi leaders can arise among informed people."[8]

Second, the ICTY option is relatively fast and secure. There is an apt saying, "justice delayed is justice denied." That is because over time, evidence gets lost, memories fade, and witnesses disappear or die. In contrast to the proposed military commissions, as an existing institution the ICTY does not require the erection of a new courtroom, prison, and offices; the appointment of prosecutors and judges; nor the hiring of translators, guards, investigators, and support staff -- all of which will take many months, if not years, to complete. And unlike federal district courts, with the addition of two new panels of ad litem judges this fall, the ICTY does not have a lengthy back log of cases.

The ICTY could conduct the trial of al-Qaeda/Taliban leaders at nearby Camp Zeist, the retired American air base where the Pan Am 103 bombing trial was held in 2000-01. The specially designed Camp Zeist court facilities, located away from any population center, are both incredibly secure and high tech, making them an ideal venue in which to try the al-Qaeda/Taliban leaders.

Third, because of the unique body of law that the ICTY applies, the ICTY option actually presents a better chance of obtaining a conviction, thus up-ending one of the critical assumptions of the Bush Administration. Unlike the principles of conspiracy and aiding and abetting applicable in a U.S. federal court, the ICTY applies the principle of "command responsibility," under which a leader can be convicted for failing to prevent or punish the crimes of his subordinates.[9] This basis of culpability renders it much easier to convict leaders than under the federal standard, which requires proof that they ordered or directly participated in the crime. And unlike U.S. military commissions, the ICTY completely rejects the "obedience to orders defense."[10] Before a U.S. military commission, a defendant who was following an order will be acquitted unless the Prosecution can show that the order was patently unlawful.[11] An order to hijack a plane might be, but an order to open a bank account and deposit funds might

---

8. Report to the President from Justice Robert H. Jackson, Chief of Counsel for the United States in the Prosecution of Axis War Criminals, June 7, 1945, *reprinted in* 39 AM. J. INT'L L. 178, 184 (Supp. 1945).

9. *See* STATUTE OF THE INTERNATIONAL CRIMINAL TRIBUNAL FOR THE FORMER YUGOSLAVIA (annexed to S/RES/827 (1993) and as amended by S/RES/1166 (1998) and S/RES/1329 (2000)).

10. *See id.*; *see also*, Prosecutor v. Erdomovic, IT-96-22 (1998).

11. *See* Calley v. Callaway, 519 F.2d 184 (5th Cir. 1975) *cert. denied*, 425 U.S. 911 (1975).

not.

In addition, with the ICTY there is no risk that the trial verdict will be overturned upon a finding of juror exposure to prejudicial pre-trial publicity. Under the U.S. Supreme Court's precedent in the *Sheppard Case*,[12] it might prove difficult if not impossible to find a jury which has not been contaminated by the extraordinary amount of news coverage concerning persons accused of playing a role in the events of September 11th. The *Sheppard* rule does not apply to the ICTY's panel of professional judges.

Fourth, the ICTY may actually be able to better protect intelligence information than federal courts and U.S. military commissions. That is because the ICTY does not apply the *Brady* Rule,[13] which in U.S. proceedings permits the defense to undertake "fishing expeditions" for exculpatory information including CIA and NSA documents in the possession of the government. And contrary to popular belief, the *Brady* Rule trumps the Classified Information Protection Act,[14] at least with respect to information a judge deems relevant to the defense after an in camera inspection. The ICTY, in contrast, operates under an MOU between the U.S. Government and the Office of the Prosecutor, under which intelligence information that is not used in court may not be disclosed to the defense, whether or not it may be exculpatory.

Fifth, the ICTY provides a greater chance of securing custody over offenders, at least with respect to persons who surface in countries other than Afghanistan or the United States. The recent stance of Spain is indicative of the problems the United States is likely to face. Spanish authorities said they would not extradite to the United States an al-Qaeda operative found in Spain who was accused of playing a major role in the September 11th attacks, unless Spain received an assurance that the United States would not apply the death penalty and would not try the accused in a military commission.[15] Other states where al-Qaeda/Taliban perpetrators surface are likely to follow Spain's example in opposing extradition to the United States for trial before a military commission, and Islamic States may oppose their extradition to the United States under any circumstances.

In this respect, the situation is analogous to that which existed in the

---

12. *See* Sheppard v. Maxwell, 384 U.S. 333 (1966).

13. *See* Brady v. Maryland, 373 U.S. 83 (1963).

14. *See* 18 U.S.C. App.3 §§ 1-16 (1994) (classified information disclosure in federal trials); 42 U.S.C. § 2014 (1994); *see also* Saul M. Pilchen & Benjamin B. Klubes, *Using the Classified Information Procedures Act in Criminal Cases: A Primer for Defense Counsel*, 31 AM. CRIM. L. REV. 191 (1994); Debra Stafford, *"Secret Trials:" A Defense Perspective*, 1988 ARMY LAW. 24 (1988).

15. *See* T.R. Reid, *Allies Oppose Death Penalty & Bush's Plan for Secret Military Tribunals*, WASH. POST, Nov. 29, 2001, at A23.

aftermath of the 1994 genocide in Rwanda. States where the Hutu perpetrators of the genocide had taken refuge refused to extradite them to the new Tutsi-dominated government in Rwanda for fear that they would not receive a fair trial. But these same states were ultimately perfectly willing to extradite the Hutus to the International Criminal Tribunal for Rwanda, which was established by the Security Council in 1994.[16] Thus, pursuing the ICTY option would increase the chances of bringing al-Qaeda/Taliban leaders to justice.

Finally, trying al-Qaeda/Taliban leaders before military commissions will undermine the protection of U.S. citizens abroad. The United States traditionally protests the trial of American citizens before foreign military courts that do not fully comport to international notions of due process. The well-known case of Lori Beronson is illustrative. Beronson was a young American journalist, who was tried and convicted of conspiring with the Tupac Amaru Revolutionary Movement by a secret military court in Peru, headed by a hooded judge.[17] American authorities successfully pressured the Peruvian government to give her a new trial before a civilian court on the grounds that the trial before the military court violated her human rights. The State Department's Annual Human Rights Report is filled with similar cases.[18] The United States will not be able to bring moral pressure to bear on such countries in support of American citizens if the United States, itself, has embraced military commissions that do not provide the same rights as applied in ordinary courts.

In contrast, the ICTY Statute incorporates the rights of the accused required by the International Covenant on Civil and Political Rights.[19] American support for a trial of al-Qaeda leaders before the ICTY would reaffirm American commitment to the rule of law, set a positive example for the rebuilding of the Afghan judicial system, and enable the United States to protect its citizens from unfair trial abroad.

CONCLUSION

President Bush has said that he wants the al-Qaeda and Taliban leaders responsible for the September 11th attacks "dead or alive" and that they should either be brought to justice or "justice should be brought to

---

16. *See* Michael P. Scharf, *Responding to Rwanda: Accountability Mechanisms in the Aftermath of Genocide*, 52 COLUM. J. INT'L AFF. 621, 628-29(1999).

17. *See* MICHAEL P. SCHARF, THE LAW OF INTERNATIONAL ORGANIZATIONS 686-89 (2001).

18. *See* U.S. Department of State, Bureau of Democracy, Human Rights & Labor, HUMAN RIGHTS REPORT (2000), *available at* http://www.state.gov/g/drl/rls/hrrpt/2000/ (last visited April 19, 2002).

19. *See* International Covenant on Civil and Political Rights, March 23, 1976, art 14, U.N. G.A. Res. 2200 (1976).

them."[20] Most people assume, therefore, that the highest level perpetrators will be extinguished militarily, and that there will be no need for trials. But it is becoming increasingly clear that Osama bin Laden and several of his lieutenants have escaped through the net American forces cast around Afghanistan, and are likely to surface in other countries. Thus, the debate over where best to try such individuals is not simply an academic exercise.

In the context of the crisis in the Balkans, the ICTY has demonstrated how international tribunals can be a particularly effective foreign policy tool. The indictments and proceedings before the ICTY, which led to the isolation and removal of rogue leaders like Radovan Karadzic and Slobodan Milosevic, generated international support for economic sanctions and the use of force against Serbia, and have produced an internationally accepted historic record of war crimes and genocide in the former Yugoslavia.

Despite the ICTY's successes, the Bush Administration has not fully considered the comparative advantages of employing the ICTY as a venue for trying the al-Qaeda and Taliban leaders responsible for the September 11th attacks. The Bush Administration appears to have self-consciously shut its eyes to the benefits of the ICTY approach as a consequence of its opposition to the Permanent International Criminal Court.[21] A final argument, that might have currency with the Bush Administration, is that by embracing an expansion of the ICTY's jurisdiction to cover the crimes against humanity committed on September 11, 2001, the Administration could begin the transformation of the ICTY into a permanent Security Council-controlled international criminal tribunal, which the Bush Administration might consider more useful and less threatening to the United States than the permanent International Criminal Court.

---

20. Vernon Loeb, *U.S. Forces to Stay in Afghanistan for a While, Bush Says, Noting "Dead or Alive is Fine With Me," President Bush Asserts Bin Laden is Not Escaping*, WASH. POST, Dec. 29, 2001, at A8.

21. *See* Michael P. Scharf, *The United States and the International Criminal Court: A Recommendation for the Bush Administration*, ILSA J. INT'L & COMP. L. 385-89 (2001).

Accountability Without Hypocrisy: Consistent Standards, Honest History

Beth Stephens*

My specialty within the field of international law is accountability, the study and development of mechanisms to hold accountable those who violate international human rights standards. As a sub-specialty, I have worked for many years on a particular approach to accountability: civil lawsuits in U.S. federal court against individuals, corporations and governments responsible for egregious human rights abuses. This paper is an edited version of a talk I gave at a symposium discussing responses to rogue states. My assigned topic was the use of civil litigation as a means of holding such states accountable for their human rights abuses.

The events of September 11, 2001 starkly reshaped our symposium. Most of the participants talked about terrorism -- its sources and sponsors, the legal regimes governing international and domestic responses to terrorism, and the use of force to combat it. Terrorism was my starting point as well. But my views about accountability for such crimes are the product of my own experiences with terrorism, experiences that have led me to a very different framework than most of those who spoke at the symposium. As a foundation for my discussion of civil remedies for terrorism and other human rights abuses, let me first explain my own close encounters with terrorism and terrorists of very different origins.

TERRORISM IN NEW YORK, TERRORISM IN NICARAGUA

Living in a suburb of New York City, I was deeply affected by the September 11th attacks. A relative of mine worked on a top floor of one of the World Trade Towers. Within minutes of the attacks I watched tele-

---

* Beth Stephens is an Associate Professor at Rutgers University School of Law-Camden and a cooperating attorney with the Center for Constitutional Rights. She has participated in the civil lawsuits described in this article in many roles, including plaintiff and field investigator as well as attorney, and written extensively on related issues. Professor Stephens lived in Nicaragua from 1983-89 and worked at the Nicaraguan government's National Commission for the Protection and Promotion of Human Rights. This is an edited version of the remarks she presented at the New England School of Law symposium, Responding to Rogue States: From Smart Bombs to Smart Sanctions, on November 9, 2001.

vised pictures of her floor engulfed in flames. I soon learned of an extraordinary stroke of luck: she was late to work that morning, standing on the street outside the Towers when a plane hit her office. Almost 200 of her co-workers died in the attack, but she survived, physically unharmed.

Such stories are all too common around New York City. Everyone I know was touched in some way. We all know people who lost relatives. Countless were saved by near misses. Many of us saw the buildings collapse. Others were displaced from homes or offices. Most still suffer the after-effects of these traumatic events.

September 11th, however, was the *second* time that terrorism has closely impacted my life. The first was in the 1980's when I lived in Nicaragua, and the state sponsor of the terrorism was my own government, the United States. Supported by President Reagan and hundreds of millions of dollars of U.S. government funding, the counterrevolutionaries (the "contra") fighting to overthrow Nicaragua's Sandinista government murdered, tortured and kidnapped civilians -- peasants, teachers, engineers, doctors, nurses, many of them my friends and neighbors.[1]

In April 1987, the terror struck my inner circle of friends. I first heard an unconfirmed report that something had happened to Ben Linder, a U.S. civil engineer designing a mini-hydroelectric plant in a mountain village in northern Nicaragua. I joined a sobbing gathering of friends as we learned the horrifying details: Ben and a group of Nicaraguans were ambushed while building a dam to measure water flow through a mountain stream. I later investigated his murder on behalf of his parents.[2] After being wounded in the legs, Ben was shot at close range, a bullet to his forehead. I met the struggling widows and children of the two Nicaraguans killed with him, one stabbed in the heart after being wounded by gunfire.

Dozens of similar accounts haunt my memories: the coffee grower tortured and murdered because he insisted on continuing to harvest his crop despite "contra" threats; the terrified fishermen who survived an attack with minor wounds by diving into the water off their boats; families with young children sleeping on the floor of a rudimentary school because their homes had been burned; teachers and doctors forced to leave their jobs in

---

1. For a fully documented description of "contra" human rights abuses, see, *e.g.*, REED BRODY, CONTRA TERROR IN NICARAGUA: REPORT OF A FACT-FINDING MISSION, SEPTEMBER 1984-JANUARY 1985 (1985).

2. The Linder investigation contributed to the filing of a lawsuit against the "contra" leaders. *See* Linder v. Calero Portocarrero, 963 F.2d 332 (11th Cir. 1992). I was also a plaintiff in a lawsuit seeking implementation of a decision of the International Court of Justice that held that the U.S. attacks on Nicaragua violated international law. *See* Committee of U.S. Citizens Living in Nicaragua v. Reagan, 859 F.2d 929 (D.C. Cir. 1988).

rural areas in the face of death threats.

Perhaps my most chilling memory is of an interview with a "contra" commander who cheerfully described torture and murder of civilians, all with a smile on his face -- a killing machine trained and paid by my own government. The Nicaraguan government also committed human rights violations, but the vast number of horrific abuses were committed by the forces funded, trained, organized and directed by the government of the United States. The war inflicted tens of thousands of casualties -- and most of their blood is on our hands.

I feel deeply for those who lost loved ones on September 11th. I cry when I read their stories and see their faces -- bewildered children, young women struggling to raise families alone, older widows and widowers facing life without their longtime partners. But I also cry when I remember the close Nicaraguan friend whose sister was killed by a bomb made in the USA. I still flinch at the memory of the colleague whose son was shot in the head by the "contra": he survived but lost much of his memory and all of his emotional control. He became an angry, often frightening dependent in his parents' home. I carry with me the anger of Ben Linder's father, who died knowing that U.S. government officials responsible for the murder of his son were never punished for their crimes.

Nicaragua in the 1980's was hardly the first or last time that the United States government has employed the tactics of terror in the effort to overthrow a government it considers unfriendly, to install a government more to its liking or to protect an unpopular government from removal by its own people: Iran and Guatemala in the 1950's, the Congo in the 1960's, East Timor and Chile in the 1970's -- the examples are numerous and well-documented.[3] The numbers of casualties mount into the hundreds of thousands, if not millions, but very few are remembered in this country. We have wiped them from our collective memory, from our history of ourselves.

I approached our symposium topic, responding to rogue regimes, marked by this understanding of history. Other speakers at our symposium recited a list of terrorist attacks aimed at U.S. military personnel and civilians. They appeared to view the United States as a generous, meek giant, fighting terrorists with both hands tied behind our back. But when I look at world events over the past fifty years, I do not see a gentle giant suffering from decades of unredressed attacks. The very different picture I see is the one seen by many people around the world: the United States as a big

---

3. One fascinating source is the collection of declassified U.S. government documents available on-line from the National Security Archive, George Washington University, *available at* http://www.gwu.edu/~nsarchiv/ (last visited Feb. 19, 2002).

bully, throwing our weight around the world. We have received an occasional punch, none more vicious than that on September 11th, and we have done much good, but we have also terrorized large swaths of the world through decades of military interventions and support for terrorist regimes and organizations.

Entering into this conversation with my friends and neighbors after September 11th, several said to me in anger, "But are you claiming that this history of U.S. wrongdoing justifies what was done to us?" Of course not. But it does, it should, affect the way in which we understand terrorism and our responses to it. The battle against terrorism is not one of pitting good people (us, and those who stand with us) against evil people (all those who attack us or do not join us in our fight). Terrorism is a tool of combat that has been employed all too frequently by far too many individuals, organizations and governments, including our own.

To wage an effective campaign against terrorism, we must first acknowledge that we too have stooped to employ it. We must support mechanisms to prevent our own use of terrorism in the future, to ensure that we stay on the side of the good. This requires applying a consistent, even-handed definition of terrorism to ourselves as well as to others -- forgoing the hypocrisy of supporting terrorism when it serves our interests, while waging war against it when we are the target. This clearheaded objectivity is sorely lacking in the U.S. response to September 11th.

One key step would be to join multilateral efforts to promote accountability through the rule of law, to empower alternatives to the unilateral use of force. With this goal, the application of civil remedies for human rights abuses is one contribution to strengthening the role of international law in response to human rights abuses around the world.

CIVIL REMEDIES FOR HUMAN RIGHTS ABUSES

The international community has worked on many fronts over the past fifty years to construct peaceful responses to violations of international law. These efforts have led to the elaboration of detailed norms governing international human rights and terrorism. Most international law before World War II had governed interactions among states and their governments. The Nazi atrocities spurred two crucial developments. First, international norms now govern a state's treatment of its own citizens. Abuses such as genocide and torture are prohibited by international law when committed within a state's own territory as well as across national lines. Second, international law increasingly addresses the misconduct of private individuals and organizations, rather than just official actions. Genocide is prohibited whether committed by government officials or by private actors. Many terrorist acts -- hijacking, hostage-taking, use of explosives --

are international crimes prohibited for private as well as state actors.

Enforcement of these international rules remains spotty. International mechanisms created over the past decades are still at their early stages of development. Ad hoc tribunals have prosecuted dozens of individuals for international crimes committed in the former Yugoslavia and Rwanda. The permanent International Criminal Court that will soon come into being -- without the participation of the United States -- will be empowered to bring to trial both citizens of member states and those who have committed designated international crimes within their territory.

As with most rules of international law, much of the enforcement of human rights and anti-terrorism norms must take place within domestic legal systems. Many nations have begun domestic criminal prosecutions of individuals for international law violations.[4] Some involve crimes committed within the territory of the forum state or against its citizens. Other prosecutions, however, rest upon the principle of universal jurisdiction. Universal jurisdiction permits, or at times requires, a state to hold accountable those responsible for certain egregious violations of international law even where the offense has no relationship to its territory or citizens. Criminal prosecutions have been initiated throughout Europe and in one African state, Senegal.[5]

In the United States domestic legal system, accountability for international violations has been implemented through civil litigation rather than criminal prosecutions.[6] This reflects in part a reluctance of our federal

---

4. The most well-publicized involved the effort to extradite former-Chilean General Augusto Pinochet from England to Spain to face criminal charges arising out of the military dictatorship in Chile. For general discussions of the Pinochet extradition battle, see Michael Byers, *The Law and Politics of the Pinochet Case*, 10 DUKE J. COMP. & INT'L L. 415 (2000); Naomi Roht-Arriaza, *The Pinochet Precedent and Universal Jurisdiction*, 35 NEW ENG. L. REV. 311 (2001).

5. See cases summarized in INT'L LAW ASS'N, FINAL REPORT ON THE EXERCISE OF JURISDICTION IN RESPECT OF GROSS HUMAN RIGHTS OFFENCES (2000) and in Beth Stephens, *Translating Filártiga: A Comparative and International Law Analysis of Domestic Remedies for International Human Rights Violations*, 27 YALE J. INT'L L. 1, 37 n.160 (2002).

6. A U.S. criminal statute defines torture committed outside of the United States as a federal crime. 18 U.S.C. § 2340a (2002) (imposing penalties for "[w]hoever outside the United States commits or attempts to commit torture"). Despite this statute, the U.S. government has refused to initiate prosecutions, even though human rights groups have presented information to the government about potential defendants present in the United States. *See* Beth Van Schaack, *In Defense of Civil Redress: The Domestic Enforcement of Human Rights Norms in the Context of the Proposed Hague Judgments Convention*, 42 HARV. INT'L L.J. 141, 148-49 (2001) (citing interview with Gerald Gray, Executive Director, Center for Justice & Accountability on June 5, 2000, discussing U.S. government refusals to

government to undertake such prosecutions. But it is also a product of characteristics of the U.S. legal system that render civil lawsuits a more viable and attractive option here. These include principles of jurisdiction and choice of law that make such lawsuits possible, as well as a legal and political culture that supports private use of the legal system to impact upon national policies.[7]

U.S. civil lawsuits for human rights violations originated as the latest application of a U.S. statute, the Alien Tort Claims Act (ATCA), which states, "[t]he district courts shall have original jurisdiction of any civil action by an alien for a tort only, committed in violation of the law of nations . . . ."[8] Originally enacted in 1789,[9] the statute was largely ignored for 200 years.[10] The modern line of ATCA human rights cases began in 1980 with the Second Circuit decision in *Filártiga v. Peña-Irala*.[11] Joelito Filártiga, a seventeen-year-old Paraguayan, was tortured to death in Paraguay in 1976 by Americo Norberto Peña-Irala, a Paraguayan police officer, in retaliation for his father's political opposition to the military dictatorship. Peña-Irala fled Paraguay in 1978 and moved to New York City; the following year, the Filártiga family sued him in federal district court seeking damages for Joelito's torture and death. The Second Circuit ruled that the lawsuit fell within the reach of the ATCA, holding that the statute incorporates the law of nations in its current form and that modern international law prohibits human rights abuses committed by government officials against their own citizens. The court concluded that "official torture is now prohibited by the law of nations" and therefore triggered jurisdiction under the ATCA.[12]

Since 1980, a series of decisions has developed both the list of human rights abuses addressed by the statute and the categories of defendants

---

initiate prosecutions). *See also* Coletta A. Youngers, *The Pinochet Ricochet*, NATION, May 8, 2000, at 5 (discussing U.S. government refusal to initiate torture prosecution).

7. For a full discussion of the procedural, structural and cultural explanations for the use of civil remedies in the United States, see Stephens, *supra* note 5, at 10-17.

8. 28 U.S.C. § 1350 (2002).

9. Judiciary Act of 1789, ch. 20, § 9(b), 1 Stat. 73, 77.

10. For the history of ATCA claims prior to *Filártiga*, see Kenneth C. Randall, *Federal Jurisdiction Over International Law Claims: Inquiries into the Alien Tort Claims Statute*, 18 N.Y.U. J. INT'L L. & POL. 1, 4-7 nn.15-17 (1985) (Twenty-one cases claimed ATCA jurisdiction prior to *Filártiga*; only two sustained jurisdiction).

11. 630 F.2d 876 (2d Cir. 1980). For the full factual background, see *id.* at 878-80.

12. *Id.* at 884.

subject to suit. Cases have addressed claims such as genocide, war crimes, summary execution, disappearance, and arbitrary detention, as well as torture; claims have been upheld when the court finds that the alleged tort satisfies the international law definition of a universally prohibited offense.[13] The cases have also upheld claims against commanding officers, officials of de facto states, U.S. and local government officials, and corporations.[14] The courts look to international law to define who can be held liable for the violation. Since the international law definition of torture requires some public action, the cases have limited torture claims to those that allege that the defendant acted under color of official authority.[15] Since genocide is defined as an international violation when committed by private persons, claims of genocide have been accepted even in the absence of state action.[16]

Despite academic debates about the intent of the ATCA,[17] all courts that

---

13. *See, e.g.*, Kadic v. Karadzic, 70 F.3d 232 (2d Cir. 1995), *cert. denied*, 518 U.S. 1005 (1996) (genocide, war crimes, summary execution, torture); Abebe-Jira v. Negewo, 72 F.3d 844 (11th Cir. 1996), *cert. denied*, 519 U.S. 830 (1996) (torture); Trajano v. Marcos, 978 F.2d 493 (9th Cir. 1992), *cert. denied*, 508 U.S. 972 (1993) (summary execution); Xuncax v. Gramajo, 886 F. Supp. 162 (D. Mass. 1995) (summary execution, torture, disappearance, cruel, inhuman or degrading treatment); Forti v. Suarez-Mason, 672 F. Supp. 1531 (N.D. Cal. 1987), *aff'd on reh'g*, 694 F. Supp. 707 (N.D. Cal. 1988) (summary execution, torture, disappearance).

14. *See, e.g.*, Wiwa v. Royal Dutch Petroleum Company, 226 F.3d 88 (2d Cir. 2000), *cert. denied*, 532 U.S. 941 (2001) (corporation); Martinez v. City of Los Angeles, 141 F.3d 1373 (9th Cir. 1998) (local government officials); Kadic v. Karadzic, 70 F.3d 232 (head of de facto state); *In re* Estate of Ferdinand E. Marcos Human Rights Litg., 978 F.2d 493 (commanding officer/former head of state); Jama v. U.S. Immigration and Naturalization Serv., 22 F. Supp. 2d 353 (D.N.J. 1998) (U.S. officials and private corporation).

15. Torture is defined by the Convention Against Torture as certain acts committed "by or at the instigation of or with the consent or acquiescence of a public official or other person acting in an official capacity." Convention Against Torture and Other Cruel, Inhuman or Degrading Treatment or Punishment, Dec. 10, 1984, art. 1(1), G.A. Res. 39/46, 39 U.N. GAOR, Supp. No. 51, at 197, U.N. Doc. A/39/51 (1984) (entered into force for the United States Nov. 20, 1994). *See Kadic*, 70 F.3d at 243 (torture proscribed by international law "only when committed by state officials or under color of law").

16. *See* Convention on the Prevention and Punishment of the Crime of Genocide, Dec. 9, 1948, art. 4, 102 Stat. 3045, 78 U.N.T.S. 277 (entered into force for the United States Feb. 23, 1989) ("Persons committing genocide . . . shall be punished, whether they are constitutionally responsible rulers, public officials or private individuals."); *Kadic*, 70 F.3d at 241-42 (holding that ATCA applies to a claim of genocide committed by a private individual).

17. *Compare* Beth Stephens, *Federalism and Foreign Affairs: Congress's Power to "Define and Punish . . . Offenses Against the Law of Nations*," 42 WM. &

have decided the issue have concluded that the statute authorizes suits in federal court by aliens for torts in violation of international law, no matter where committed. For example, in *Trajano v. Marcos*, the Ninth Circuit stated, "[w]e start with the face of the statute. It requires a claim by an alien, a tort, and a violation of international law."[18] The Supreme Court has never addressed the issue, but has repeatedly denied petitions seeking review of the issue.[19] A modern statute, the Torture Victim Protection Act (TVPA), creates a right to sue for torture or extrajudicial execution, extending that right to sue to U.S. citizens as well as aliens.[20] The legislative report accompanying the TVPA stresses its support for the application of the ATCA to international human rights cases.[21]

Civil litigation can play an important role in the search for accountability for human rights violations. Litigation that leads to an enforceable money judgment offers the plaintiff compensation for the injuries inflicted by the defendant, punishes those responsible for those harms, and serves as a deterrent to future abuses. Even if collection of the judgment is impossible, civil lawsuits still serve the accountability drive, offering benefits to the victim, to the human rights movement, and to society as a whole.[22] Many such lawsuits lead to judicial recognition of developing

---

MARY L. REV. 447, 520-24, 520-21 n.305 (2000) (citing cases and additional articles) *and* Beth Stephens, *The Law of Our Land: Customary International Law as Federal Law After Erie*, 66 FORDHAM L. REV. 393 (1997) (discussing role of customary international law in debate over ATCA), *with* Curtis A. Bradley & Jack L. Goldsmith, *Federal Courts and the Incorporation of International Law*, 111 HARV. L. REV. 2260 (1998) *and* Curtis A. Bradley & Jack L. Goldsmith, *Customary International Law as Federal Common Law: A Critique of the Modern Position*, 110 HARV. L. REV. 815 (1997).

18. 978 F.2d at 499.

19. *See supra* notes 13-14 and accompanying text. Even if this interpretation of the ATCA is rejected by the Supreme Court or lower courts, civil litigation for human rights abuses committed abroad could be litigated in the United States pursuant to the more recent statutes discussed in what follows.

20. The Torture Victim Protection Act, 28 U.S.C. § 1350 note (2002). The TVPA is limited by the requirement that the torture or extrajudicial execution be committed "under actual or apparent authority, or color of law, or any foreign nation." *Id at* § 2(a).

21. *See* H.R. REP. NO. 102-367, at 3 (1991), *reprinted in* 1992 U.S.C.C.A.N. 84, 86 (noting that the ATCA "has . . . important uses and should not be replaced").

22. *See* BETH STEPHENS & MICHAEL RATNER, INTERNATIONAL HUMAN RIGHTS LITIGATION IN U.S. COURTS 233-38 (1996) (explaining value of civil remedies):

> [T]he plaintiffs in these cases are concerned about much more than money. They take tremendous personal satisfaction from filing a lawsuit, forcing the defendant to answer in court or to abandon the United States, and creating an official record of the human rights abuses in-

rules of law, creating a precedent for future policy development. The judicial proceeding may produce a full factual investigation, identifying those responsible and leading to public recognition of victim's injuries and the defendant's culpability. The defendant may be punished by this public exposure as well as by legal consequences such as deportation or the denial of visas.[23] The cases assist the human rights movements in the countries where the abuses took place, offering a focus for education and organizing about human rights violations. They also help educate people in the United States about international law, human rights and accountability. Although criminal prosecutions may entail more severe penalties for the defendant, civil remedies play an important role where criminal prosecution is not possible, and complement criminal prosecutions where they take place.

One of the reasons civil litigation is attractive in the United States is the possibility of punitive damages awards. Such judgments serve as a kind of community condemnation of the defendant, even where they cannot be collected. Punitive damages express society's revulsion for the defendant's conduct, and thus resemble criminal penalties. "[P]unitive damages advance the interests of punishment and deterrence, which are also among the interests advanced by the criminal law . . . ."[24] United States civil litigation thus has much in common with criminal prosecutions in other legal systems, particularly systems in which criminal prosecutions can be initiated by private parties.[25]

---

flicted on them or their families.

Id. at 234. *See also* Jose E. Alvarez, *Rush to Closure: Lessons of the Tadic Judgment*, 96 MICH. L. REV. 2031, 2101-04 (1998) (explaining value of civil remedies); Beth Stephens, *Conceptualizing Violence Under International Law: Do Tort Remedies Fit the Crime?*, 60 ALBANY L. REV. 579, 603-05 (1997) (same).

23. For example, the defendant in one U.S. case, Hector Gramajo, lost his visa from the United States as a result of the judgment in Xuncax v. Gramajo, 886 F. Supp. 162 (D. Mass. 1995) *Guatemalan ex-army general barred from United States*, REUTERS, Sept. 13, 1995.

24. Browning-Ferris Industries of Vermont, Inc. v. Kalco Disposal, Inc., 492 U.S. 257, 275 (1989). "The Court's cases abound with the recognition of the penal nature of punitive damages." *Id.* at 297 (O'Connor, J., dissenting). As the Court said in Beckwith v. Bean, "[t]he principle of permitting damages, in certain cases, to go beyond naked compensation, is for example, and [for]the punishment of the guilty party for the wicked, corrupt, and malignant motive and design which prompted him to the wrongful act." 98 U.S. 266, 277 (1878).

25. *See* Stephens, *supra* note 5, at 18-21, 44-46.

An early effort to bring a claim of terrorism under the ATCA was unsuccessful. Although the three-judge panel produced three splintered opinions,[26] even the judge most supportive of human rights litigation rejected the terrorism claim:

> [T]he nations of the world are so divisively split on the legitimacy of such [terrorist] aggression as to make it impossible to pinpoint an area of harmony or consensus. . . . Given this division, I do not believe that under current law terrorist attacks amount to law of nations violations.[27]

The problem identified by Judge Edwards -- lack of consensus about what constitutes illegitimate terrorism -- persists despite intensive worldwide attention to the issue. However, it may be possible to assert a terrorism claim under the ATCA by arguing that consensus exists as to subparts of the international violation.[28] Applying this approach, at a minimum, terrorism claims probably can be based upon hijacking, hostage-taking, bombings, and other offenses condemned by existing international conventions. In addition, a future case might argue that a consensus exists as to a larger subset of terrorism. Indeed, Professor Harold Koh argued fifteen years ago that the terrorism claim in *Tel-Oren* could have been sustained had the judges recognized that the acts alleged in the complaint fell within a core of universally condemned conduct.[29] Koh defines this core to encompass an "organized and deliberate attack upon innocent civilians without a collateral military target . . . ."[30] Where the facts of a case fall within such a core, it might be possible to litigate an ATCA case under the rubric of terrorism.

As part of an anti-terrorism initiative, Congress has also authorized civil suits for victims of terrorism pursuant to a separate statute.[31] The provision authorizes a U.S. national "injured in his or her person, property, or business by reason of an act of international terrorism" to sue for treble

---

26. *See* Tel-Oren v. Libyan Arab Republic, 726 F.2d 774, 775-98 (Edwards, J., concurring), 798-823 (Bork, J., concurring), 823-27 (Robb, J., concurring) (D.C. Cir. 1984), *cert. denied*, 470 U.S. 1003 (1985).

27. *Id.* at 795 (Edwards, J., concurring).

28. Applied to cruel, inhuman or degrading treatment, this approach led a court to find that a consensus as to the exact parameters of the tort was unnecessary where consensus existed as to the particular conduct at issue in that case. *See* Xuncax v. Gramajo, 886 F. Supp. 162, 187 (D. Mass. 1995).

29. *See* Harold H. Koh, *Civil Remedies for Uncivil Wrongs: Combating Terrorism through Transnational Public Law Litigation*, 22 TEX. INT'L L.J. 169, 205 n.116 (1987). *Id.*

30. *Id.*

damages.[32] "[I]nternational terrorism" is defined by statute as violent or dangerous acts that would constitute crimes if committed within the United States, that "appear to be intended (i) to intimidate or coerce a civilian population; (ii) to influence the policy of a government by intimidation or coercion; or (iii) to affect the conduct of a government by assassination or kidnapping," and that occur outside of the United States or "transcend national boundaries."[33] In two reported cases, courts have upheld jurisdiction under this provision over claims of terrorism filed by relatives of U.S. citizens killed in Israel.[34]

CIVIL LITIGATION AND ROGUE STATES

Each of these statutes has been used primarily against individuals, associations and corporations, not states. These uses reflect the impact of U.S. Foreign Sovereign Immunities Act, which affords foreign states immunity from suit unless a claim falls within a list of enumerated exceptions.[35] The broadest exceptions permit suit for torts committed within the United States and for certain commercial matters. In an important limitation, the Supreme Court held that ATCA claims do not create an automatic exception to the statutory definition of foreign sovereign immunity; that is, such suits are not permitted unless they fall within one of the standard statutory exceptions.[36]

---

31. 18 U.S.C. §2333 (2002)

32. 18 U.S.C. § 2333(a) (2002).

33. 18 U.S.C. § 2331(1)(a)-(c) (2002).

34. *See* Estate of Ungar v. Palestinian Auth., 153 F. Supp. 2d 76, 100 (D.R.I. 2001) (upholding subject matter jurisdiction over claims alleging that defendants were complicit in shooting death of U.S. citizen in Israel); Boim v. Quranic Literacy Institute, 127 F. Supp. 2d 1002, 1021 (N.D. Ill. 2000) (denying motion to dismiss claim by parents of man killed in Israel who alleged that defendants had engaged in international terrorism by supplying funding to groups responsible for the attack).

35. The Foreign Sovereign Immunities Act (FSIA) is codified at 28 U.S.C. §§ 1330, 1602-1611 (2002). The exceptions are set forth in section 1605(a)(1)-(7).

36. *See* Argentine Republic v. Amerada Hess Shipping Corp., 488 U.S. 428, 439 (1989) (refusing to recognize an FSIA exception for claims of gross human rights abuses) *But see* Princz v. Federal Republic of Germany, 26 F.3d 1166, 1176-84 (D.C. Cir. 1994) (Wald, J., dissenting) (arguing that the FSIA should not be read to bar civil suits for violations of universally accepted international norms). The FSIA does permit human rights lawsuits to proceed where the claims fall within the statutory exceptions. *See, e.g.*, Siderman de Blake v. Republic of Argentina, 965 F.2d 699, 723 (9th Cir. 1992) (applying exceptions for waiver of immunity, extraterritorial torts having an impact in the United States, and expropriation of property); Letelier v. Republic of Chile, 488 F. Supp. 665, 674 (D.D.C. 1980) (applying tort exception to claim arising out of assassination of political

The English courts have reached a similar conclusion, protecting the sovereign immunity of the government of Kuwait in a suit alleging responsibility for torture.[37] This decision was recently upheld by the European Court of Human Rights in a 9-8 decision finding that the grant of immunity did not violate the European Convention's guarantee of a right of access to the courts.[38] The dissenters argued that all governments have an international law obligation to provide access to their courts to adjudicate claims of torture, an obligation that cannot be evaded by the application of domestic law statutory immunities.

Litigation against sovereign states may be possible where the defendant government is not recognized by the United States: the FSIA does not apply if the United States government does not support the defendant's claim to be a sovereign state. A lawsuit against the Taliban, for example, might have been possible despite the FSIA, given that neither the United States nor the United Nations recognized the Taliban as the government of Afghanistan.

In addition, a 1996 amendment to the FSIA permits civil suits for torture, extrajudicial killing and other abuses against a small group of foreign governments.[39] The new exception permits suits by U.S. citizens when the defendant government has been designated by the U.S. government as a "state sponsor of terrorism."[40] The State Department has designated seven such countries,[41] and judgments have been issued pursuant to the statute against both Cuba and Iran.[42] The statute renders some abusive regimes subject to suit for human rights abuses, an important goal, but it does so through non-uniform standards, setting the United States as the sole arbiter of which governments should be held accountable and which should

---

opponent within the United States).

37. *See* Al-Adsani v. Government of Kuwait, 107 I.L.R. 536 (Eng. C.A. 1996).

38. *See* Al-Adsani v. United Kingdom, Application No. 35763/97 (Nov. 21, 2001) *available at* http://hudoc.echr.coe.int/ (visited February 21, 2002).

39. 28 U.S.C. § 1605(a)(7) (2002 Supp. V).

40. 28 U.S.C. §§ 1605(a)(7)(A), (B)(ii) (2002 Supp. V).

41. The foreign States designated as sponsors of terrorism for the purposes of this statute are Iran, Cuba, Syria, Iraq, Libya, Sudan, and North Korea. 22 C.F.R. § 126.1(d) (2001).

42. *See, e.g.*, Elahi v. Islamic Republic of Iran, 124 F. Supp. 2d 97, 100 (D.D.C. 2000); Flatow v. Islamic Republic of Iran, 999 F. Supp. 1, 12-13 (D.D.C. 1998); Alejandre v. Cuba, 996 F. Supp. 1239, 1247-48 (S.D. Fla. 1997). Cases have also been filed against Iraq and Libya. *See, e.g.*, Rein v. Socialist People's Libyan Arab Jamahiriya, 162 F.3d 748, 754 (2d Cir. 1998) (denying defendant's motion to dismiss); Daliberti v. Republic of Iraq, 97 F. Supp. 2d 38, 40 (D.D.C. 2000) (same).

benefit from immunity.

CONCLUSION

Accountability for human rights violations, including terrorism, whether committed by individuals, associations, corporations or governments, is essential to our future well-being. The horrors of September 11th brought that truth home to those of us living in the United States. But many people in many parts of the world have lived with the terror of state-sponsored human rights violations and state-sponsored terrorism, as well as the abuses of non-state fanatics. If our campaign to rid the world of terrorism truly includes them and the crimes directed against them, we will have contributed to a revolutionary advance towards reducing such atrocities. If we limit our focus to the abuses that endanger ourselves, we both abandon our sisters and brothers around the world and doom our efforts to failure. In the battle against terrorism, partial solutions are no solutions at all.

Terrorism Supported by Rogue States: Some Foreign Policy Questions Created by Involving U.S. Courts

Barry E. Carter*

The United States and other civilized countries should aggressively deal with rogue states that support terrorism.

Fortunately, there are many tools available for these efforts. These tools include the use of military force (as we have just demonstrated in Afghanistan), covert actions, and a variety of economic sanctions against a country or group that supports terrorists. These sanctions -- which preferably would be in union with other countries, or unilateral if necessary -- can include freezing assets, as well as ending or limiting U.S. government programs (ranging from landing rights to foreign aid), cutting off exports to or imports from the country supporting terrorists, denying it credit or investment funds, and working with the World Bank and other multilateral banks to block loans to the country.[1]

---

* © Barry E. Carter and New England Law Review, 2002. Professor of Law and Director, Program in International Business and Economic Law, Georgetown University Law Center. Prof. Carter wishes to acknowledge the ideas, comments, and assistance from his recent research assistants and from his Georgetown colleagues at a summer seminar on a draft of this article. A shorter version of this paper was delivered at the Annual Meeting of the American Society of International Law on March 15, 2002; the presentation will be reprinted in the *Proceedings* from the Annual Meeting.

1. There has been much analysis and debate on the use and effectiveness of international economic sanctions. *See, e.g*, Barry Carter, *Study of U.S. Unilateral Sanctions 1997-2001* (USA*Engage 2002), *available at* http://www.usaengage.org/ 2002sanctions/index.html; RICHARD HAASS, ECONOMIC SANCTIONS AND AMERICAN DIPLOMACY (Brookings 1998); GARY C. HUFBAUER, JEFFREY J. SCHOTT, & KIMBERLY A. ELLIOTT, ECONOMIC SANCTIONS RECONSIDERED: HISTORY AND CURRENT POLICY (Institute for International Economics 2d ed. 1990) (third edition forthcoming in 2002); PRESIDENT'S EXPORT COUNCIL, UNILATERAL ECONOMIC SANCTIONS: A REVIEW OF EXISTING SANCTIONS AND THEIR IMPACTS ON U.S. ECONOMIC INTERESTS WITH RECOMMENDATIONS FOR POLICY AND PROCESS IMPROVEMENTS (June 1997); National Ass'n of Manufacturers, *A Catalog of New U.S. Unilateral Economic Sanctions for Foreign Policy*

Clearly, the U.S. courts also have a role. For example, individual terrorists should be subject to criminal and civil actions whenever we can obtain jurisdiction.

The more complex questions revolve around suits against rogue states, which engage in terrorism or which provide material support to terrorists. Amendments that started in 1996 to the U.S. Foreign Sovereign Immunities Act (FSIA)[2] now essentially allow a new category of suits against a foreign state by: (1) U.S. nationals, (2) who are seeking money damages for personal injury or death, (3) "that was caused by an act of torture, extrajudicial killing, aircraft sabotage, hostage taking, or the provision of material support or resources . . . for such an act if such act or provision of material support is engaged in by an official, employee, or agent of such foreign state while acting within the scope of his or her office, employment, or agency . . . ."[3] These suits are an exception to foreign sovereign immunity under new subsection (a)(7) of Section 1605.

This exception is limited to suits against states that are designated by the Executive Branch as a state sponsor of terrorism. Seven states are now so designated: Cuba, Iran, Iraq, Libya, North Korea, Sudan, and Syria. (They are also often viewed as "rogue states.")

There are a growing number of cases against these states -- especially Iran and Iraq -- that are based on the (a)(7) exception, with many of the suits resulting in very large damage judgments. In 1996, the first judgment under (a)(7) resulted from Cuba's deliberate shooting down of two civilian Cessna aircraft, operated by a group called Brothers to the Rescue, over international waters. The personal representatives of the three U.S. citizens who were killed sued the Republic of Cuba and the Cuban Air Force

---

Purposes 1993-96 (March 1997); Sen. Jesse Helms, *What Sanctions Epidemic?*, FOREIGN AFF., Jan.-Feb. 1999, at 2 (comments on the NAM study and other studies); BARRY E. CARTER, INTERNATIONAL ECONOMIC SANCTIONS: IMPROVING THE HAPHAZARD U.S. LEGAL REGIME (1988).

2. The key provisions of the FSIA, as codified and amended, are at 28 U.S.C. §§ 1330, 1332, 1391(f), 1441(d), and 1602-1611. The FSIA contains "a comprehensive set of [U.S.] legal standards governing claims of immunity in every civil action against a foreign state or its . . . agencies or instrumentalities." Verlinden B.V. v. Central Bank of Nigeria, 461 U.S. 480, 488 (1983). It also provides the legal standards for service of process against foreign states, and the standards for the attachment and execution of that state's assets if there is a judgment against it.

The basic approach of the Act is to provide a blanket of immunity to a foreign state from the jurisdiction of U.S. courts, subject to a set of specific exceptions. These exceptions include, among others, actions where the foreign state has waived its immunity or has engaged in certain commercial activities. As indicated above, the (a)(7) exception was first added in 1996.

3. 28 U.S.C. § 1605(a)(7) (2001).

for damages under (a)(7). The U.S. district court awarded the plaintiffs a default judgment for $187.6 million in compensatory and punitive damages.[4]

A judgment that even better highlights what is happening with the new laws is *Flatow v. Islamic Republic of Iran*.[5] In this case, the U.S. district court awarded $247.5 million in damages, including $225 million in punitive damages, to the administrator of the estate (and father) of Alisa Michelle Flatow, against the defendants, the Islamic Republic of Iran, the Iranian Ministry of Information and Security, and three Iranian officials. Ms. Flatow, a 20-year-old American student at Brandeis University, was killed in April 1995 while riding a bus in the Gaza Strip during an independent foreign study program in Israel. A suicide bomber drove a van loaded with explosives into the bus, destroying it and killing eight persons, including Ms. Flatow, who suffered several hours of pain and suffering before dying of her injuries.

Although Iran and the other defendants did not appear in court to respond to the complaint, the court entered a default judgment against the defendants after a two-day evidentiary hearing that proceeded in the manner of a nonjury trial.[6] The court heard evidence that the Shaqaqi faction of Palestine Islamic Jihad claimed responsibility and perpetrated the terrorist act. The court concluded that the Shaqaqi faction's "sole source" of funding was the Islamic Republic of Iran. The court also found that Iran helped train the faction and that the other defendants were involved in Iran's provision of material support and resources.

These large awards pale in comparison to the case that is still working its way through the U.S. courts. It involves Libya's apparent role in the bombing of Pan Am Flight 103 over Lockerbie, Scotland. All 259 persons on board were killed, many of them U.S. citizens. Representatives of many of those killed brought suit, seeking damages of about $10 billion. After the case had been dismissed prior to the 1966 amendment to the FSIA, the (a)(7) amendment, which includes a 10-year statute of limitations, led the U.S. Court of Appeals for the Second Circuit to decide that the district court had jurisdiction.[7]

There are credible reports that this case is on the verge of being settled,

---

4. Alejandre v. Republic of Cuba, 996 F. Supp. 1239 (S.D. Fla. 1997).

5. 999 F.Supp. 1 (D.D.C. 1998).

6. Section 1608(e) of the FSIA provides in part: "No judgment by default shall be entered by a court of the United States . . . against a foreign state . . . unless the claimant establishes his claim or right to relief by evidence satisfactory to the court." 28 U.S.C. § 1608(e) (2001).

7. Rein v. Socialist People's Libyan Arab Jamahiriya, 162 F.3d 748 (2d Cir. 1998) (Calabresi, J.).

with the Qaddafi government of Libya apparently willing to make a payment of about three billion dollars to the plaintiffs. Qaddafi apparently is seeking to normalize relations with the United States, and he is willing to pay substantial compensation to achieve this.[8] If so, Libya will not only be the only rogue, or terrorist-supporting, state that has appeared in U.S. courts to contest these (a)(7) cases,[9] but it will also be the first case where one of these states has provided funds to settle the case or satisfy a judgment. It would be the one success story so far under the 1996 statute. Nevertheless, a settlement here still does not overcome the need to address the questions about the (a)(7) exception that are raised below.

Besides these three cases, there has been over a score of other cases against Iran, Iraq, and Libya. In most of the recent decisions, the U.S. district courts are awarding about $300 million in punitive damages per suit.[10] Moreover, with terrorist activity continuing and the (a)(7) exception including a 10-year statute of limitations, more such suits would seem likely. Indeed, as a result of the events of September 11th, new suits are being filed, alleging that at least Iran and/or Iraq provided material support to the aircraft hijackers who killed about 3,000 Americans and others.

---

8. President Qaddafi had already turned over two Libyans accused of participating in the bombing to the Scottish Court, meeting specially in The Hague, Netherlands. The court convicted one. Qaddafi's actions there went a long way toward mending fences with the Europeans.

9. On one occasion, Iraq appeared in a (a)(7) case essentially to challenge the applicability of this FSIA exception on a motion to dismiss, and was unsuccessful. Daliberti, et al. v. Republic of Iraq, 97 F. Supp.2d 38 (D.D.C. 2000). Iraq then appeared no further in the case, and a different judge held a hearing and entered a default judgment against Iraq. Daliberti, et al., v. Republic of Iraq, et al., 175 F. Supp.2d 36 (2001). Iraq did not appear in a later (a)(7) case. Hill, et al., v. Republic of Iraq et al., 175 F. Supp.2d 36 (2001).

10. *See e.g.*, Anderson v. Islamic Republic of Iran, 90 F. Supp. 2d 107 (D.D.C. 2000); Sutherland v. Islamic Republic of Iran, 151 F. Supp. 2d 27 (D.D.C. 2001; Jenco v. Islamic Republic of Iran, 154 F. Supp. 2d 27(D.D.C. 2001); Hill v. Republic of Iraq, 175 F. Supp. 2d 36 (D.D.C. 2001); Weinstein v. Islamic Republic of Iran, 184 F. Supp. 2d 13 (D.D.C. 2002). For a comprehensive list of (a)(7) suits through January 25, 2002, see DAVID M. ACKERMAN, SUITS AGAINST TERRORIST STATES (Cong. Res. Serv. Jan. 25, 2002), *available at* http://fpc.state.gov/documents/organization/8045.pdf.

By way of comparison, the federal compensation fund for the victims of the attacks on September 11th has issued regulations indicating that it will make awards ranging from about $1 million to $4 million. The victims' estates, though, can also sue terrorists for additional damages.

THE IMPLICATIONS OF THE LARGE DAMAGE AWARDS

One's initial reaction to these awards, and the possibility of future large awards, might well be that these countries are getting what they deserve that "they ought to pay a price." The implementation of this statute, however, raises some important questions for the long term. What follows is a brief analysis of the experience under the statute and the possible future dynamics.

First, the foreign governments that are identified as supporting terrorists and that lose these cases usually have little interest in paying the judgments. Moreover, except possibly for Iraq and Libya, these foreign governments generally have only limited assets in the United States.[11] Usually these are the assets that were frozen when a rupture occurred in the relations between the two countries. (Countries that might run athwart of U.S. foreign policy learned not to leave large assets in the United States after the United States froze about $12 billion of Iranian assets during the 1979-81 hostage crisis.) As a result, successful plaintiffs in these (a)(7) suits are hard-pressed to find attachable assets to satisfy their default judgments.

Second, in reaction to this difficulty of collection, Congress passed a special law, the Victims of Trafficking and Violence Protection Act of 2000.[12] Pursuant to its terms, the U.S. Government has now paid out about $350 million, from U.S. taxpayer funds, to the successful plaintiffs in nine cases against Iran, with the justification that these claims are being subrogated to the U.S. Government to collect eventually against Iran. The statute might allow just a few other plaintiffs also to collect from the U.S. Government. Not only does this law raise questions why the U.S. Government should be making such large payments to a few plaintiffs, there is no similar provision for compensating later successful plaintiffs against Iran or for covering any judgments against Iraq or any other terrorist state, except for one case against Cuba. Recognizing this problem, the U.S. Congress in 2001 directed President Bush to submit soon "a legislative proposal to establish a comprehensive program to ensure fair, equitable,

---

11. The Office of Foreign Assets Control in the U.S. Department of Treasury estimated that the United States has blocked, as of the end of December 2001, the following net assets in the United States of state sponsors of terrorism: Cuba--$112 million; Iran--$252 million; Iraq--$1.704 billion; Libya--$1.177 billion; North Korea--$30 million; Sudan--$27 million; Syria--$104 million. OFFICE OF FOREIGN ASSETS CONTROL, TERRORIST ASSETS REPORT: CALENDAR YEAR 2001 ANNUAL REPORT TO THE CONGRESS ON ASSETS IN THE UNITED STATES OF TERRORIST COUNTRIES AND INTERNATIONAL TERRORISM PROGRAM DESIGNEES (2002).

12. Pub. L. No. 106-386, § 2002(f)(1), 114 Stat. 1543 (2000). See discussion of the statute and its implementation in ACKERMAN, *supra* note 9, at 14-25.

and prompt compensation for all United States victims of international terrorism . . . that occurred or occurs on or after November 1, 1979."[13]

Third, the question remains how new U.S. plaintiffs or even the U.S. Government (with its subrogated claims) will receive payment from these foreign governments. Since most terrorist states are unlikely to offer new funds or allow more assets to be vulnerable to attachment, it is now a likely eventuality that successful judgment-creditors will not be able to obtain full satisfaction while these governments are in power. Then, when a moderate government takes over in Cuba, Iran, Iraq or elsewhere, and that moderate government wants to do business with the United States, these judgment-creditors will insist on payment of their large judgments, or will seek to attach that country's assets when commercial activities resume between the United States and that country.

Traditionally, outstanding claims between the United States and another country are often settled by a foreign claims settlement agreement when the two countries resume relations.[14] These agreements usually serve to protect a foreign country from any liability beyond the agreed loss of frozen assets or any other assets that the foreign state might agree to contribute.

In the present situation, however, there are serious questions whether a foreign claims settlement agreement is a reasonable approach for the United States in light of the large outstanding, and likely future, judgments under the (a)(7) exception, and the limited assets now frozen by the U.S. Government. The new moderate foreign government might well conclude for economic or domestic political reasons that it does not have the funds, or is unwilling, to contribute more to the settlement agreement. The new government might conclude that it is easier to do business with Europe or Japan.

On the other hand, if a proposed agreement were simply going to distribute the limited frozen assets available at the time of the negotiations, the judgment-creditors would strongly object to the Executive Branch and the Congress (which has shown itself to be sympathetic to the American plaintiffs) because the agreement would only provide the judgment-creditors with just a few cents for each dollar of their judgment. This might well block the conclusion of the agreement.[15]

---

13. Departments of Commerce, Justice, and State, the Judiciary, and Related Agencies Appropriations Act, 2002, § 626, Pub. L. 107-77, 114 Stat. 748 (2002).

14. Dames & Moore v. Regan, 453 U.S. 654, 680-81 (1981).

15. The Executive Branch would hesitate to make the executive agreement a so-called sole or Presidential agreement, which does not require Congressional approval, because of the political opposition by the victims' families and others,

Also, since these judgments are based on federal law (the FSIA), there is some question whether an executive agreement, either entered into by the President alone or one approved by Congress, could provide that the judgment-creditors receive less than full compensation, without giving them a takings claim against the U.S. Government in the U.S. Court of Claims.[16]

Given these questions about the present implementation of the statute and possible future scenarios, it would seem time to consider possible changes to the (a)(7) exception.

First, the exception now includes a provision that the foreign state be afforded "a reasonable opportunity to arbitrate the claim in accordance with accepted international rules of arbitration" *if* "the act occurred in the foreign state against which the claim has been brought."[17] Presumably, this provision reflects some recognition that the foreign state might have reason to question the impartiality of the U.S. courts. In any event, might this provision be expanded to offer the reasonable opportunity for international arbitration in any (a)(7) case? Why should arbitration be available if Iran is alleged to support terrorist torture or murder in Iran itself, but not be available when an alleged Iran-supported terrorist act occurred in Israel or Lebanon? Among the relevant considerations for expanding the situations calling for an initial attempt at arbitration is the possibility that a rogue state might be more likely to participate in an international arbitration than in a U.S. court proceeding, where the foreign country might have concerns about the impartiality of the American judge. On the other hand, from the plaintiff's standpoint, an arbitral award is much more likely to be enforced around the world than a U.S. court judgment. Over 120 countries, including the United States and the European Union states, are parties to the New York Convention on the Recognition and Enforcement of Foreign Arbitral Awards, which provides a generally effective means for obtaining recognition and enforcement of international arbitral awards through the domestic courts of each of these countries.[18]

---

which would likely be reflected in blocking steps in Congress. A congressional-executive agreement, however, might well not obtain the required majority votes in the Senate and House of Representatives.

16. The judgment-creditors could attempt to proceed in the U.S. Court of Claims for a taking. *See* Dames & Moore v. Regan, 453 U.S. 654, 679-83 (1981). The plaintiff there had obtained a federal court judgment on its breach of contract claims, a state law cause of action. *Id.* at 666. As for whether compensation would be allowed, see, for example, Abrahim-Youri v. United States, 139 F.3d 1462 (Fed. Cir. 1997), *cert. denied sub nom.* Gurney v. United States, 524 U.S. 951 (1998).

17. 28 U.S.C. § 1605(a)(7)(B)(i).

18. 21 U.S.T. 2517, T.I.A.S. No. 6997, 330 U.N.T.S. 38. *See generally*

Second, should there be a limit on the damage awards? Tucked away in the 2000 law that provided compensation to several of the successful plaintiffs was a provision that repealed a 1998 provision that allowed plaintiffs to seek punitive damages against the foreign states. The amendment, however, still allows plaintiffs to seek punitive damages against foreign agencies or instrumentalities. The impact, then, of the 2000 amendment is not clear. In any case, there might be an argument for limiting the total damage awards per victim against foreign states, agencies, or instrumentalities. That way, the likely carryover to successor moderate governments would be less burdensome on future relations.

Third, as Prof. Anne-Marie Slaughter of Harvard has suggested,[19] should (a)(7) essentially be repealed and these suits not be allowed against foreign governments or their agencies and instrumentalities? A related 1996 statute, the so-called Flatow amendment, still could continue and allow suits against foreign officials, employees, or agents of these terrorist-supporting states.[20] That way the judgments are personal to the individuals who have supported the terrorist acts, and do not continue against a new government with new officials.

Fourth, should the judgments against a terrorist state be conditioned to lapse upon a major change in the government of the state? As part of the condition, the U.S. Executive Branch would make the determination that there was a sufficiently genuine change in the government. This approach would, incidentally, provide an incentive (though admittedly of little likely impact) for a country to change a government that supports terrorists. To be sure, the traditional rule in international law is that a successor government succeeds to these debts, but this is not an ironclad rule. Moreover, the FSIA is a U.S. statute and the terms of a damage award under it can be statutorily defined.

Above all, just as the U.S. Congress recently asked the Bush Administration for a legislative proposal for a "comprehensive program to ensure fair, equitable, and prompt compensation for all [U.S.] victims of

---

GARY B. BORN, INTERNATIONAL COMMERCIAL ARBITRATION (2d ed. 2001).

19. Anne-Marie Slaughter & David Bosco, *Sue Terrorists, Not Terrorist States*, WASH. POST, October 28, 2000, at A25; *see generally* Anne-Marie Slaughter & David Bosco, *Plaintiff's Diplomacy*, FOREIGN AFF., Sept.-Oct. 2000, at 102.

20. Pub. L. 104-208, Div. A, Title I, § 101 (c), Sept. 30, 1996, 110 Stat. 3009-172 (reprinted at 28 U.S.C. § 1605 note). In *Hill v. Republic of Iraq*, 175 F. Supp. 2d 36 (D.D.C. 2001), Judge Jackson issued an order against President Saddam Hussein for $300 million in punitive damages in an (a)(7) action brought by several people who had been detained and mistreated in 1990 after the Iraqi invasion of Kuwait. The court characterized Saddam Hussein as an "instrumentality" of Iraq (and hence subject to punitive damages), but it would seem that the court could also have reached the same result under the Flatow amendment.

international terrorism,"[21] now would seem to be the time to consider seriously some changes to the FSIA. Terrorism should be fought. Arguably, however, there might be conditions or a limit on the large judgments that might well miss their intended targets, but might hamper future U.S. relations with new, moderate foreign governments.

21. Departments of Commerce, Justice, and State, the Judiciary, and Related Agencies Appropriations Act, 2002, § 626, Pub. L. 107-77, 114 Stat. 748 (2002).

The United States Sanctions Response to the Attacks of September 11, 2001: A Synopsis of Remarks at the NESL Rogue Regimes Conference

Himamauli Das*

The United States has responded to the September 11th attacks through a "broad campaign at home and abroad against terrorism, including by taking military action in Afghanistan."[1] A key component of the U.S. response to this crisis has been through the use of economic sanctions and other restrictions to further a range of foreign policy goals.

The United States has not acted alone in responding to the horrendous attacks on the Pentagon and the World Trade Center on September 11th. The United Nations reacted quickly and constructively to the events of September 11th. Other nations and bodies have likewise responded positively to the September 11th terrorist attacks.

These remarks will provide a brief synopsis of the sanctions-related response of both the United States and international community to the September 11th attacks.

I. THE UNITED STATES SANCTIONS RESPONSE TO THE SEPTEMBER 11TH ATTACKS

The United States' immediate response to the September 11th attacks was designed to "combat and destroy" the "financial underpinnings of global terrorism,"[2] while at the same time taking steps to further U.S. coa-

---

* The author is currently an attorney-adviser for economic and business affairs in the Office of the Legal Adviser, U.S. Department of State. The views expressed herein are the author's own and are not necessarily the views of the United States government or the Department of State.

1. Report by the United States to the Counter Terrorism Committee pursuant to paragraph 6 of Security Council Resolution 1373 of 28 September 2001, Implementation of U.N. Security Council Resolution 1373.

2. President Declares National Emergency, To the Congress of the United States (Sept. 24, 2001) *at* http://www.whitehouse.gov/news/releases/2001/09/20010924.html [hereinafter President's Message to Congress].

lition building and international cooperation in the war against terrorism, including the modification or lifting of certain restrictions against India and Pakistan.[3]

Through both unilateral and multilateral sanctions-related measures, the United States has sought to prevent and suppress acts of terrorism, to deny the use of financing and financial services to terrorists and terrorist organizations, and to facilitate the sharing of information about terrorism and the funding activities in support of terrorism. These efforts complement one another and improve the overall ability to combat terrorism and terrorist financing.

On a unilateral basis, President Bush invoked his authorities under the International Emergency Economic Powers Act (IEEPA)[4] on September 23, 2001 to issue Executive Order 13224 (E.O. 13224), which, in general terms, provide a means to disrupt the financial support network for terrorist organizations, by blocking their assets as well as those of their front organizations, agents, and associates, and other entities and individuals that provide services or assistance to them.[5]

E.O. 13224 provides for sanctions against terrorists and their supporters that are potentially broader in scope than previous sanctions laws and executive orders directed at combating terrorism, including Executive Order 12947 of January 25, 1995[6] or section 219 of the Immigration and

---

3. On September 22, 2001, the President waived most nuclear-related sanctions imposed on India and Pakistan. *See* 66 Fed. Reg. 50095 (Sep. 22, 2001) (determining that the continued imposition of the sanctions imposed under § 102(b)(2)(B), (C), and (G) were not in the national security interest of the United States, pursuant to § 9001(b) of the Department of Defense Appropriations Act, 2000, Pub. L. No. 106-79 (2000) and waiving pursuant to § 9001(a) sanctions imposed under section 101 and 102 of the Arms Export Control Act (AECA), section 2(b)(4) of the Export-Import Bank Act of 1945, and section 620E(e) of the Foreign Assistance Act of 1961, as amended). In particular, these waivers allowed the United States to vote for loans to Pakistan by international financial institutions. On October 27th, President Bush signed into law legislation that exempted for FY 2002 (and provided the President with authority to waive through FY 2003) assistance restrictions imposed on Pakistan as a result its military coup and its default on certain loans. *See* Pub. L. No. 107-57, 115 Stat. 403 (Oct. 27, 2001).

4. *See* 50 U.S.C. § 1701 (2001).

5. *See* Exec. Order No. 13,224, 66 Fed. Reg. 49,079 (Sept. 25, 2001). President Bush designated twenty-seven terrorism-linked individuals and entities in the Annex to E.O. 13224.

6. Exec. Order No. 12,947, 60 Fed. Reg. 5079 (Jan. 23, 1995) *as amended by* Exec. Order No. 13,099, 63 Fed. Reg. 45,167 (Aug. 20, 1998). President Clinton designated twelve terrorist organizations in the Annex to E.O. 12947. Executive Order 13099 amended E.O. 12947 by adding Osama bin Laden, Al-Qaida, and two other organizations to the list of persons designated in the Annex to Executive Order 12947. *See* Exec. Order No. 13,099, 63 Fed. Reg., at 45,167.

Nationality Act,[7] as amended. Executive Order 12947 authorizes, among other things, the blocking of the assets of foreign entities and individuals designated by the President in the Annex to the Order, as amended by Executive Order 13099 of August 20, 1998, as well as those subsequently determined by the Secretary of State, in consultation with the Secretary of the Treasury and the Attorney General, to have committed, or to pose a significant risk of committing, acts of violence that have the purpose or effect of disrupting the Middle East Peace Process.[8] The national emergency declared in Executive Order 12947, however, is limited to the unusual and extraordinary threat posed by the grave acts of violence committed by foreign terrorists that disrupt the Middle East peace process, and does not have the global reach of Executive Order 13224.

Similarly, section 219 of the Immigration and Nationality Act, as added by section 302 of the Antiterrorism and Effective Death Penalty Act of 1996,[9] authorizes the Secretary of State, in consultation with the Secretary of the Treasury and the Attorney General, to designate an organization as a "foreign terrorist organization" ("FTO") if he finds that: a) the organization is a foreign organization; b) the organization engages in terrorist activity, as defined; and c) the terrorist activity of the organization threatens the security of US nationals or the national security of the United States.[10] Several consequences flow from the designation of an organization as an FTO, including the blocking of its assets by U.S. financial institutions, the ineligibility of certain members of the FTO for U.S. visas, and the imposition of criminal penalties for the provision of material support or resources to such organizations.

---

7. *See* 8 U.S.C. § 1189 (2001).

8. *See* Exec. Order No. 12,947, at §1(a)(ii)(A), 60 Fed. Reg. at 5079. E.O. 12947 also authorizes, *inter alia*, the designation of foreign persons found "to assist in, sponsor, or provide financial, material, or technological support for, or services in support of . . . acts of violence [that have the purpose or effect of disrupting the Middle East peace process]." *Id.* at § 1(a)(ii)(B). This contrasts against the scope of collateral designations under E.O. 13224, which authorizes the designation of those persons determined "to assist in, sponsor, or provide financial, material, or technological support for, or *financial or other services to or in support of, . . . acts [that threaten the security of U.S. nationals or the national security, foreign policy, or economy of the United States]* or [ ]persons [designated pursuant to the Order,]" and additional persons "*associated with*" certain categories of persons designated in or under the Order. Exec. Order No. 13,224, at § 1(d), 66 Fed. Reg. at 49080 (emphasis added).

9. *See* Pub. L. No. 104-132, title III, § 321 (Apr. 24, 1996).

10. *See* 8 U.S.C. § 1189(a)(1) (2001). There are currently thirty groups designated as FTOs. By statute, designations expire after two years unless the Secretary redesignates the organization for an additional two-year period upon a finding that the relevant circumstances "still exist." *See* 8 U.S.C. § 1189(a)(4) (2001).

In contrast, E.O. 13224 specifically authorizes the blocking of the property and interests in property, including bank deposits, of those designated in the Annex by the President, as well as those subsequently designated by the Secretary of State or the Secretary of the Treasury, as authorized in the Order. It specifically targets foreign entities and individuals determined to have committed, or to pose a significant risk of committing, acts of terrorism that "threaten the security of U.S. nationals or the national security, foreign policy, or economy of the United States," as determined by the Secretary of State, in consultation with the Secretary of the Treasury and the Attorney General.[11] It may further be used to target other entities or individuals that are acting for or on behalf of, or are owned or controlled by, persons designated in or pursuant to the Order; that assist in or provide support or financial or other services to those entities and individuals designated in or under the Order; or are associated with certain categories of entities and individuals designated in or under the Order, as determined by the Secretary of the Treasury, in consultation with the Secretary of State and the Attorney General.[12] This may include non-governmental or charitable organizations, as well as financial institutions in third countries that provide financial or other services to or for persons designated in or pursuant to the Order.

While these powers are broad in scope, President Bush, in his Message to Congress on Executive Order 13224, "committed to exercising [these powers] responsibly, with due regard for the culpability of the persons and entities potentially covered by the order, and in consultation with other countries."[13] This Executive Order, in a unique departure from previous executive orders imposing sanctions, also allows for a "lesser action" than the complete blocking of assets to be taken if the Secretary of the Treasury, in consultation with the Secretary of State and the Attorney General, deems it to be consistent with the national interests of the United States.[14] Some of the factors that the President has noted may be considered in deciding whether a lesser action is consistent with the national interests of the United States include the impact on the United States or international financial system of the blocking of assets; the extent of cooperation provided by the foreign person with U.S. authorities; the degree of knowledge that the foreign person had in the terrorist-related activities of the designee; the extent of the relationship between the foreign person and the designated person; and the impact of the blocking or other measures on the foreign person.[15] This authority to apply a "lesser action" has to date not

11. *See* Exec. Order No. 13,224, at § 1(b), 66 Fed. Reg. at 49079.
12. *Id.*, § 1(c)-(d).
13. President's Message to Congress, *supra* note 2, at 1.
14. *See* Exec. Order No. 13,224, at § 5, 66 Fed. Reg. at 49081.
15. *See* President's Message to Congress, *supra* note 2, at 2.

been applied.

The Order, by its terms, is as an element of a broader multilateral strategy to combat terrorism and terrorist financing. It specifically directs the Secretary of State, the Secretary of the Treasury and other agencies to cooperate and coordinate with other countries, including through existing and future multilateral and bilateral agreements and arrangements.[16] Many nations have responded to these cooperative efforts by freezing the assets of terrorists and terrorist supporters.[17]

In the Annex to the Executive Order, the President designated eleven terrorist organizations, twelve individual terrorist leaders, three charitable or humanitarian organizations that operate as fronts for terrorist financing and support, and one business entity that operates as a front for terrorist financing or support.[18] Each of the individuals and entities originally designated by the President is linked to Osama bin Laden or to the Al-Qaida network. Including subsequent designations by the Secretaries of State and the Treasury, a total of 191 entities and individuals (as of March 31, 2002) have been designated under E.O. 13224. These designations include all of the foreign organizations previously designated as Foreign Terrorist Organizations (FTOs) by the Secretary of State, many of which are unrelated to either Osama bin Laden or the Al-Qaida network, highlighting the global scope of U.S. efforts to combat terrorist financing. In addition, a number of financial networks, such as those entities linked to the al-Barakaat network, and charitable organizations, including the Holy Land Foundation for Relief and Development, have also been designated by the Secretary of the Treasury.[19] Under E.O. 13224, the Secretary of State has also designated several Irish terrorist groups and a Spanish group that were also designated by the Council of the European Union in a Common Position adopted on December 27, 2001,[20] in an

16. *See* Exec. Order No. 13224, at ¶ 6, 66 Fed. Reg. at 49081.

17. For example, the Secretary of the Treasury reported in testimony before the Committee on Financial Services of the House of Representatives on March 1, 2002 that $104 million in terrorism-linked assets have been frozen globally, and pointed specifically to the cooperation provided by the United Arab Emirates in freezing the assets of Al-Barakaat, a financial network designated by the Secretary of the Treasury pursuant to E.O. 13224.

18. *See* Exec. Order No. 13,224, annex, 66 Fed. Reg. at 49079; President's Message to Congress, *supra* note 2, at 2.

19. *See generally* Office of Foreign Assets Control, http://www.treas.gov/ofac/ (last visited Mar. 18, 2002) (listing the rounds of designations under E.O. 13224).

20. *See* Council Common Position of 27 December 2001 on the application of specific measures to combat terrorism, 2001/931/CFSP, *reprinted in* Official Journal of the European Communities, at L 344/93 (Dec. 28, 2001). On the same day, the Council also adopted a regulation on specific restrictive measures directed

effort to further multilateral efforts to combat terrorism.

II. THE INTERNATIONAL RESPONSE TO THE SEPTEMBER 11TH ATTACKS

The objectives underlying E. O. 13224 gained significant support from the international community at the outset. Within forty-eight hours of the terrorist attacks, both the Security Council and the General Assembly passed resolutions making clear that not only the perpetrators of terrorist attacks, but also those who support them, must be held accountable.[21]

The U.N. Security Council adopted resolution 1368 on September 12th, in which it, among other things, condemned in the strongest terms the horrifying terrorist attacks of September 11th and regarded such acts, like any act of international terrorism, to be a threat to international peace and security. While resolution 1368 did not impose any binding obligations on member states under Chapter VII of the U.N. Charter, it called on member states to work together urgently to bring to justice the perpetrators, organizers and sponsors of terrorist attacks and stressed that those "aiding, supporting or harboring" the "perpetrators, organizers and sponsors" of terrorism will be held accountable. The Security Council further expressed its readiness to take all necessary steps to combat all forms of terrorism.[22]

The Security Council acted again on September 28, 2001 by adopting resolution 1373, which includes binding measures on member states under Chapter VII of the U.N. Charter.[23] Security Council resolution 1373 is a far-reaching measure that aims at limiting the ability of terrorists and terrorist organizations to operate internationally by requiring member states to take certain concrete steps in combating terrorism. Specifically, it requires member states to prevent and suppress the financing of terrorism and to criminalize the wilful provision or collection of funds with the intention that the funds should be used, or in the knowledge that they are to be used, in order to carry out terrorist acts.[24] It further requires that the funds and other financial assets and economic resources of persons who commit, or attempt to commit, terrorist acts or participate in or facilitate

---

against certain persons and entities with a view to combating terrorism, which identified standards and procedures for the designation of terrorists and their supporters. *See* Council Regulation (EC) No. 2580/2001 (Dec. 27, 2001), *reprinted in* Official Journal of the European Communities, L 344/70 (Dec. 28, 2001).

21. *See* U.N. S. C. Res. 1368, U.N. SCOR, 56th Sess., 4370th mtg., U.N. Doc. S/RES/1368 (2001); U.N. G. A. Res. 56/1, U.N. GAOR, 56th Sess., 1st plen. mtg., U.N. Doc. A/RES/56/1 (2001) (adopting this resolution on Sept. 12, 2001).

22. *See* U. N. S. C. Res. 1373, U.N. SCOR, 56th Sess., 4385th mtg., U.N. Doc. S/RES/1373 (2001) [hereinafter U.N. SCOR 1373].

23. *See id.*

24. *See id.* at ¶ 1(a)- (b).

the commission of terrorist acts, among others, be frozen.[25] The resolution also targets terrorists' ability to move across international borders and find safe-haven, to solicit and move funds, and to acquire weapons.[26] It further requires states that do not have laws criminalizing terrorist activity and support to enact such laws immediately.

Given its far-reaching scope, the international community is paying careful attention to how resolution 1373 is being implemented. The Counter Terrorism Committee established by resolution 1373 is charged with monitoring implementation of the resolution, and will highlight best practices, identify gaps, and help coordinate advice and technical assistance to states that need it. Such careful attention is likewise being paid to the implementation of U.N. Security Council resolution 1390. Resolution 1390 complements resolution 1373 by specifically obliging member states to take specific measures against the Taliban, Osama bin Laden, and members of the Al-Qaida organization, as well as others, as designated by the sanctions committee established pursuant to U.N. Security Council resolution 1267.[27]

A number of other multilateral efforts have been taken in preventing and suppressing the financing of terrorism.[28] For example, the Financial Action Task Force on money laundering ("FATF") met in October 2001 and adopted eight special recommendations on terrorist financing and agreed to expand FATF's mandate to include terrorist financing, in addition to money laundering. Other groups -- including NATO, the G-7, the G-8, the OSCE (Organization for Security and Cooperation in Europe), APEC (Asian Pacific Economic Cooperation) forum, the Gulf Cooperation Council and the Organization of American States (OAS) -- are separately meeting or have separately met to address joint efforts to combat terrorism and terrorist financing. These efforts include developing mechanisms for identifying terrorists whose assets should be frozen and for sharing and protecting information regarding the designation of terrorism-linked individuals and entities, and other measures to combat terrorism on a broader basis.

---

25. *See id.* at ¶ 1(c).

26. *See id.* at ¶ 2.

27. *See* U.N. S.C. Res. 1267, U.N. SCOR, 54th Sess., 4051st mtg., U.N. Doc. S/RES/1267 (1999) [hereinafter U.N. SCOR 1267].

28. In addition, there are twelve U.N. conventions that are directed at combating various specific aspects of terrorism. Resolution 1373 calls upon member states to become parties as soon as possible to the relevant international conventions and protocols related to terrorism, including the International Convention for the Suppression of the Financing of Terrorism. *See* U.N. SCOR 1373,*supra* note 22, at ¶ 3(d). In December, the United States Senate granted its advice and consent to ratification of the U.N. Convention for the Suppression of the Financing of Terrorism and the UN Convention for the Suppression of Terrorist Bombings.

III. REFORM OF AFGHANISTAN-RELATED SANCTIONS

A third — and recent — development in the sanctions-related response to the September 11th attacks occurred after the Taliban were removed from power in Afghanistan in late December 2001. This response included the lifting by the United States of the broad trade and transactions embargo on those territories of Afghanistan that had been determined to be controlled by the Taliban and the unblocking of certain assets over which the Afghan Interim Authority was able to exercise authority. The U.N. Security Council, and its Afghan sanctions committee, also took action to modify the scope of the sanctions regime imposed pursuant to United Nations Security Council resolutions 1267 (1999) and 1333 (2000).

A. The Pre-September 11th Afghanistan Sanctions

On July 4, 1999, pursuant to his authority under IEEPA, President Clinton declared in Executive Order 13129 (E.O. 13129) a national emergency with respect to the unusual and extraordinary threat to the "national security and foreign policy of the United States" posed by the actions and policies of the Taliban in Afghanistan, "in allowing territory under its control in Afghanistan to be used as a safe haven and base of operations for Usama bin Laden and the Al-Qaida organization who have committed and threaten to continue to commit acts of violence against the United States and its nationals." [29] E.O. 13129 imposed broad sanctions against the Taliban and with respect to Taliban-controlled territories of Afghanistan.

E.O. 13129 blocked the property and interests in property of the Taliban[30] and of persons determined by the Secretary of the Treasury, in consultation with the Secretary of State and the Attorney General, "to be owned or controlled by, or to act for or on behalf of, the Taliban" or "to provide financial, material, or technological support for, or services in support of" the Taliban or those persons determined to be owned or controlled, or to act for or on behalf of, the Taliban.[31] The Order prohibited any transaction or dealing by United States persons or within the United States in such blocked property, "including the making or received of any contribution of funds, goods, or services to or for the

---

29. Exec. Order No. 13,129, 64 Fed. Reg. 36759 (July 7, 1999).

30. The "Taliban" is defined in section 4(b) of the Order to mean, among other things, "the political/military entity headquartered in Kandahar, Afghanistan and that as of the date of [the Order] exercises de facto control over the territory of Afghanistan described in paragraph [4(d) of the Order]."

31. *Id.*, § 1.

benefit of the Taliban" or other persons designated pursuant to the Order.[32]

In addition, E. O. 13129 imposed a comprehensive embargo on "goods, software, technology (including technical data), or services" to the Taliban or persons designated pursuant to the Order, as well as to "the territory of Afghanistan controlled by the Taliban."[33] The embargo included a ban on both the exportation and re-exportation of such items from the United States or by United States persons, wherever located, and the importation of such items into the Untied States. The "territory of Afghanistan controlled by the Taliban" was defined to mean the "territory referred to as the 'Islamic Emirate of Afghanistan,'" among other things, and included a number of provinces of the country of Afghanistan.[34] The effect of the embargo was to prohibit virtually all financial and other transactions or trade by U.S. persons with the majority of Afghanistan, absent a license from the Office of Foreign Assets Control of the Department of the Treasury.[35]

Following the action by the United States, the United Nations Security Council, acting under Chapter VII of the Charter of the United Nations, adopted resolution 1267 on October 15, 1999, which deplored the "fact that the Taliban continues to provide safe haven to Usama bin Laden and to allow him and others associated with him to operate a network of terrorist training camps from Taliban-controlled territory and to use Afghanistan as a base from which to sponsor international terror" and noted "the indictment of Usama bin Laden and his associates by the United States," for, among other things, the bombings of the U.S. embassies in Nairobi, Kenya, and Dar es Salaam, Tanzania, on August 7, 1998.[36] The Security Council demanded that the Taliban turn over Osama bin Laden "to appropriate authorities in a country in which he has been indicted, or to appropriate authorities in a country where he will be returned to such a country, or to appropriate authorities in a country where he will be arrested and effectively brought to justice."

It further decided that member states were required to freeze the funds and other financial resources of the Taliban, or by any undertaking owned or controlled by the Taliban as designated by the sanctions committee established pursuant to paragraph 6 of U.N. Security Council Resolution 1267 ("1267 Committee"), and to deny permission for any Taliban-owned, -leased, or –operated aircraft to take off from or land in their territory,

---

32. *Id.*, § 2(a).

33. *Id.*, § 2(b)-(c).

34. *Id.*, § 4(d). The term "territory of Afghanistan controlled by the Taliban" was modified to include the City of Kabul by Public Notice 3151 of October 21, 1999, 64 Fed. Reg. 58879 (Nov. 1, 1999).

35. *See generally*, 31 C.F.R. Part 545.

36. U.N. SCOR 1267, *supra* note 27.

unless the particular flight was approved in advance by the 1267 Committee on the grounds of humanitarian need, including religious obligations such as the performance of the Hajj.[37]

By adopting resolution 1333 on December 19, 2000, the Security Council expanded the measures required to be taken against the Taliban.[38] In addition to condemning the continuing use of the Taliban-controlled areas for the sheltering and training of terrorists and planning of terrorist acts, the Security Council reiterated its deep concern over the continuing violations of international humanitarian law and of human rights, particularly discrimination against women and girls, and over the significant rise in the illicit production of opium.

Notably, the Security Council expanded the range of restrictions imposed on the Taliban and Taliban-controlled areas of Afghanistan. In addition to the measures imposed on member states pursuant to resolution 1267, the Security Council imposed an embargo on arms and arms-related material, including weapons and ammunition, military vehicles and equipment, paramilitary equipment, and spare parts to the Taliban and Taliban-controlled territories; prohibited member states from providing technical advice, assistance or training related to Taliban military activities; and required member states to withdraw all military-related personnel from Afghanistan.[39] The Security Council also required U.N. member states to take other measures, including closing the offices of the Taliban and Ariana Afghan Airlines, preventing the sale of the chemical acetic anhydride, used in the production of opium, by their nationals or from their territories to Taliban-controlled territory, and to deny aircraft the permission to take off, land in, or overfly their territories if it had taken off or would land in Taliban-controlled territory, unless approved by the 1267 Committee. As noted above, resolution 1333 also expanded the scope of economic sanctions against the Taliban to impose targeted economic sanctions against Osama bin Laden and Al-Qaida and their associates, as designated by the 1267 Committee.

B. Sanctions after the ouster of the Taliban

Both the United States and the U.N. Security Council quickly acted to modify the sanctions regime against Afghanistan after the Taliban were removed from power. These measures effectively ended the geographically-based restrictions, including the U.S. embargo, on the territory of Afghanistan controlled by the Taliban, and focused sanctions on the Tali-

37. *See id.* at ¶¶ 4(a)-(b).

38. *See* U.N. S.C. Res. 1333, U.N. SCOR, 55th Sess., 4251st mtg., U.N. Doc. S/RES/1333 (2000) [hereinafter U.N. SCOR 1333].

39. *See id.* at ¶ 5.

ban and those individuals and entities associated with it.

The Security Council acted initially to lift measures specifically targeted at certain entities that had been linked to the Taliban by virtue of its control of Afghanistan. It adopted resolution 1388 on Janaury 15, 2002, which terminated the requirement imposed in resolution 1333 that member states close the offices of Ariana Afghan Airlines in their territories, and decided that the measures in resolution 1267 requiring the freezing of certain designated Taliban-related assets no longer applied to the Afghan national airlines.[40] Likewise, the 1267 Committee, either on its own initiative or following Security Council action, acted several times to delist several Afghan government-controlled entities, including the Afghan central bank, that it had placed on the consolidated list of entities and individuals designated pursuant to resolutions 1267 and 1333 whose assets member states were required to freeze.[41] With appropriate independent action by member states to unfreeze the assets of these entities, the Afghan Interim Authority will be able to access previously frozen Afghan government funds to help meet its significant financial needs, such as the payment of government salaries and running of government operations.

The United States followed the Security Council actions by unblocking the assets of several Afghan-government entities, including the Afghan central bank and Ariana Afghan Airlines. On January 24, 2002, the United States also determined that the Taliban no longer controlled any territory in Afghanistan, and modified the definition of "the territory of Afghanistan controlled by the Taliban" to reflect this.[42] Accordingly, while the restrictions against most financial and trade transactions with the Taliban and those designated pursuant to E. O. 13129 continue to remain in place, U.S. persons are now able to engage in most trade and transactions with Afghanistan.

Finally, the Security Council adopted resolution 1390 on January 16, 2002, which continued the measures imposed by paragraph 8(c) of resolution 1333[43] and took note of continued application of the measures im-

---

40. *See* U. N. S. C. Res. 1388, U.N. SCOR, 57th Sess., 4449th mtg., U.N. Doc. S/RES/1388 (2002). Resolution 1388 also lifted with respect to Ariana Afghan Airlines the restriction in resolution 1267 that member states deny permission for Taliban-owned, -leased, or –operated aircraft from taking off or landing in their territories. *See id.* at ¶ 1; U.N. SCOR 1267, *supra* note 27, at ¶ 4(a).

41. For a list of actions taken by the 1267 Committee, see Press Releases Concerning the Security Council Committee Established Pursuant to Resolution 1267 (1999), http://www.un.org/Docs/sc/committees/Afghanistran/AfghanPressEng.htm (last visited Mar. 18, 2002).

42. Public Notice 3899 of January 24, 2002, 67 Fed. Reg. 4301 (Jan. 29, 2002).

43. Several of the measures imposed by UNSCR 1333, including paragraphs

posed by paragraph 4(b) of resolution 1267, in accordance with paragraph 2 of resolution 1390.[44] Paragraph 2 of U.N. Security Council Resolution 1390 directs member states to take certain specific measures with respect to "Usama bin Laden, members of the Al-Qaida organisation and the Taliban and other individuals, groups, undertakings and entities associated with them," as designated by the 1267 Committee.[45] These include, inter alia: (1) the freezing of their "funds and other financial assets or economic resources"; (2) the prevention of the entry into or the transit through their territories of designated individuals; and (3) the prevention of the direct or indirect supply, sale and transfer of arms and arms-related material to designated individuals and entities.[46]

IV. CONCLUSION

The use of sanctions by the United States to respond to the September 11th attacks reflects certain broad considerations regarding the use of economic sanctions. Sanctions can be sharpened to respond to specific foreign policy crises, such as the targeted response to the September 11th attacks, Sudan, and other crises. At the same time, the Executive Branch should retain sufficient flexibility in administering sanctions to achieve specific foreign policy goals.

Most importantly, sanctions can be a valuable foreign policy tool if they are properly designed, implemented and applied as part of a coherent strategy. Sanctions should target foreign governments or specific groups to minimize unintended harm to other important interests. A multilateral approach is usually preferable to a unilateral one. Unilateral economic sanctions should be used only in conjunction with the consideration of other options. Finally, sanctions should be imposed on food only under

---

5, 8, 10, and 11, were established only for a twelve-month period from January 19, 2001. *See* U.N. SCOR 1333, *supra* note 38, at ¶¶ 22-24.

44. *See* U. N. S. C. Res. 1390, U.N. SCOR, 57th Sess., 4452nd mtg., at ¶ 1, U.N. Doc. S/RES/1390 (2002) [hereinafter U.N. SCOR 1390].

45. The consolidated list of entities and individuals designated pursuant to UNSCR 1267 and 1333, and now, UNSCR 1390, by the 1267 Committee has become a particularly important mechanism for identifying entities and individuals linked to the Taliban, Osama bin Ladin and the Al-Qaida organization, and to impose the obligation on all member states to freeze their funds and financial assets or economic resources.

46. U.N. SCOR 1390, *supra* note 44, at ¶ 2. The Security Council further requested the 1267 Committee to "promulgate expeditiously such guidelines and criteria as may be necessary to facilitate the implementation of paragraph 2" of resolution 1390. *Id.* at ¶ 5(d). The 1267 is currently working on revised procedures and guidelines for additions or deletions to the consolidated list of individuals and entities designated pursuant to resolutions 1267 (1999) and 1390 (2002).

extraordinary circumstances, a principle that underlies the Trade Sanctions Reform and Export Enhancement Act of 2000, as amended.[47]

The United States sanctions response is one component of a broad campaign against terrorism, which includes law enforcement, military, diplomatic, and numerous other elements. The sanctions imposed under Executive Order 13224 are narrowly targeted on terrorists and their supporters, as part of a broader multilateral strategy to combat terrorist financing, including through U.N. Security Council resolutions 1267, 1333, 1373 and 1390, and through multilateral institutions such as FATF, the G-8, and others.

47. Pub. L. No. 106-387, Title IX (Oct. 28, 2000).

Managing "Smart Sanctions" Against Terrorism Wisely

Peter L. Fitzgerald*

INTRODUCTION

Economic sanctions, tools that seek to deprive terrorist organizations of the financial ability to support and conduct operations such as the attacks on Washington and New York, are among the U.S. government's most powerful weapons in the war on terrorism. Sanctions greatly expand the reach of the government's own anti-terrorism efforts, by effectively enlisting the aid of the financial and trading communities in identifying and blocking access to assets belonging to terrorist organizations and their supporters. However, there is a significant difference between simply serving political goals by announcing sanctions, and promulgating controls that are actually effective in combating terrorism.

Osama bin Laden and his Al Qaeda network were targeted by a number of different sanctions well before the events of September 11th, as part of a series of anti-terrorist programs that date back to the mid-1990's. Additionally, other sanctions were aimed at the Taliban in Afghanistan, and many other states believed to support or sponsor terrorist activities. Following the attacks, this basic anti-terrorist sanctions framework was bolstered with a new Presidential Executive Order on Terrorism[1] as well as a U.N. Security Council Resolution[2] calling for similar sanctions by other nations. Given the new appreciation of the very real need to deprive terrorists of the ability to finance attacks like those that occurred late last

---

* Peter L. Fitzgerald, Professor of Law, Stetson University College of Law. All rights reserved. Professor Fitzgerald was formerly counsel to the IBM Export Regulation Office, and recently testified before the Congressionally created Judicial Review Commission on Foreign Assets Control. Portions of this article appear in *Tightening the Screws: The Economic War Against Terrorism*, 66 NAT. INT. 76 (Winter 2001 - 2002).

1. *Blocking Property and Prohibiting Transactions with Persons who Commit, Threaten to Commit, or Support Terrorism*, Exec Order No. 13,224, 66 Fed. Reg. 49,079 (Sept. 25, 2001).

2. On September 28, 2001 the Security Council adopted the resolution on Threats to International Peace and Security Caused by Terrorist Acts, S.C. Res. 1373, U.N. SCOR, 56th Sess., 4385th mtg., U.N. Doc. S/RES/1373 (2001).

year, one might ask how these controls can be made more practical and effective in the ongoing war on terrorism than they evidently proved to be in the past?

The government's anti-terrorist sanctions are neither isolated nor unique regulatory programs. Rather, they are but a small part of twenty-one similar, yet ostensibly separate, economic sanctions programs administered by the Treasury Department's Office of Foreign Assets Control (OFAC). When Saddam Hussein threatens peace in the Middle East; when Slobodan Milosevic conducts ethnic cleansing in the republics of the former Yugoslavia; when human rights are abused in Sudan or Burma; when regional peace and stability is threatened in Liberia or Sierra Leone; when international terrorism, narco-trafficking, or weapons proliferation problems arise; economic sanctions are frequently the government's first -- and principal-- tool for dealing with these and a host of other similar complex and intractable international policy issues.

The Terrorist Sanctions Regulations (TSR), promulgated in 1996, require "U.S. persons" (i.e., U.S. citizens or residents wherever they may be located, and U.S. business entities including their overseas branches) to block or "freeze" the assets of those who threaten the Middle East peace process.[3] The TSR also prohibits unlicensed transactions with such parties.[4] When the President signed Executive Order No. 13,224 on September 23, 2001, these controls were expanded to cover all global terrorism.[5] The Foreign Terrorist Organizations Sanctions Regulations (FTOSR), published in 1997, criminalize providing support or resources to foreign terrorist organizations designated by the Secretary of State. Additionally, a broad range of U.S. affiliated financial institutions and businesses providing financial services are required to block funds in which the designated terrorist organizations, or their agents, have any interest.[6] Bin Laden and Al Qaeda are listed among the targets of both of these sanctions programs.[7] Moreover, unlicensed financial dealings with the governments of states which have been designated by the Secretary of State as supporters of terrorism pursuant to the Export Administration Act -- Cuba, Iran, Iraq, Libya, North Korea, Sudan, and Syria -- are criminalized under OFAC's Terrorism List Governments Sanctions Regulations

3. *See* 31 C.F.R. pt. 595 (2001).

4. *See id.*

5. *See Blocking Property and Prohibiting Transactions with Persons who Commit, Threaten to Commit, or Support Terrorism*, Exec. Order No. 13,224, 66 Fed. Reg. 49,079 (Sept. 25, 2001).

6. *See* 31 C.F.R. pt. 597 (2001).

7. *See infra* notes 49-58, and accompanying text.

(TLGSR) issued in 1996.[8] The Treasury Department also maintains separate stand alone economic sanctions programs targeted at each of these countries, with the exception of Syria.[9] Among its other country-specific sanctions programs, OFAC also prohibited unlicensed dealings with the Taliban-controlled portions of Afghanistan and required U.S. persons to block all assets in which the Taliban has any interest under the Taliban (Afghanistan) Sanctions Regulations (TASR).[10]

As an alternative to the use of force, there are compelling reasons for employing such sanctions. Over the last twenty-five years the United States became known for its willingness to use sanctions for a variety of purposes, either unilaterally or in conjunction with its allies. The policy debates associated with the use of economic sanctions usually focus upon the question of whether unilateral sanctions are effective or valuable -- whether they represent "good" policy.[11] Indeed, since less than one-third of the funds attributed to Osama bin Laden or his network were frozen by U.S. action alone, the value of the cooperation of more than 161 other nations in this effort is clear.[12] Multilateral financial sanctions are always more effective than unilateral measures, and multilateral cooperation should be encouraged.

Unfortunately, the preoccupation with the unilateralism debate overwhelms and obscures an equally important issue. The manner in which sanctions programs are actually implemented and administered has immediate and practical consequences. Sanctions can be a very powerful and effective tool, uniting both the government's and the civilian community's

---

8. *See* 31 C.F.R. pt. 596 (2001).

9. *See id.* pts. 500, 515, 535, 538, 550, 560, 575 (2001).

10. *See id.* pt. 545 (2001). The TASR, which were issued on January 11, 2001, implement controls initially imposed by an Executive Order signed on July 6, 1999. *See* Executive Order No. 13,129, 64 Fed. Reg. 36,759 (July 7, 1999). On January 24, 2002, following the military action in Afghanistan, the State Department determined that the Taliban no longer "controlled" any territory in Afghanistan. *See* 67 Fed. Reg. 4301 (2002).

11. *See generally* Peter L. Fitzgerald, *Pierre Goes Online: Blacklisting and Secondary Boycotts in U.S. Trade Policy*, 31 VAND. J. TRANSNAT'L. L. 1 (1998) [hereinafter BLACKLISTING AND SECONDARY BOYCOTTS].

12. *See US-EU Designation of Terrorist Financiers Fact Sheet* TREASURY NEWS, May 3, 2002, *available at* http://www.treas.gov/press/releases/po3070.htm (last visited May 21, 2002); *Remarks by Treasury Secretary Paul O'Neill on New US-Saudi Arabia Terrorist Designations*, TREASURY NEWS, Mar. 11, 2002, *available at* http://www.treas.gov/press/releases/po1086.htm (last visited May 21, 2002); *Terrorist Financing Fact Sheet*, Treasury News, Dec. 20, 2001, *available at* http://www.treas.gov/press/releases/po886.htm (last visited May 21, 2002); *Shutting Down the Terrorist Financial Network*, Treasury News, Dec. 4, 2001, *available at* http://www.treas.gov/press/releases/po841.htm (last visited May 21, 2002).

resources in pursuit of a common objective. The importance of enlisting non-governmental parties in the war on terrorism, for example in foreclosing the terrorists' ability to move or launder funds through "unofficial" channels such as the *hawala* or *hundi* money networks,[13] is now more clear than ever but the devil is in the details. If these programs are not properly implemented, the government's objectives may be undermined or frustrated. Similarly, the burdens imposed on those who seek to comply in good faith with the government's policies may be unnecessarily increased if the regulatory controls are not properly crafted and administered. Moreover, the details, the precision sought by these anti-terrorist sanctions, also distinguishes them from many of the other OFAC programs. The anti-terrorist programs are part of a new wave of controls promulgated over the last several years known as "smart sanctions."[14]

SMART SANCTIONS

Blacklisting particular individuals and entities is at the heart of smart sanctions, and OFAC's anti-terrorist programs. Historically, sanctions were employed to economically isolate entire countries, so they were primarily geographically based controls. Blacklisting particular individuals or entities was also employed with these geographic sanctions and embargoes, but as a secondary tool to restrict the ability of the targeted destination to undermine the adverse impact of the sanctions through "corporate cloaks" or "fronts" operating in third countries.[15] During World War II, for example, President Roosevelt extended the prohibitions on dealing with the Axis powers to particular individuals and entities operating in neutral countries with his "Proclaimed List of Certain Blocked Nationals."[16] With smart sanctions, however, blacklisting is no longer a secondary tool, rather it becomes the primary tool for achieving the gov-

---

13. *See* Gerald P. O'Driscoll, Jr., Brett D. Schaefer, and John C. Hulsman, *Stopping Terrorism: Following the Money,* HERITAGE FOUNDATION BACKGROUNDER NO. 1479 *available at* http://www.heritage.org/library/backgrounder/bg1479.html (last visited May 21, 2002).

14. *See* Peter L. Fitzgerald, *Drug Kingpins and Blacklists, Compliance Issues with U.S. Economic Sanctions (Part 1)*, 4 J. MONEY LAUNDERING CONTROL 360, 362-66 (2001)[hereinafter COMPLIANCE ISSUES, PART 1]; Peter L. Fitzgerald, *"If Property Rights Were Treated Like Human Rights, They Could Never Get Away with This": Blacklisting and Due Process in U.S. Economic Sanctions Programs,* 51 HASTINGS L. J. 73, 88-110 (1999)[hereinafter BLACKLISTING AND DUE PROCESS].

15. *See* COMPLIANCE ISSUES, PART 1, *supra* note 14, at 362-65.

16. *Authorizing a Proclaimed List of Certain Blocked Nationals and Controlling Certain Exports*, Proclamation No. 2497, 6 Fed. Reg. 3555 (July 19, 1941).

ernment's objectives.[17]

The essence of smart sanctions is narrowly targeting the controls at the individual or entity concerned, within or without a particular geography or nation. This level of precision is becoming increasingly important as concern mounts over the humanitarian costs on innocent civilian populations of broader geographically-based sanctions. The new focus on the collateral costs of broad sanctions even prompted the U.S. and the United Kingdom to formally explore "smart sanctions" in connection with the renewal of the Iraqi "oil for food" program,[18] and led the United Nations to consider smart -- or "targeted" -- sanctions in its Interlaken Process.[19] This three-year study effort, coordinated by the Swiss and German missions to the UN, focused on improving the effectiveness of financial sanctions while minimizing their potential for collateral damage. The study culminated in the release of a series of recommendations for the design and implementation of smart sanctions programs,[20] and substantial revisions to the UN's Iraqi sanctions program.[21] Thus, the movement towards smart sanctions is influencing not only the U.S. and other nations' unilateral sanctions programs, but also those multilateral sanctions efforts coordinated under the auspices of the Security Council.

Smart sanctions are sometimes also associated with a subtle but significant shift in the government's objectives for imposing the controls. Rather than seeking the practical and effective economic isolation of an enemy nation as a prelude or adjunct to war, the motivation for these less comprehensive controls is frequently more overtly political. Smart sanctions are frequently imposed to demonstrate political leadership or to claim the moral high ground for domestic or international political purposes. Until very recently, there was often less concern for whether financial sanctions would actually limit the actions of their intended target than might be the case, for example, with the product oriented controls aimed at restricting

---

17. *See* COMPLIANCE ISSUES, PART 1, *supra* note 14, at 365-66.

18. *See* ANNUAL REPORT OF THE SECURITY COUNCIL COMMITTEE ESTABLISHED BY RESOLUTION 661 (1990) CONCERNING THE SITUATION BETWEEN IRAQ AND KUWAIT (July 27, 2001), *available at* http://www.un.org/Docs/sc/committees/IraqKuwait/738e.pdf.

19. *See e.g., Security Council Sanctions Committees: an Overview, available at* http://www.un.org/Docs/sc/committees/INTRO.htm (last visited May 21, 2002).

20. *See* WATSON INSTITUTE FOR INTERNATIONAL STUDIES, TARGETED FINANCIAL SANCTIONS: A MANUAL FOR DESIGN AND IMPLEMENTATION, (Oct. 2001) *available at* http://www.brown.edu/Departments/Watson_Institute/programs/gs/targetedfinsan/TFS.pdf.

21. *See* S.C. Res. 1409, U.N. SCOR, 57th Sess., 4531st mtg, U.N. Doc. S/RES/1409 (2002).

trade in weapons or weapons related technologies.[22] It is unlikely, after all, that a blacklisted organization like the Armed Islamic Group or an avowed terrorist like Osama bin Laden would open accounts at a U.S. affiliated bank or institution under their own names. The difficult problem for the government in crafting more effective financial sanctions is how to best engage the financial and trading communities in identifying and restricting the actions of those who might "front" for someone like bin Laden. While the political desire to be seen as "taking action" on complex international problems might be satisfied by simply announcing and promulgating a sanctions program, the result may be that the resources and attention to detail required to follow through and give these controls teeth is sometimes lacking. Recent events show that simply staking out the political or moral high ground alone is not enough, greater attention to the practical issues associated with implementing these blacklist-based controls is needed in order to improve their effectiveness.

THE ISSUES

The issues associated with OFAC's sanctions fall in two broad categories, the structural or institutional impediments to more effective controls, and particular problems created by the agency's program administration practices.

Structural and Institutional Impediments to More Effective Sanctions

The relatively meager resources devoted to supporting these programs constitute one of the most significant institutional impediments to more effective sanctions. With a staff of roughly one hundred and a budget of approximately $25 million, OFAC currently manages the licensing, compliance, and enforcement aspects of twenty-one separate sets of sanctions, each with its own distinct regulatory and policy requirements, and the newly established Foreign Terrorist Asset Tracking Center.[23] Given the breadth of its responsibilities OFAC is easily overwhelmed when either the Congress or the Executive branch impose new sanctions programs, or a change in policy direction occurs. For example, in 1994 when Congress mandated that "family remittances" to Cuba must be individually licensed, OFAC was required to divert resources usually devoted to its other pro-

---

22. *See* COMPLIANCE ISSUES, PART 1, *supra* note 14, at 365-66; BLACKLISTING AND DUE PROCESS, *supra* note 14, at 88-90.

23. *See*, R. Richard Newcomb, *Response to Written Questions*, November 14, 2000, *in* JUDICIAL REVIEW COMMISSION ON FOREIGN ASSET CONTROL, FINAL REPORT TO CONGRESS, App. B, at 47-52 (Jan. 2001)[hereinafter COMMISSION REPORT].

grams.[24] This resulted in substantial delays in processing licenses or responding to queries concerning the agency's other sanctions programs.[25] Most of those who deal with these sanctions programs on a practical level, both within government and the regulated community, would agree that more resources are required to properly administer the increasing number of programs OFAC is asked to manage. Improving the effectiveness of these controls, however, requires more than simply providing additional personnel or funding. A change in OFAC's basic regulatory approach, a change in how the agency actually uses its resources, is also necessary.

As a small office within the Treasury Department, and partly as a result of being the institutional successor to the WWII era Office of the Alien Property Custodian, OFAC has often labored in relative isolation from both other government agencies and the regulated community. The Director of OFAC has stated that the Office is primarily a law enforcement agency engaged in "economic warfare," and should not be considered as a "trade control" agency like the Commerce Department's Bureau of Industry and Security (BIS)[26] or even the State Department's Office of Defense Trade Controls (DTC).[27] This self image -- along with limited resources and a lack of integration with the other government agencies that control financial and trade transactions -- means that OFAC never developed a working partnership with the regulated community, those who must im-

---

24. *See,* R. Richard Newcomb, *Response to Written Questions*, November 6, 2000, *id*, at 37-38.

25. Interestingly, there are no statutory or regulatory time periods in which OFAC must act to render a decision on any license application, a petition to "unblock" assets, or other similar requests. Even apart from the period when the demand for Cuban family remittance licensing impacted its operations, OFAC has sometimes been quite slow to act. In some cases it has taken as long as two years to receive a licensing decision, and in others the agency refused to supply any formal response whatsoever. *See* BLACKLISTING AND DUE PROCESS, *supra* note 14, at 115-16. The congressionally created Judicial Review Commission on Foreign Asset Control formally recommended that OFAC take steps to improve its responsiveness in dealing with the regulated community following a $1 million study of OFAC's practices administering its various sanctions programs. *See,* COMMISSION REPORT, *supra* note 23, at 138-39 (Recommendation #5). However, no formal action has been taken on any of the Commission's recommendations as yet.

26. Formerly known as the Bureau of Export Administration, the Commerce Department changed the organization's name in April 2002 to more accurately reflect the scope of the Bureau's responsibilities and activities. *See Industry and Security Programs; Change of Agency Name*, 67 Fed. Reg. 20,630 (April 26, 2002).

27. *See* Keith Kendrick, *In Economic Battle, Soldiers Fight on Carpet,* WASH. POST, Sept. 19, 1990, at A21; Benjamin Weiser, *Sanctions Czar: Big Gun or Loose Cannon?*, WASH. POST, Aug. 4, 1996, at H1; and BLACKLISTING AND SECONDARY BOYCOTTS, *supra* note 11, at 10-11.

plement the dictates of OFAC's programs into their routine business processes.

The actual effectiveness of OFAC's regulatory programs, not unlike the tax code, is much more dependant upon voluntary implementation and compliance with its rules by the regulated community than on the agency's own enforcement actions. For the financial and trading communities to properly augment the government's anti-terrorist efforts, or any other sanctions, the controls themselves must be commercially practicable and attuned to the demands of business. Through advisory committees and a variety of other means the BIS, and to a lesser degree the DTC, built cooperative relationships with the regulated community which helps provide the government with a better understanding of what does and does not work in the marketplace, and in the various internal compliance programs businesses establish to meet the government's requirements.[28] Although this type of cooperation might be particularly useful in improving the effectiveness of OFAC's financial sanctions in an era of rapidly developing e-commerce technologies and new online business practices, it has not occurred.

As a result of its past isolation, and the lack of resources to devote to this task, OFAC' s institutional culture tends to regard the regulated community as an adversary rather than a potential ally.[29] Accordingly, the lack of a good working partnership with those who must comply with its rules, OFAC's "stakeholders" to use corporate jargon, is a fundamental flaw in the way these sanctions are currently formulated and administered. OFAC's usual mode of operation is to "command," without considering whether its orders make sense in the 21st century business environment, rather than enlisting the regulated community in a cooperative effort to achieve a goal which both those in and out of government support, in this case, winning the war against terrorism.[30]

These resource constraints and historic institutional attitudes permeate the problems associated with OFAC's administrative practices. The Judicial Review Commission on Foreign Asset Control, which was

---

28. *See e.g.*, U.S. DEPT. OF COMMERCE, *History of the President's Export Council*, *available at* http://www.ita.doc.gov/td/pec/history.html (last visited May 21, 2002); and U.S. DEPT. OF COMMERCE, *Technical Advisory Committees, available at* http://bxatac.doc.gov/ (last visited May 21, 2002).

29. *See* Peter L. Fitzgerald, *Drug Kingpins and Blacklists, Compliance Issues with U.S. Economic Sanctions (Part 2)*, 5 J. Money Laundering Control 66 (2001) [hereinafter COMPLIANCE ISSUES, PART 2].

30. The need for the agency to work more closely with the regulated community permeated the various recommendations the Congressionally created Judicial Review Commission on Foreign Asset Control made following its examination of OFAC's administrative practices. *See* COMMISSION REPORT, *supra* note 23.

established by Congress in connection with a new set of smart sanctions on drug trafficking imposed under the Foreign Narcotics Kingpin Designation Act of 1999[31] and chaired by Larry D. Thompson (now the Deputy Attorney General), examined OFAC's operations in a year long $1 million study with a particular emphasis on issues associated with procedural due process.[32] Without digressing into the merits of the policies behind the Kingpin Act sanctions, or any other sanctions program, the Thompson Commission identified a number of issues relating to program implementation and administration which cut across all of OFAC's sanctions programs. The Commission summarized its findings in a Final Report to the House and Senate Intelligence Committees which was submitted in January, 2001. Among its twelve major recommendations,[33] the Thompson Commission strongly urged the creation of a formal advisory committee under the Federal Advisory Committee Act.[34] It stated that:

> [s]uch advisory committees already exist in connection with other functions of the Department of the Treasury, as well as in connection with the Department of Commerce's administration of export control laws, a regulatory area similar in important respects to the administration of the economic sanctions laws. The creation of an advisory committee regarding U.S. economic sanctions programs would provide a forum for dialogue that would both promote greater understanding by OFAC of the problems the regulated community confronts in seeking to comply with U.S. economic sanctions laws, and greater understanding by the regulated community of how OFAC enforces those laws.[35]

Additionally, the Thompson Commission urged that Congress appropriate additional funding for OFAC's operations, significantly noting that all of its other recommendations regarding reforming the agency's administrative practices were conditioned on the agency having more resources available to remedy the problems it identified.[36]

Recognizing the importance of improving the dialogue between the government and the regulated community, the National Foreign Trade Council -- a leading business organization representing many of the country's largest exporters and multinational businesses-- recently declared that working with OFAC to develop concrete measures reflecting the Commission's recommendations is among the NFTC's top priorities for

---

31. 21 U.S.C. §1908 (2000).
32. *See supra* note 25.
33. COMMISSION REPORT, *supra* note 23, at 125-47.
34. 5 U.S.C. app. 2 §9 (2000).
35. COMMISSION REPORT, *supra* note 23, at 146 (Recommendation #11).
36. *See id.* at 146-47 (Recommendation #12).

2002.[37]

Program Administration Practices Which Impede More Effective Sanctions

There are also a number of problems associated with OFAC's administration of its anti-terrorist and other sanctions programs. Four topics are of particular importance: the structure of OFAC's regulatory controls; the manner in which it publishes its controls, decisions, and other actions; the avenues available to challenge erroneous decisions; and the agency's enforcement practices.

Firstly, OFAC administers more than a score of different largely standalone and self-contained regulatory programs, many of which are driven by divergent policy objectives.[38] Each time a new set of sanctions is ordered an entirely new set of regulatory controls is created. With so many programs there are often delays in promulgating the detailed regulations implementing the new controls.[39] For example, it took nearly eighteen months for the Taliban (Afghanistan) Sanctions Regulations to appear in the Federal Register following the President's announcement of the sanctions.[40] Additionally, the agency interprets and applies the same regulatory wording differently in various programs, depending upon the intended target of the sanctions, making it difficult to discern a coherent approach to the controls. In fact, in the anti-terrorist sanctions and other recently published programs, OFAC now routinely states that it reserves the right to apply "differing interpretations of similar language" in each of its various sanctions programs.[41]

---

37. National Foreign Trade Council, *NFTC's 2002 Goals & Priorities, available at* http://www.nftc.org/2002%20Priorities.html (last visited May 20, 2002).

38. *See* 31 C.F.R. ch. V (2001).

39. Delays issuing the detailed regulations implementing particular sanctions programs have been a chronic problem for more than a decade. *See* COMPLIANCE ISSUES, PART 2, *supra* note 29, at 66-67. OFAC also has a similar problem in not promptly removing outdated controls. *See id.* at 68-69.

40. The Executive Order imposing the Taliban sanctions was issued in July, 1999. *See* Exec. Ord. 13,129, 64 Fed. Reg. 36,759 (July 7, 1999). The TASR themselves were not issued until January of 2001. *See* 66 Fed. Reg. 2,726 (Jan. 11, 2001).

41. The very first provision in the TSR, for example, states:

> This part is separate from, and independent of, the other parts of this chapter . . . . Differing foreign policy and national security contexts may result in differing interpretations of similar language among the parts of this chapter. No license or authorization contained in or issued pursuant to those other parts authorizes any transaction prohibited by this part. No license or authorization contained in or issued pursuant to any other provision of law or regulation authorizes any transaction pro-

The reason for the proliferation of these regulatory programs is largely historical, and greatly influenced by the fact that lawmakers and regulators tend to focus on the policy concern of the moment prompting a particular set of new sanctions. However, there are only so many forms that financial sanctions, trade sanctions, travel limitations, and contract restrictions may take. Despite the variety of policy objectives which may influence the precise combination of sanctions "tools" applied in any given case, the tools themselves are fairly standard. Rather than issuing new regulations for each and every target, OFAC should restructure its regulations into a single master set of possible sanctions, and then selectively invoke those that are appropriate to whatever target and policy objective is at hand. In doing so, greater speed, simplicity, and uniformity of approach would be achieved without any real loss of flexibility in applying the controls.

Secondly, there are a host of issues associated with the manner in which OFAC publishes its various regulations and actions, and in particular, the manner in which entries are added or changed in OFAC's blacklists. Historically, OFAC focused most of its attention on providing "actual notice" of its wishes to the U.S. financial community through the Federal Reserve Bank's Fedwire system, and a number of similar vehicles, but largely eschewed timely publication of its actions and regulations in the Federal Register. There was even a time when OFAC resorted to what was colloquially known as the "radiator distribution system," where the only way for those not in the financial industry to learn of OFAC's actions was to check for hardcopy handouts on the heating radiator in the entry vestibule to the Treasury Annex building where OFAC is located.[42] In recent years, OFAC has made extensive use of its web page and the Internet to distribute information, but continues to place less emphasis on timely publication in the official Federal Register.[43] In essence, the Internet has become a "virtual radiator distribution system" for the agency. While this use of technology may be laudable, and is an improvement over the old radiator system, it does not comport with the Administrative Procedure Act and the Federal Register Act requirements to publish official actions in the Federal Register before they become legally effective.[44] Moreover, failure to use the Federal Register -- which provides "constructive notice" of official actions -- can also complicate or undermine OFAC's enforcement activi-

---

hibited by this part.

31 C.F.R. § 595.101(a) (2001). Virtually identical provisions are found in the other two anti-terrorist programs, the TLGSR, *see id.* § 596.101, and the FTOSR, *see id.* § 597.101, as well as in the sanctions aimed at the Taliban, *see id.* § 545.101.

42. *See* COMPLIANCE ISSUES, PART 2, *supra* note 29, at 69-70.

43. *See id.*, at 69-73.

44. *See* 5 U.S.C. §§ 552(a)(1), 553 (2000); 44 U.S.C. § 1505(a) (1994).

ties.[45] Although OFAC sometimes asserts that it is time consuming to publish in the Federal Register, it is nevertheless the government's own legally mandated official publication process.[46] Moreover, urgent material can be published quickly when the need arises. The Iraqi blocking orders promulgated following the invasion of Kuwait[47] and even the new Executive Order on Terrorism,[48] for example, both appeared in the Federal Register within twenty-four hours of their becoming effective.

Additionally, on a practical level, having multiple official and unofficial means of distribution which provide differing information greatly complicates the task of those who are trying in good faith to stay current and comply with OFAC's controls. This is particularly critical with smart sanctions because of the vital importance of knowing who is or is not blacklisted. For example, "Usama Bin Ladin" was first identified as a sanctioned terrorist in an August 22, 1998, Executive Order,[49] and formally added to OFAC's blacklist of "Specially Designated Terrorists" in a Federal Register notice on July 1, 1999.[50] However, his name did not appear in the "changes list" to the blacklist available on OFAC's website until October 12, 1999, when the spelling was changed to "Bin Laden" and the Al Qaeda organization was added - along with multiple spelling variations and aliases.[51] This updated information from OFAC's website did not in turn appear in the Federal Register until eight months later, on June 23, 2000.[52] The cycle continued as other changes to the terrorist

---

45. *See infra* notes 95-97 and accompanying text.

46. The Thompson Commission noted that the Federal Register should be used for OFAC's blacklist designations, and that the agency should also make greater use of the Federal Register to seek notice and comment concerning proposed regulations and actions whenever possible, as well as make greater use of official publication channels to enhance the "transparency" of its operations. *See* COMMISSION REPORT, *supra* note 23, at 130 (Recommendation #2), 142-143 (Recommendation #9), and 144-145 (Recommendation #10).

47. Exec. Order No. 12,722, 3 C.F.R. 294 (1990); Exec. Order No. 12,723, 3 C.F.R. 296 (1990).

48. Exec Order No. 13,224, 66 Fed. Reg. 49,079 (Sept. 25, 2001).

49. Exec Order No. 13,099, 63 Fed. Reg. 45,167 (Aug. 25, 1998).

50. *See Blocked Persons, Specially Designated Terrorists, Foreign Terrorist Organizations, and Specially Designated Narcotics Traffickers: Additional Sudanese Government Designations and Supplementary Information, and Removal of One Individual*, 64 Fed. Reg. 35,575, at 35,576 (July 1, 1999).

51. *See* entry for Oct. 12, 1999 in Office of Foreign Assets Control, *SDN Changes 1999*, http://www.ustreas.gov/offices/enforcement/ofac/sdn/sdnew99.txt (last visited May 21, 2002).

52. *See Blocked Persons, Specially Designated Nationals, Specially Designated Terrorists, Foreign Terrorist Organizations, and Specially Designated Narcotics Traffickers; Addition of Persons Blocked Pursuant to 31 CFR Part 597, or Executive Order 13129*, 65 Fed. Reg. 39,100-103 (June 23, 2000).

blacklist in the June 23rd Federal Register notice did not appear on the OFAC website until August 18, 2000.[53] Moreover, following the attacks on September 11th and the release of the President's Executive Order on September 25, 2001, OFAC used its website to publish important additional "descriptor information" clarifying its terrorist blacklist entries, creating a new blacklist category of "Specially Designated Global Terrorist" (SDGT) for bin Laden and the others identified in the Executive Order, and making a variety of other changes to over two hundred terrorist blacklist entries,[54] none of which appeared in the Federal Register until October 26, 2001.[55]

This also highlights one of the basic limitations of smart sanctions, that is their complete dependence upon the quality and extent of the information contained in the blacklist. For example, despite sanctions aimed at bin Laden having first been imposed in 1998, the spelling "Osama bin Laden" currently being used in most reports was not found in the list appended to the President's September 23rd Executive Order,[56] and did not appear on

---

53. *See* COMPLIANCE ISSUES, PART 2, *supra* note 29, at 73.

54. *See* Office of Foreign Assets Control, *Recent OFAC Actions (Oct. 12, 2001)*, http://www.ustreas.gov/offices/enforcement/ofac/actions/20011012.html (last visited May 21, 2002); Office of Foreign Assets Control, *Recent OFAC Actions (Oct. 5, 2001)*, http://www.ustreas.gov/offices/enforcement/ofac/actions/20011005.html (last visited May 21, 2002); Office of Foreign Assets Control, *Recent OFAC Actions (Oct. 3, 2001)*, http://www.ustreas.gov/offices/enforcement/ofac/actions/20011003.html (last visited May 21, 2002); Office of Foreign Assets Control, *Recent OFAC Actions (Oct. 1, 2001)*, http://www.ustreas.gov/offices/enforcement/ofac/actions/20011001.html (last visited May 21, 2002). It is also striking that the changes made to the terrorist blacklist on both October 1st and 3rd also purport to be retroactive to September 24, 2001, the date of the President's Executive Order.

55. *See Blocked Persons, Specially Designated Nationals, Specially Designated Terrorists, Foreign Terrorist Organizations, and Specially Designated Narcotics Traffickers: Additional Designations of Terrorism – Related Blocked Persons*, 66 Fed. Reg. 54,404 (Oct. 26, 2001). OFAC issued another Federal Register notice on November 15, 2001, which announced a number of other blacklist changes, primarily associated with its narco-trafficking sanctions. The only "terrorist" related item in the November notice was the designation of the "Real IRA," and its various aliases, as an FTO. *Additional Designations and Removal of Persons Listed in Appendix A to 31 CFR Chapter V and Appendix I of 31 CFR Part 539, Weapons of Mass Destruction Trade Control Regulations*, 66 Fed. Reg. 57,371 at 57,379-380 (Nov. 15, 2001). Interestingly, this November notice did not note that the "Real IRA" was also blacklisted as an SDGT nearly two weeks earlier on OFAC's website. *Compare id.*, *with* Office of Foreign Assets Control, *Recent OFAC Actions (Nov. 2, 2001)*, http://www.ustreas.gov/offices/enforcement/ofac/actions/20011102.html (last visited May 21, 2002).

56. *See* Exec. Order No. 13, 244, 66 Fed. Reg. 49,079 at 49,083 (Sept. 25, 2001).

OFAC's website until October 1, 2001, with a notation that it should be applied retroactively to the date of the Executive Order.[57] Moreover, that spelling did not appear in the blacklist published in the Federal Register until October 26, 2001.[58] Thus, even though Osama bin Laden was a government target nearly three years before the recent attacks, an automated screen run against the official version of OFAC's terrorist blacklist would likely have missed the name altogether before late last October. Even if the screening incorporated the additional "unofficial" information found on OFAC's website, no match would have occurred based upon the information provided prior to October 1st -- slightly more than two weeks after the events of September 11th.

Entirely apart for the question of the sufficiency of the information provided for particular blacklist entries, the extent to which OFAC relies upon its website in apparent disregard of the Federal Register publication requirements is equally significant. For example, in addition to the chronology detailed above, on November 2, 2001, the scope of the regulatory controls applicable to Foreign Terrorist Organizations (FTOs) were effectively expanded by their dual designation as FTO/SDGTs by a simple notice on OFAC's website.[59] Five days later sixty-two additional parties were blacklisted as SDGTs by means of entries in the website's "What's New" file.[60] The focus of the anti-terrorist sanctions was then further expanded with similar SDGT designations of three Hamas-controlled organizations on December 4th, seven Pakistani individuals or entities on December 20th and 26th, one Spanish and five Irish-related entities on December 31, 2001,[61] and four Afghani and Pakistani SDGTs with multi-

---

57. *See* Office of Foreign Assets Control, *Recent OFAC Actions (Oct. 1, 2001)*, http://www.ustreas.gov/offices/enforcement/ofac/actions/20011001.html (last visited May 21, 2002).

58. *See Blocked Persons, Specially Designated Nationals, Specially Designated Terrorists, Foreign Terrorist Organizations, and Specially Designated Narcotics Traffickers: Additional Designations of Terrorism-Related Blocked Persons*, 66 Fed. Reg 54,404 at 54,405, 54,407, and 54,409 (Oct. 26, 2001).

59. *See* Office of Foreign Assets Control, *Recent OFAC Actions (Nov. 2, 2001)*, http://www.ustreas.gov/offices/enforcement/ofac/actions/20011102.html (last visited May 21, 2002). The dual designation of these parties as both FTOs, sanctioned under the FTOSR, and SDGTs, sanctioned under the TSR, means that both sets of regulations control any dealing with the named entities.

60. *See* Office of Foreign Assets Control, *Recent OFAC Actions (Nov. 7, 2001)*, http://www.ustreas.gov/offices/enforcement/ofac/actions/20011107.html (last visited May 21, 2002).

61. *See* Office of Foreign Assets Control, *Recent OFAC Actions (Dec. 26, 2001)*, http://www.ustreas.gov/offices/enforcement/ofac/actions/20011226.html (last visited May 21, 2002); Office of Foreign Assets Control, *Recent OFAC Actions (Dec. 20, 2001)*, http://www.ustreas.gov/offices/enforcement/ofac/actions/

ple aliases on January 9, 2002.[62] Moreover, amidst this other activity, the "What's New " file noted that on December 14th three U.S.-based organizations were added to the anti-terrorist blacklists as "BPI-PA,"[63] or blocked pending an investigation regarding further actions pursuant to the USA Patriot Act.[64] The six-month anniversary of the World Trade Center and Pentagon attacks passed before any of these blacklisting actions appeared in the government's official Federal Register.[65] Additionally, when the blacklisting notice actually did appear in the Federal Register, on March 19, 2002,[66] it was both incomplete and already outdated. The changes detailed in OFAC's Federal Register notice this past March failed to highlight the dual re-designation of the FTOs as SDGTs,[67] the blocking

---

20011220.html (last visited May 21, 2002); and Office of Foreign Assets Control, *Recent OFAC Actions (Dec. 4, 2001)*, http://www.ustreas.gov/offices/enforcement/ofac/actions/20011204.html (last visited May 21, 2002).

62. *See* Office of Foreign Assets Control, *Recent OFAC Actions (Jan. 9, 2002)*, http://www.ustreas.gov/offices/enforcement/ofac/actions/20020109.html (last visited May 21, 2002).

63. *See* Office of Foreign Assets Control, *Recent OFAC Actions (Dec. 14, 2001)*, http://www.ustreas.gov/offices/enforcement/ofac/actions/20011214.html (last visited May 21, 2002).

64. *See infra* notes 79-87, and accompanying text for a discussion of some of the concerns associated with effectively blacklisting a party pending a decision to formally designate them in one or another of OFAC's official blacklists.

65. OFAC asserts that its use of its website for notification of blacklisting actions is fully effective. It prefaces the file it posts on its website summarizing changes to the blacklist throughout the year with the following:

> The latest changes may appear here [on the website] prior to their publication in the Federal Register, and *it is intended that users rely on changes indicated in this document that post-date the most recent Federal Register publication* with respect to a particular sanctions program .... Such changes reflect official actions of OFAC, and will be reflected as soon as practicable in the Federal Register .... *Users are advised to check the Federal Register and this electronic publication routinely for additional names or other changes to the listings.*

Office of Foreign Assets Control, *Changes to List of Specially Designated Nationals and Blocked Persons Since January 1, 2002*, http://www.ustreas.gov/offices/enforcement/ofac/sdn/t11sdnew.pdf (last visited May 21, 2002) (emphasis added).

66. *Additional Designation of Terrorism-Related Blocked Persons*, 67 Fed. Reg. 12,644 (Mar. 19, 2002).

67. *Compare id.*, *with* Office of Foreign Assets Control, *Recent OFAC Actions (Nov. 2, 2001)*, http://www.ustreas.gov/offices/enforcement/ofac/actions/20011102.html (last visited May 21, 2002).

There was, however, another Federal Register notice issued by the State Department on the same date as OFAC's blacklist update, March 19, 2002. *See Designations of Terrorists and Terrorist Organizations Pursuant to Executive Order 13224 of September 23, 2001*, 67 Fed. Reg. 12,633 (Mar. 19, 2002). The State Department notice updated the list of foreign terrorist groups that were subject to the sanctions imposed under Section 1(b) of Executive Order 13,224, the

of U.S. parties under the USA Patriot Act,[68] or the blacklisting of the Spanish, Irish and Pakistani parties which occurred up through January 9, 2002.[69] The changes published in March also failed to note the removal of

---

order which effectively expanded OFAC's TSR to include all terrorism globally. Nevertheless, the State Department notice did not expressly use either the FTO or SDGT designations commonly employed in the OFAC blacklists. Accordingly, while the intent may have been to trigger sanctions under both the FTOSR and the TSR as expanded by the President's Executive Order, there was no clear statement to that effect in any of the OFAC blacklist information published in the Federal Register (as opposed to the agency's website). Thus, based upon the information provided on OFAC's website on November 2, 2001, all those blacklisted as FTOs (and subject to the controls of the FTOSR) are also generally blacklisted as SDGTs (and subject to the controls of the TSR) - even though there are inconsistencies in the agency's use of the dual FTO/SDGT designation subsequent to November 2nd. In contrast, based upon the information OFAC published in the Federal Register, one could easily be led to conclude that the TSR do not govern dealings with FTOs. It is only when the additional information provided in the State Department notice is considered, with an understanding of the interrelationship between the State Department and Treasury Department designations, that one would understand that both sets of OFAC regulations apply to the named parties. Moreover, this requires that those seeking to comply with the controls now monitor the notices published in the Federal Register by two different Governmental Departments. Prior to this time, the practice had been for OFAC to publish the names of those designated by the Secretary of State as foreign terrorist organizations in conjunction with the OFAC blacklists. *See e.g.* the naming of the "Real IRA" as an FTO in *Additional Designations and Removal of Persons Listed in Appendix A to 31 CFR Chapter V and Appendix I of 31 CFR Part 539, Weapons of Mass Destruction Trade Control Regulations*, 66 Fed. Reg. 57,371, 57,379-380 (Nov. 15, 2001), discussed *supra* at note 55.

68. *Compare Additional Designation of Terrorism-Related Blocked Persons*, 67 Fed. Reg. 12,644 (Mar. 19, 2002), *with* Office of Foreign Assets Control, *Recent OFAC Actions (Dec. 14, 2001)*, http://www.ustreas.gov/offices/enforcement/ofac/actions/20011214.html (last visited May 21, 2002).

69. *Compare Additional Designation of Terrorism-Related Blocked Persons*, 67 Fed. Reg. 12,644 (Mar. 19, 2002), *with* Office of Foreign Assets Control, *Recent OFAC Actions (Dec. 31, 2001)*, http://www.ustreas.gov/offices/enforcement/ofac/actions/20011231.html (last visited May 21, 2002); Office of Foreign Assets Control, *Recent OFAC Actions (Dec. 26, 2001)*, http://www.ustreas.gov/offices/enforcement/ofac/actions/20011226.html (last visited May 21, 2002). The New Year's Eve designation on OFAC's website of the Irish and Spanish organizations as SDGTs is, however, implicitly noted in the State Department's March 19, 2002, Federal Register notice. Interestingly, OFAC failed to indicate that these Irish and Spanish groups should also be considered FTOs following the State Department's action. The State Department's Federal Register notice also marks the first appearance in the Federal Register of OFAC's December 26th dual designation of several Afghani and Pakistani groups as FTOs/SDGTs but, as is the case with all the parties identified in the State Department notice, fails to employ the FTO or SDGT designator commonly used with

three Taliban entities from the blacklist on February 12th, the designation of twenty-one new Spanish SDGTs on February 26th, the addition of entities in both Bosnia-Herzegovina and Somalia to the terrorist blacklist on March 11th, or modifications to the address information for two SDGT's locations in the Netherlands on March 15th, all of which were previously announced on OFAC's website.[70] Similarly, OFAC used its website to subsequently identify three other entities as FTO/SDGTs on March 27th, name ten new SDGTs on April 19th, blacklist nine more SDGTs on May 3rd, and provide new "aka" information for two older SDGT entries on May 15th, all without any corresponding Federal Register announcement or notice.[71] This dramatically illustrates the difficulty facing those who seek to comply with the government's anti-terrorist blacklist controls, and the rather extraordinary need to monitor and compare multiple official and unofficial sources for current information. It also substantially undermines the ability of the financial and trading communities to help achieve the government's objective of blocking access to these assets and therefore limits the effectiveness of the controls.

OFAC's blacklists. *See Designations of Terrorists and Terrorist Organizations Pursuant to Executive Order 13224 of September 23, 2001*, 67 Fed. Reg. 12,633 (2002) and the discussion of the confusion between the two Department's Federal Register notices *supra* note 55.

70. *Compare Additional Designation of Terrorism-Related Blocked Persons*, 67 Fed. Reg. 12,644 (Mar. 19, 2002), *with* Office of Foreign Assets Control, *Recent OFAC Actions (Mar. 15, 2002)*, http://www.ustreas.gov/offices/enforcement/ofac/actions/20020315.html (last visited May 21, 2002); Office of Foreign Assets Control, *Recent OFAC Actions (Mar. 11, 2002)*, http://www.ustreas.gov/offices/enforcement/ofac/actions/20020311.html (last visited May 21, 2002); Office of Foreign Assets Control, *Recent OFAC Actions (Feb. 26, 2002)*, http://www.ustreas.gov/offices/enforcement/ofac/actions/20020226.html (last visited May 21, 2002); and Office of Foreign Assets Control, *Recent OFAC Actions (Feb. 12, 2002)*, http://www.ustreas.gov/offices/enforcement/ofac/actions/20020212.html (last visited May 21, 2002). It should perhaps also be noted that the March Federal Register notice also failed to note either the designation twenty-seven Mexican SDNTKs in January or the removal of another SDNTK designation in early March. *See* Office of Foreign Assets Control, *Recent OFAC Actions (Mar. 7, 2002)*, http://www.ustreas.gov/offices/enforcement/ofac/actions/20020307.html (last visited May 21, 2002); and Office of Foreign Assets Control, *Recent OFAC Actions (Jan. 31, 2002)*, http://www.ustreas.gov/offices/enforcement/ofac/actions/20020131.html (last visited May 21, 2002).

71. During this same time period OFAC did, however, publish a Federal Register notice concerning changes another of its blacklists. In April it published a notice that it was adding five organizations to the blacklist associated with its Balkans sanctions. *Addition of Persons to Appendix A to 31 CFR Chapter V*, 67 Fed Reg. 16308 (Apr. 5, 2002).

Moreover, in addition to sometimes being under-inclusive because of a failure to provide or properly distribute appropriate information, blacklisting mechanisms may also be over-inclusive. The more common the name of the blacklisted individual or entry, the more likely unproductive false matches will occur when transactions are screened against the blacklist. If a name such as Ahmed or Muhammed is as common in Arabic as Jones or Smith is in English, for example, the number of false matches could easily overwhelm those administering blacklist compliance processes. Accordingly, in order to have a practical and effective control, and not just a political statement, more is called for than simply promulgating a list of names. The accuracy, nature, and amount of information provided in the government's blacklist is crucial in determining the commercial practicability of the control.[72] Indeed, a substantial portion of the U.N. recommen-

---

72. The case of Johnnie Thomas, a seventy-year-old African-American grandmother who was repeatedly and erroneously stopped at various airports by the FBI's "no fly" or "master terrorist" list, illustrates the problems that can arise when blacklist based controls involve common names. "John Thomas Christopher" was one of the aliases used by Christian Michael Longo -- a 28 year old, blue-eyed, reddish-blond male --but the alias also appears in the FBI's "no fly" list as John or Johnny Thomas. The FBI list was shared with Immigration, Customs, and the U.S. Marshals, and eventually wound up in the hands of the airlines who stopped Mrs. Thomas, conducted special searches of her person and luggage, temporarily took her passport, and prevented her from flying on her scheduled flights on several occasions since March 15, 2002. None of the governmental agencies or private companies involved would take responsibility for removing the name from the blacklist, even after learning that Longo had been captured and imprisoned since January, despite Mrs. Thomas' pleas for assistance to the FBI, Transportation Security Administration, and her representatives in Congress. Recently, Mrs. Thomas' plight has caught the attention of the media, in part due to a rap song she wrote describing her travails:

> It was the March 15th in 2002,
> decided to visit my grandchildren for a week or two.
> 'Where are you going' asked the man at USAir,
> 'New York, to Boston, my daughter's Vineyard lair.'
>
> He typed into the computer, turned a pale shade of gray.
> 'Don't think you'll be flying our shuttle today.
> I'll have to have your passport, you're a security risk.
> Your name has come up on the master terrorist list.'
> Say,'uh-huh, uh-huh, uh-huh, uh-huh, uh-huh!'
>
> The agent called the security guard who looked at me askance
> as she approached me aggressively in a kung fu stance.
> 'Mrs. Thomas is not a problem,' the worried man said.
> But if looks could have killed, I would have been dead.
> Say,'uh-huh, uh-huh, uh-huh, uh-huh, uh-huh!'
>
> 'We can't catch the anthrax, that may be true,
> but by golly, Mrs. Thomas, we did get you!'

dations for implementing smart sanctions focuses upon the crucial role the proper distribution of complete and accurate information plays in making the controls effective.[73] It is with processes and problems such as these where a cooperative partnership between the regulated community and the government might be useful in devising more effective controls.

Thirdly, as noted earlier, it is unlikely that criminals or terrorists will walk in the door of any U.S. affiliated business using their correct names. Thus, multilateral support for pursuing those at the heart of these controls is certainly crucial to their effectiveness, extending the reach of the sanctions to other jurisdictions which the terrorists use to conduct their dealings. Where the domestic blacklists associated with smart sanctions have their real value and enhance the effectiveness of a control program, however, is in blocking assets and transactions with those who launder or "front" for the targeted criminals and terrorists. These are what the Thompson Commission called "Tier II" parties, when they were looking at the narco-trafficking controls.[74] The first tier of blacklisted parties would be the designated terrorists (or drug kingpins in the Commission's study) at the core of the sanctions, and this second tier would be comprised of those who knowingly deal with the other blacklisted parties or, in the terms of the new Executive Order, are "associated" with blacklisted terrorists. This significant conceptual distinction is implicit, but not self-evident, in OFAC's blacklists. The current blacklist controls do not distinguish between those who are the essential targets of the controls, and those who may have a more peripheral involvement with the target. The distinction, in the current system, lies primarily in the discretion the agency exercises in deciding whether or not to blacklist a particular party at all.

As conceivably there might be an almost endless chain of parties who dealt with parties who dealt with those on the blacklist, unless the government is extraordinarily careful when it does decide to add a name to the blacklist, there can be substantial problems associated with these Tier II designations. Mistakes can occur where innocent transactions could result in an "inappropriate" blacklisting entry. Under the regulatory scheme

---

Dierdre McNamer, *Here's Johnnie*, THE NEW YORKER, May 13, 2002, at 34; John Ydstie, *Johnnie Thomas Being Stopped at Airports*, NPR: All things Considered, May 15, 2002, *available at* 2002 WL 3496262.

73. *See* WATSON INSTITUTE FOR INTERNATIONAL STUDIES, TARGETED FINANCIAL SANCTIONS: A MANUAL FOR DESIGN AND IMPLEMENTATION, at 94-98, *available at* http://www.brown.edu/Departments/Watson_Institute/programs/gs/targetedfinsan/TFS.pdf (last visited May 21, 2002); *see also* COMMISSION REPORT, *supra* note 23, at 130 (Recommendation #2), 142-43 (Recommendation #9), and 144-45 (Recommendation #10).

74. *See* COMMISSION REPORT, *supra* note 23, at 45-54.

there is little redress. OFAC's regulations do permit parties to administratively request reconsideration of their being blacklisted,[75] or to seek that frozen funds be "unblocked."[76] However, there is no right under the regulations to review the factual basis or reasons for the initial blacklisting decision. In fact OFAC once issued, and then quickly withdrew, a regulation providing for such review - in part based upon a concern for protecting intelligence sources and methods.[77] Moreover, given the deference the courts give to Congressional and Executive decisions touching upon foreign affairs, there is very little practical ambit for judicial review of OFAC's actions in this area.[78]

One case which highlights the extraordinary power vested in OFAC's blacklisting authority is that involving Salah Idris, the owner of the El-Shifa pharmaceutical plant in Northern Sudan which was bombed in 1998, along with bin Laden's camps in Afghanistan, in retaliation for the attacks on the U.S. embassies in Kenya and Tanzania. Intelligence sources originally thought that the El-Shifa plant was owned by bin Laden and producing chemical weapons precursors. Salah Idris was believed to be acting for bin Laden and OFAC ordered the Bank of America to block roughly $24 million of assets held in Idris's name in the United Kingdom, "pending a decision" on whether to add his name to the terrorist blacklist. Idris sued challenging OFAC's authority to take such an action in the absence of being named in any published blacklist.[79] Idris's assets were voluntarily unblocked in May, 1999, when the government declined to fight the suit. He is now seeking compensation from the U.S. government for the damage to the pharmaceutical plant.[80]

---

75. 31 C.F.R. § 501.807 (2001).

76. *Id at* § 501.806 (2001).

77. As initially promulgated in 1998 OFAC's Reporting and Procedures Regulations (RPR) provided that those seeking administrative reconsideration of their blacklisting could review the factual basis for the agency's actions, although sensitive material could be redacted. 31 C.F.R. § 501.807 (1998). However, less than a year later, OFAC withdrew the ability to obtain this material without explanation. *See Reporting and Procedures Regulation: Procedure for Requests for Removal from List of Blocked Persons, Specifically Designated Nationals, Specifically Designated Terrorists, Foreign Terrorist Organizations, Specifically Designated Narcotics Traffickers, and Blocked Vessels*, 64 Fed. Reg. 5,614 (Feb. 4, 1999). Parties may still request that a blacklisting decision be administratively reconsidered, but they no longer have a regulatory right to review the basis for the agency's actions. *See* 31 C.F.R. § 501.807 (2001).

78. *See* BLACKLISTING AND DUE PROCESS, *supra* note 14, at 135-60.

79. *See* Vernon Loeb, *A Dirty Business*, WASH. POST, July 25, 1999, at F1; Daniel Pearl, *After the Bombings*, WALL ST. J., October 28, 1998, at A1.

80. Vernon Loeb, *Review of Sudan Attack Sought, Bill asks Court to Study U.S. Bombing of Plant, Possible Redress*, WASH. POST, September 27, 2000, at A5;

Interestingly, when the Kingpin Act was passed OFAC persuaded the Congress to provide the statutory authority to order assets blocked "pending" a decision on whether to designate someone as a blacklisted narcotrafficker.[81] In a slightly different fashion, Executive Order No. 13,224 specifically provides that "no prior notice of a listing or determination" under the anti-terrorist sanctions is required,[82] and in the USA Patriot Act Congress similarly amended the International Emergency Economic Powers Act to specifically authorize blocking assets "pending" an investigation into further actions.[83] The Thompson Commission was highly critical of this type of provision, and urged the repeal of that section of the Kingpin Act. Additionally, in terms which seem particularly forceful if one regards terrorism as an assault on the "rule of law," the Thompson Commission urged that blacklisted parties:

> be accorded the right to challenge their designation in an administrative hearing on the record pursuant to standard APA procedures, provided that they file their challenge within sixty (60) days after publication of their designation in the Federal Register. Additionally, [the government] should establish that no Tier II designation by the Secretary of the Treasury is legally effective until it is published in the Federal Register . . . [84][and that] U.S. businesses . . . should be able to rely upon clearly defined standards governing Tier II designations [which would enable them to] better know how to undertake due diligence efforts to minimize the possibility that they are doing business (or may in the future) with a potential client or partner who will be [a] designated entity . . . .[85]

The Commission continued to note that post-blacklisting administrative procedures could afford adversely affected parties, like Idris, basic due process of law without seriously affecting the government's control objectives.[86] Additionally, legitimate concerns for protecting intelligence sources and methods are perhaps best handled on a case by case basis, rather than by the current blanket denial of the right to review the basis for OFAC's blacklisting decisions.[87] These are especially important points given that the master OFAC blacklist already contains several thousand

---

Milt Bearden, *U.S. Should Admit its Mistake in Sudan Bombing*, WALL ST. J., May 20, 1999, at A20; David S. Cloud, *U.S. Unfreezes Accounts in Suit on Sudan Bombing*, WALL ST. J., May 4, 1999, at A8.

81. 21 U.S.C. §1905(a)(2) (2000).

82. Exec. Order No. 13,224, §10, 66 Fed. Reg. 49,079 at 49,081 (2001).

83. *See* Pub. L. 107-56 §106, October 26, 2001, *amending* 50 U.S.C. §1702 (2000).

84. COMMISSION REPORT, *supra* note 23, at 130 (Recommendation #2).

85. *Id.* at 135 (Recommendation #3).

86. *Id.* at 131 (Recommendation #2).

87. *Id.* at 132-133 (Recommendation #2).

names, and can be expected to grow dramatically as the war on terrorism proceeds.

Fourthly, OFAC's enforcement practices, and the lack of publicity regarding its enforcement actions, further undermines the effectiveness of the agency's sanctions. The letter of the anti-terrorist sanctions, like most other OFAC sanctions programs, imposes strict liability for any unlicensed dealings with blacklisted parties.[88] Nevertheless, hundreds of thousands of financial transactions are processed daily by various U.S. affiliated parties. The New York based Clearing House Interbank Payments Company, alone, handles a staggering 200,000 payments worth $1.1 trillion through its funds transfer system on an average day.[89] Given the breadth of the regulatory controls, inadvertent violations of the sanctions are presumably quite common, but there appears to be little public information regarding OFAC's enforcement activities.[90] For example, there are only eight readily available press releases concerning OFAC enforcement actions since 1990, and these only address the Cuban or Libyan sanctions programs.[91] The agency has historically been very reluctant to disclose information regarding its enforcement activities or any settlements it may negotiate to forestall more formal proceedings. The weekly Corporate

---

88. *See* COMPLIANCE ISSUES, PART 2, *supra* note 29, at 75-77.

89. *See* Written Submission of the New York Clearing House Association *in* COMMISSION REPORT, *supra* note 23, at app. D52.

90. *See* COMPLIANCE ISSUES, PART 2, *supra* note 29, at 75-77.

91. *See Treasury Penalizes Caterpillar Industrial for Sales to Libya*, TREASURY NEWS, October 1, 1999, *available at* http://fedbbs.access.gpo.gov/library/download/FAC_MISC/pr100191.txt (last visited May 21, 2002); *Treasury Penalizes Two Former Congressmen for Libyan Dealings*, TREASURY NEWS, May 6, 1993, *available at* http://fedbbs.access.gpo.gov/library/download/FAC_MISC/pr050693.txt (last visited May 21, 2002); *Treasury Raps Banks, Brewer, for Libyan Sanctions Violations*, TREASURY NEWS, July 27, 1992, *available at* http://fedbbs.access.gpo.gov/library/download/FAC_MISC/pr072792.txt (last visited May 21, 2002); *Treasury Cracks Down on Americans in Libya*, TREASURY NEWS, March 24, 1992, *available at* http://fedbbs.access.gpo.gov/library/download/FAC_MISC/pr032492.txt (last visited May 21, 2002); *Treasury Penalizes Phibro Energy for Dealing with Libyan Company*, TREASURY NEWS, December 20, 1991, *available at* http://fedbbs.access.gpo.gov/library/download/FAC_MISC/pr122091.txt (last visited May 21, 2002); *Treasury Announces Civil Penalties for Cuba Violations*, TREASURY NEWS, April 15, 1999, *available at* http://fedbbs.access.gpo.gov/library/download/FAC_MISC/pr041599.txt (last visited May 21, 2002); *Treasury Collects $20,000 for Cuban Embargo Violations*, TREASURY NEWS, August 28, 1997, *available at* http://fedbbs.access.gpo.gov/library/download/FAC_MISC/pr082897.txt (last visited May 21, 2002); *Merck Settles Treasury Cuban Embargo Matter for $127,000*, TREASURY NEWS, October 24, 1995, *available at* http://fedbbs.access.gpo.gov/library/download/FAC_MISC/pr102495.txt (last visited May 21, 2002).

Crime Reporter recently noted that such secrecy when enforcing the law was "unusual" compared to the practices of other agencies concerned with corporate and business compliance practices such as the Federal Trade Commission, Justice Department, Occupational Safety and Health Administration, or the Securities and Exchange Commission, which tend to inform the public as soon as any enforcement action is filed. When the publication filed a Freedom of Information Act Request for more enforcement data and sought an explanation why the agency took a different approach, OFAC simply stated, "we do not give interviews about internal procedures."[92]

However some additional information about the agency's enforcement practices may be found in documents like the "Departmental Offices Performance Report for FY 2000," which are largely circulated internally. That Report indicates that during FY 2000 over 7,300 banking transactions were blocked affecting just under $500 million, and that OFAC's Compliance Division sent out close to a thousand inquiries which led to over 500 warning letters regarding questionable transactions. The Report also discloses that during the same time period "over 1,000 compliance cases [were instituted] -- the majority of them concerned with OFAC's terrorism-related programs, with 132 ultimately referred for civil penalties and 73 referred for criminal investigation."[93] Testimony by the Director of OFAC before the Senate Appropriations Subcommittee on the Treasury further indicates that this ultimately resulted in six criminal prosecutions.[94] However, little or none of this information reaches the regulated community, much less the general public. The lack of publicity regarding particular cases is perhaps understandable, as no business wants to be seen as associated with terrorists or other similar blacklisted parties. Nevertheless, the combination of the sweeping scope of the controls and the lack of publicity regarding the enforcement of those controls on even an aggregate

92. OFAC informed the Corporate Crime Reporter, in response to its FOIA request, that there were between 100-150 enforcement cases settled between May 1998 and May 2000. Specific details were only provided regarding ten of those cases, but that details on additional cases would be released "on a rolling basis." The ten cases involved alleged violations of the Cuban, Hatian, Iranian, North Korean, and Sudanese sanctions, which were settled for amounts ranging from $2,000 to $40,000. *See Treasury Begins to Release OFAC Enforcement Documents*, CORPORATE CRIME REPORTER, Vol. 16, No. 9 (March 4, 2002).

93. *Program Performance Report, Fiscal Year 2000*, at 172, *available at* http://www.ustreas.gov/gpra/2000rpt.pdf (last visited May 21, 2002).

94. R. Richard Newcomb, Director of the Office of Foreign Assets Control, *Statement before the Senate Appropriations Committee Subcommittee on Treasury and General Government*, (May 10, 2001) *available at* http://www.ustreas.gov/press/releases/po370.htm.

level combines to create an impression -- rightly or wrongly -- that the controls are either unenforced or only selectively enforced at best.

In fact, OFAC's sanctions programs are not enforced as strictly as they are written. Whether due to resource constraints or policy considerations, there does appear to be an element of selectivity in the agency's enforcement practices. For example, in recent hearings before the Treasury and General Government Subcommittee of the Senate Appropriations Committee the Director of OFAC acknowledged that tens of thousands of Americans travel to Cuba each year without the license approval required under the Cuban sanctions. Under the Bush Administration, civil penalty actions against these violators increased by almost 250% over the last year of the Clinton Administration, to a total of 766 cases, but only seven officials are available to monitor enforcement of the Cuban program.[95] Even this modest level of enforcement prompted criticism from Senator Byron Dorgan, who noted that the defendants in these cases included Marilyn Meister, a 75-year-old retired schoolteacher who went on a Canadian sponsored bike trip, and Cevin Allen, 56, who scattered his missionary parents' ashes at the Cuban church they helped found. Senator Dorgan quipped that these cases were a "waste of resources" unlikely to affect the object of the controls, the Castro government, and that "one would think when we're engaged in the war on terrorism, we would be going after terrorists, rather than tracking down little old ladies who are taking bicycle trips to Cuba."[96] When the government does go after the terrorists, however, the agency's inconsistent enforcement record also hampers the prosecution effort. Defense attorneys for John Walker Lindh, the Californian accused of fighting with the Taliban challenged two of the ten counts in the case against him on the basis that the government is engaged in a selective prosecution. They assert that Lindh is the only person ever charged with "supplying services to the Taliban" contrary to the TASR, despite publicly available reports of commercial transactions with the Taliban by at least two large U.S. based companies, and paid consultantcy services by at least two other individuals -- one of whom is on Lindh's defense team.[97]

To help define its enforcement practices, OFAC developed a series of

---

95. Carolyn Skorneck, Trav*elers to Cuba Complain of Fines*, AP Online, February 11, 2002, *available at* 2002 WL 13774966. Following former President Jimmy Carter's visit to Cuba, however, the Bush Administration announced that it was committed to upholding the Cuban embargo. *See* Elisabeth Bumiller, *Bush Ties Easing of Cuba Embargo to Change in Rule*, NY TIMES, May 21, 2002, at A1.

96. Christopher Marquis, *Bush Plans to Tighten Sanctions, Not Ease Them*, NY Times, May 15, 2002.

97. Katherine Q. Seelye, *Lindh Lawyers Claim "Selective Prosecution,"* NY TIMES, May 17, 2002, at A21.

internal policies regarding the use of warning letters, establishing de minimis levels below which enforcement actions are not taken as a matter of prosecutorial discretion (at least for a first offense), and specifying both aggravating and mitigating factors to be considered during settlement negotiations. However, prior to its production to the Thompson Commission, none of this guidance was available to the public.[98] Additionally, OFAC's refusal to release any information concerning the imposition of civil penalties or settlements, adverse licensing decisions, or its general policies regarding the exercise of its prosecutorial discretion actually exacerbates the public perception problem. Given the breadth of the controls as written, the lack of agency guidance concerning their actual application and enforcement makes compliance appear to be overly burdensome and commercially impracticable.

Most companies will not knowingly become involved in terrorism or other similarly sanctioned activities. In a competitive marketplace, however, legitimate businesses will be skeptical of the need to screen even the most trivial of their dealings against a blacklist in order to catch those few transactions which are prohibited -- which typically must be done in order to comply with these sanctions as currently written -- unless they believe their competitors are also required to do so. The lack of visible enforcement activity by the agency thus contributes to the impression that, for legitimate businesses, ignoring OFAC's controls is a relatively low risk proposition. This is further exacerbated when the sanctions themselves are not perceived as being commercially reasonable, or attuned to modern business demands -- such as the shift to online business transactions conducted at "Internet speed."[99] Here again a stronger partnership with the regulated community might help formulate controls which are seen to be more practicable, which would then promote greater compliance and effectiveness in achieving the practical blocking of prohibited transactions. A stronger relationship would also provide new avenues for the government to communicate its enforcement practices and successes.

It is also important that Congress and the Executive branch work together for a more coherent approach to sanctioning those engaged in terrorist activities. Otherwise, policies and legal measures which may appear sound in isolation may conflict in practice. The problem of what to do with the funds that get blocked pursuant to these sanctions highlights some of these difficulties. Typically, when another country's assets are

---

98. *See* COMMISSION REPORT, *supra* note 23, at app. C.

99. For a general discussion of the application of these controls to online businesses see Peter L. Fitzgerald, *Hidden Dangers in the E-Commerce Data Mine: Governmental Customer and Trading Partner Screening Requirements*, 35 INT'L LAW. 47 (2001).

blocked or frozen under one of these sanctions programs, they are simply held as "bargaining chips" until more normal governmental relations are resumed. The competing claims are then offset -- usually through the Foreign Claims Settlement Commission or a similar process -- and the funds disbursed. However, blocking assets belonging to criminals and terrorists is much more like forfeiture -- it is highly unlikely that Osama bin Laden or Al Qaeda will ever resume "normal" relations with the U.S. and seek to negotiate the recovery of any blocked assets.[100] How these assets will be handled raises a number of issues, and Congress and the Executive branch will need to closely coordinate their policies regarding the disposition of the funds blocked pursuant to the anti-terrorist sanctions.

Some confusion has already occurred. For example, in 1996 Congress amended the Foreign Sovereign Immunity Act to permit civil suits for damages arising from terrorist acts against those countries designated by the Secretary of State as sponsors of terrorism. Subjecting such states to civil liability was thought to be an added sanction, bolstering the pressure on them to renounce their support for terrorism. Since that time nearly $1 billion has been awarded in various suits against Iran and Cuba. However, the assets available to satisfy these claims are blocked pursuant to OFAC's sanctions. Accordingly, faced with plaintiffs who suffered terrible harm but were unable to collect upon their awards, Congress passed the Victims of Trafficking and Violence Protection Act. This Act authorizes paying these plaintiffs' awards from the U.S. Treasury, in exchange for subrogating their claims against the foreign states to the U.S. government.[101] This, in turn, is raising questions over whether the U.S. government's payment of civil awards directed against foreign terrorist states defeats the purpose of permitting such suits to be instituted in the first place. Similar issues can be expected to arise concerning the assets blocked under the anti-terrorist programs, and the victims of the attacks in Washington and New York.

CONCLUSION

The war on terrorism will not be won by military action alone. Smart sanctions, economic controls specifically directed at those who conduct or support terrorist activities, will be an integral part of the effort. These sanctions must be designed to have a practical and substantial impact in the commercial marketplace, and cannot be simply political statements. The details of the implementing regulations, and OFAC's administration of those regulatory controls, is of vital importance in creating more effec-

100. *See* COMPLIANCE ISSUES, PART 2, *supra* note 29, at 75, n.427.
101. Pub. L. No. 106-386, Div. C, § 2002, 114 Stat. 1541-1543 (2000).

tive anti-terrorist sanctions. In order to develop more effective controls the regulated community must become a full partner with the government in this effort, and that will require a conscious and perhaps difficult change in approach. Moreover, as terror constitutes an assault on the "rule of law," it becomes all the more important to observe the basic legal principles and due process in administering these programs. Working with those who must implement its orders, OFAC should make a concerted effort to simplify and clarify its various processes, as greater "transparency" in its controls will actually promote greater voluntary compliance with its sanctions programs. Blacklists, and other actions which OFAC takes, should be timely published in the official Federal Register. Those who become ensnared in the process should have an opportunity to review the facts leading to their blacklisting whenever possible, and a real opportunity to challenge those facts. OFAC should also align its regulatory controls with its enforcement practices to reflect its internal de minimis guidelines, and consider establishing either a "knowledge" requirement for any violation, or crafting a "safe harbor" provision to minimize the problem of inadvertent violations occurring. The agency should also publish guidelines regarding its enforcement practices, and give much greater publicity to those cases it does prosecute.

Lastly, the Thompson Commission Report comprises three volumes and several thousand pages of testimony, evidence, and written submissions regarding OFAC's sanctions programs. It is the single most comprehensive source on U.S. economic sanctions practice and procedure compiled since WWII, and makes numerous recommendations which deserve serious consideration by the Congress, the Executive Branch, and interested members of the regulated community and public. Unfortunately, only a few copies of the Commission's Final Report to Congress were prepared and delivered to the House and Senate Intelligence Committees. It deserves to be reprinted and made more widely available, as there are few better starting points for anyone interested in seeing the anti-terrorist and other OFAC sanctions programs made more effective in achieving their stated policy goals.[102]

---

102. While only a few hard copies of the Thompson Commission Report were printed, it is available in microfiche at selected Federal Depository Libraries under SuDoc. Y 3.2:F 76/F 49; Y 3.2:F 76/F 49/V.1-2; GPO Item No. 1089 (MF); and a text-searchable version of the Report in Adobe "pdf" format is available online at http://www.law.stetson.edu/JudicialReviewCommission/.

NOTES

Light Pollution in the United States: An Overview of the Inadequacies of the Common Law and State and Local Regulation

I. INTRODUCTION

"[T]he typical urban or suburban observer might only see a few hundred of the brightest stars, and none of the more elusive objects. We have abused the darkness. We have lost the faint lights." – Chet Raymo[1]

In the late nineteenth century, Thomas Alva Edison and other scientists engineered what could possibly be one of the greatest inventions of all time – the electric light bulb.[2] Today, the options available in electric il-

---

1. CHET RAYMO, THE SOUL OF THE NIGHT: AN ASTRONOMICAL PILGRIMAGE 24 (1985).

2. *See* COLLIER'S ENCYCLOPEDIA 719, 722 (MacMillan Educational Co., 1992). "When Edison invented the first practical electric light bulb in 1879, he could hardly have anticipated its transformative effect on American life in the 20th century." Joe Rey-Barreau, *Illuminating America's Century*, HFN THE WEEKLY NEWSPAPER FOR THE HOME FURNISHING NETWORK, May 31, 1999, at 65. The Chairman and CEO of General Electric, Jack Welch, believes "it is impossible to imagine a world without the electric light bulb [and] street lights ..." Ann Chambers, *GE Backs Edison Preservation Effort*, POWER ENGINEERING, Mar. 1, 2000, at 16. More than 100 years ago in 1893, the newly established General Electric Company boasted that "its lamps extended 'in an unbroken line around the earth; they shine in the palace of the Mikado as well as in the Opera House of Paris.'" THOMAS F. O'BOYLE, AT ANY COST: JACK WELCH, GENERAL ELECTRIC AND THE PURSUIT OF PROFIT 23 (1998). Even today, General Electric manufactures more than fifty percent of the lightbulbs sold in the United States. *See id.* at 57. As this Note will argue, Mr. Welch's cast of a rosy glow on the lighting industry may not be as harmless as he and his company portrays it to be. *See infra* notes 44-94 and accompanying text.

lumination have advanced significantly since the first rudimentary bulb.[3] The various modes of lighting are of crucial importance because electric illumination is used in virtually all areas of daily life, including in the home and automobile, for commercial purposes, and within the community.[4]

The proliferation of outdoor light use, particularly that light projecting into the night skies, has developed into a relatively new area of environmental degradation – light pollution.[5] Astronomers first noticed light pollution approximately one century after Edison's historical accomplishment.[6] Often described as "the presence of excessive illumination in locations where it is not desired,"[7] light pollution has been discovered to have

---

3. *See* COLLIER'S ENCYCLOPEDIA, *supra* note 2, at 722-25.

4. *See id.* at 720.

5. *See* April Reese, *Take Back the Night*, E MAGAZINE, May 2000, at 22. *See also* Penny Jewkes, *Light Pollution: A Review of the Law*, J. PLAN. & ENV'T L., Jan. 1998, at 10 (describing the non-legal treatment of light pollution in England). "In law, however, the term 'light pollution' is not recognised, although the potential and actual consequences of inappropriate lighting may be subject to various legal responses." *Id.* Across the United States, there is a growing awareness of and reaction to light pollution as an environmental problem. *See, e.g.*, Doug Irving, *Mayor Wants to Tone Down Glow*, THE OREGONIAN, Nov. 21, 2000, at D2 (Portland, Oregon); Graeme Zielinski, *Astronomers Try to Illuminate Region on Threats to Night Skies*, WASH. POST, Sept. 25, 2000, at B1 (Washington, D.C., Virginia, and Maryland); Jay Apperson, *Dim View of Life's Bright Lights*, THE BALT. SUN, Sept. 17, 2000, at 1A (Maryland); *Dimming Light Pollution*, PROVIDENCE J.-BULL., Dec. 18, 2000, at 4B (Block Island, Rhode Island); Terri Williams, *Astronomer Fights for 'Good Sky,'* THE DALLAS MORNING NEWS, Dec. 22, 2000, at 1M (Mesquite, Texas); *Dark Crusade: Flower Mound Urged to Curb Light Pollution*, THE DALLAS MORNING NEWS, Aug. 10, 2000, at 1H (northern Texas); Chris Reinolds, *Bright Lights in Need of Dimmer Switch*, THE ATLANTA J. & CONST., Nov. 9, 2000, at 9 (Atlanta, Georgia and surrounding areas); Juliet V. Casey, *Report Critical of Night Lighting*, ALBUQUERQUE JOURNAL, Aug. 26, 2000, at 1 (New Mexico); Arthur H. Rotstein, *Stargazers Call for Lights Out*, THE DETROIT NEWS, Sept. 1, 2000, at 12 (reporting on developments in Arizona); Mary Giunca, *Glaring: We're Being Blinded By Our Own Light*, WINSTON-SALEM J., Oct. 22, 2000, at 1 (North Carolina); *Let There Be Less Light*, THE CAP. TIMES, Oct. 15, 1999, at 10A (editorial describing light pollution in Wisconsin); Joe Bauman, *Active Group in Utah Valley Scans the Skies*, DESERET NEWS, July 5, 2000, at B2 (Utah); Keri Buscaglia, *Light Pollution is a Glowing Concern*, CHI. TRIB., Jan. 31, 2001, at 4 (noting light pollution concerns of suburban Chicago residents, elected government officials and community planners).

6. *See* Joe Bower, *The Dark Side of Light*, AUDUBON, Mar.-Apr. 2000, at 94 (one of the more recent articles published about light pollution, this featured article provides general light pollution information as well as specific data related to how birds are affected by lights).

7. Michael J. Brown, *Light Pollution*, ENVTL. L. N.Y., Apr. 1998, at 49. The British Astronomical Association and other organizations in England have identified

negative effects on many things, including humans, wildlife, the environment and energy consumption.[8] As a result, some states and several cities and towns are beginning to react with legislation aimed to curb light pollution.[9]

In Part II, this Note will outline the development of electric light and its uses.[10] Despite the benefits of electric lighting, Part III will address the harmful impacts of light on our natural environment.[11] Part IV discusses the history of early and more recent common law claims of excessive lighting, argued as either trespass or nuisance claims.[12] Part V surveys the recent developments in state statutory legislation and local ordinances that have begun to alleviate the problem of excessive light.[13]

Since the results and remedies of light as trespass or nuisance claims under the common law have been inconsistent and insufficient, and state and local regulations have been relatively slow to react to the growing problem of light pollution, this Note will argue that many of the issues associated with excessive light will increase and persist without more uniform and more stringent legislation across jurisdictions. Part VI of this Note suggests that because results and remedies in the more familiar types of pollution cases argued under the common law will not likely be sufficient, and that because state and local regulations for light pollution are too inconsistent among jurisdictions, there may soon be a need for federal legislation regarding light pollution similar to some provisions of the Noise Control Act of 1972.[14] Minimally, federal legislation should (1) aim to study more adequately the effects of light pollution as it relates to humans,

---

three categories of light pollution: sky glow, glare and light trespass. *See* Jewkes, *supra* note 5, at 10. In addition to the three aforementioned problems associated with light pollution, energy waste is considered a fourth problem. *See* Brown, *supra* this note, at 49.

8. *See* Marina Murphy, *Revealing the Dark Side of Light: Artificial Light*, CHEMISTRY AND INDUS., Oct. 2, 2000, at 627 (stating that "[e]xcessive artificial lighting squanders resources, fouls the air, blocks our view of the night sky and has adverse effects on both animal and human health"). *See also* Reese, *supra* note 5, at 23; Bower, *supra* note 6, at 94-96 (highlighting some bird species, such as passerines and an endangered Hawaiian seabird, that have been negatively affected by light pollution).

9. *See, e.g., infra* note 188-89 and accompanying text. *See also* Daniel M. Nonte, *Summerfield Decides to Turn Down Lights, Keep Small-Town Feel*, GREENSBORO NEWS & RECORD, Sept. 6, 2000, at B7 (announcing a recently enacted Summerfield, North Carolina ordinance aimed at curbing light pollution).

10. *See infra* notes 17-43 and accompanying text.

11. *See infra* notes 44-94 and accompanying text.

12. *See infra* notes 95-184 and accompanying text.

13. *See infra* notes 185-262 and accompanying text.

14. *See infra* notes 263-304 and accompanying text. *See also* 42 U.S.C. §§ 4901-4918 (1999) (Noise Control Act).

wildlife, and dark skies as a natural resource,[15] and (2) provide for a more uniform system of regulating light pollution, which will be crucial as it relates to reducing harm to humans and the environment around the nation and the world.[16]

II. DEVELOPMENT OF ELECTRIC LIGHT

Society has used light in various forms including fire, torches, candles and oil lamps for many millenniums,[17] but it was not until the latter part of the nineteenth century that scientists began having success with electrical means of illumination.[18] Although the first incandescent lamps were rudimentary and unsuitable for commercial use, in 1879 Thomas Alva Edison pioneered a practical incandescent light bulb that has served as a model for modern lighting.[19] Today, the three major sources of electrical illumination are the incandescent, electric-discharge and fluorescent lamps.[20]

---

15. *See infra* notes 295-300 and accompanying text.

16. *See infra* notes 301-03 and accompanying text.

17. *See* DAN RAMSEY, EFFECTIVE LIGHTING FOR HOME & BUSINESS 4-5 (1984) (providing a brief history of lighting). *See generally* NADJA MARIL, AMERICAN LIGHTING: 1840-1940 (1989) (providing a detailed history of the various lamps used, both before and after the development of electric lighting, particularly for indoor use).

18. *See* COLLIER'S ENCYCLOPEDIA, *supra* note 2, at 719, 722. In 1802, Sir Humphry Davy first demonstrated that electric current passed through strips of metal could create incandescence from the heat created. *See id.* At that point in 1802, however, there was no efficient way to utilize this source of light. *See id.*

19. *See* MARIL, *supra* note 17, at 73-74. "October 21, 1879 is cited as the date of the invention of the first successful incandescent light bulb." *Id.* at 74. Interestingly,

> the light bulb was looked upon as a novelty item which was to be seen and admired but not to be taken seriously. The earliest electric fixtures either used broad flat shades which did nothing to shield the bulb's glare, but rather either framed the light or prevented it from escaping towards the ceiling. Some used no shades at all. The major difference in the light offered by the electric bulb was that it focused downward, rather than upwards like the gas fixtures and the kerosene fixtures generally in use.

Id. This is somewhat ironic given that we appear today to be in the predicament that too much light is, in fact, still shining upwards. *See infra* notes 44-94 and accompanying text. For an interesting account of one modern day historian's attempt to study pre-industrial age, pre-electric light nighttime hours, see Joyce Wolkomir & Richard Wolkomir, *"When Bandogs Howle & Spirits Walk": Studying the Nighttime Hours Across the Centuries, Says Historian Roger Ekirch, Sheds Light on Preindustrial Society*, SMITHSONIAN, Jan. 1, 2001, at 38. In his studies, Ekirch has discovered the meaning of a dark, light-free night to people throughout history, including the consideration of night as a separate season or a time when spirits did evil things. *See id.* at 40-41.

20. *See* COLLIER'S ENCYCLOPEDIA, *supra* note 2, at 722-24. Incandescent lighting is the primary source of light for residential use. *See id.* at 720. *See also* Energy Efficiency and Renewable Energy Network (EREN), U.S. Department of Energy,

The development of electric illumination has allowed individuals, communities, and commercial entities to conduct activities that were once restricted to daylight hours.[21] Personal use of electric lighting includes indoor and outdoor residential use, as well as interior and exterior lighting of automobiles.[22] In fact, residential consumption of electricity for lighting has increased forty percent since 1970.[23] Additionally, communities use lighting on streets and buildings to help provide for additional safety at night.[24] Lighting roadways aims to minimize dangers to drivers and pedestrians,[25] while lighting buildings, parking lots and other outdoor areas is

---

Energy Efficient Lighting, *available at* http://www.eren.doe.gov/erec/facsheets/eelight.html (last visited Apr. 19, 2002). This type of lighting is the least expensive to buy, most expensive to operate and is more inefficient than other types. *See id.* In commercial use, indoor lighting is primarily comprised of fluorescent lamps. *See id.* Outdoor lighting for safety and commercial use is largely attained through high-intensity discharge and low-pressure sodium lamps. *See id.* Many of today's lighting regulations designate what kind of light bulbs and fixtures may be used. *See, e.g.*, N.M. STAT. ANN. § 74-12-6 (Michie 2000) ("[n]o new mercury vapor outdoor lighting fixtures shall be sold or installed after January 1, 2000"); ARIZ. REV. STAT. ANN. § 49-1104 (West 1999) ("[n]o new mercury vapor outdoor light fixtures shall be installed after the effective date of this section. No replacement equipment other than bulbs for mercury vapor light fixtures shall be sold . . . after January 1, 1991 and the use of mercury vapor light fixtures is prohibited after January 1, 2011"). "A main component of new legislation is the elimination of mercury-vapor lamps – which used more energy and added more glare – vs. the newly mandated high-pressure sodium lamps . . . ." Phaedra Haywood, *New Lights Installed to Keep Skies Dark and Stars Bright*, SANTA FE NEW MEXICAN, Jan. 17, 2001, at P1.

21. *See* Rey-Barreau, *supra* note 2, at 65. "Notions about how and when to work or socialize changed. Any event that was possible in the day could now take place at night." *Id.*

> In order to demonstrate to the public that electric lighting was practical, Edison developed the first power company to provide current for lights installed in New York City in 1882. Electric incandescent light spread through the country. The demand for light encouraged the formation of power companies, and the availability of current encouraged greater use of electric light.

RAMSEY, *supra* note 17, at 5.

22. *See* Bower, *supra* note 6, at 96. "The average single-family home currently consumes 1,500 kilowatt-hours a year for lighting . . . ." *Id.*

23. *See id.*

24. *See* COLLIER'S ENCYCLOPEDIA, *supra* note 2, at 722. There are an estimated fifteen million street lights in the United States, and over sixty million worldwide. *See* Andrew Meadows, *Street-Light Device Company to Relocate to Lexington, S.C.*, THE STATE, July 9, 1999, at CS.

25. *See* Brown, *supra* note 7, at 58. However, "[m]unicipalities often install numerous street lights without clear criteria for how much lighting is warranted. A major goal of increased light is typically enhanced traffic safety, but safety can actually be compromised by glare . . . ." *Id.*

intended to increase safety at night by discouraging crime.[26]

Commercial entities also use electric lighting for various business-related purposes such as safety, advertising and emphasizing architectural features.[27] Certain entities, such as airports,[28] sports arenas,[29] parking

---

26. *See id.* at 58-59. Generally, people feel that increased light means increased safety, and as a result municipalities are often asked to provide additional lighting in the community. *See id.* at 59. However, there does not appear to be an increase of crime at night when it is dark. *See id. See also* U.S. DEPT. OF JUSTICE, CRIMINAL VICTIMIZATION IN UNITED STATES, 1998 STATISTICAL TABLES: NATIONAL CRIME VICTIMIZATION SURVEY, NCJ 181585 at Tables 59-60 (2000). (reporting that of the incidents for which there is known data, 54.6% of violent crimes occurred during the daytime while only 44.2% occurred at night, and 37.2% of property crimes occurred during the day while only 36.8% occurred at night); Reese, *supra* note 5, at 23 (A representative of the International Dark Sky Association (IDA) pointed out that "overly bright security lights can actually force neighbors to close shutters, which means that if any criminal activity" occurs, it will go unnoticed). Further, glaring lights can prevent drivers from seeing clearly, thus diminishing safety on the roads. *See* Brown, *supra* note 7, at 49.

27. *See* COLLIER'S ENCYCLOPEDIA, *supra* note 2, at 722. Ironically, Discovery Communications, the parent company of a cable television channel that airs science programs, will build its new global headquarters near Washington, D.C. with a 300-foot tower shooting a strong beam of light into the evening sky. *See* Bill Triplett, *Astronomers Fume Over Night Light*, 405 NATURE 987, 987 (June 2000) (discussing how the addition of Discovery Communications' light shooting tower will only add more light pollution to the already overlit Washington, D.C. area).

28. *See, e.g., Ivanpah Valley Airport Public Lands Transfer Act: Hearing on S. 930 Before the Subcomm. on Forests and Public Land Management of the Senate Comm. on Energy and Natural Resources*, 106th Cong., (1999) (statement of John Reynolds, Regional Director, Pacific West Region, National Park Service, U.S. Department of the Interior). Mr. Reynolds, in his testimony to persuade the subcommittee to veto a bill that would allow an airport on certain public lands, stated that, among other things,

> Another potentially significant impact to the Mojave National Preserve from the proposed airport is the deterioration of the natural quiet and the current night sky darkness that visitors enjoy at the park. Light pollution is becoming a recognized problem to many rural and rustic areas, such as the Mojave Desert. Currently, opportunities to enjoy natural quiet and the natural darkness of the nighttime are being slowly impacted by development . . . a nearby airport with runway lights, tower lights, and other lighting requirements would adversely change the dark night landscape and quiet character of the Mojave National Preserve.

Id.

29. *See, e.g.*, Hansen v. Indep. Sch. Dist., 98 P.2d 959 (Idaho 1939) (involving night baseball games at a high school field); Bd. of Educ. v. Klein, 197 S.W.2d 427 (Ky. 1946) (involving night football games at a high school field); Rhudy v. Fairfield Univ., 2000 WL 1269296, at *5 (Conn. Super. Ct. Aug. 18, 2000) (granting injunctive

lots,[30] shopping centers[31] and gas stations,[32] are notorious for using a significantly greater amount of outdoor lighting for business than other industries due to the nature of their activities. In fact, "in commercial buildings, more electricity is now used for lighting than anything else."[33]

Despite the many beneficial uses of electric light, there are some economic drawbacks. Certainly, the amount of money spent to keep the lights on is costly.[34] According to the U.S. Department of Energy, an estimated $37 billion is spent on electricity for lighting annually in the United States.[35] This figure represents one-quarter of all money spent on electric-

---

relief for individuals who lived adjacent to Fairfield University's University Field which used lights bright enough "as though a team of searchlights [were] pointed directly at [the plaintiff's home]"); Amphitheaters, Inc. v. Portland Meadows, 198 P.2d 847 (Or. 1948) (involving nighttime racing at a race track, discussed *infra* at notes 124-40, 182-84); Fuentes v. Bd. of Supervisors, 2000 WL 1210446 (Va. Cir. Ct. 2000) (involving up to fourteen lighted athletic fields to be built under special permit in a residential-conservation district).

30. *See, e.g.*, Essick v. Shillam, 32 A.2d 416 (Pa. 1943) (involving a parking lot as part of a supermarket business in a residential area that would be illuminated at night). *See also* Jay Apperson, *Dim View of Life's Bright Lights*, BALT. SUN, Sept. 17, 2000, at 1A. "The lights in the Towson Place parking lot – described as the aurora borealis of Baltimore County by one government official – have brought complaints from [local residents]." *Id. See also infra* note 157 and accompanying text (noting the outcome in *Hansen v. Indep. Sch. Dist. No. 1*).

31. *See, e.g.*, Richard Turcsik, *Blinded by the Light: Reducing Exterior Lighting Makes for Lower Energy Costs and Better Neighbors*, PROGRESSIVE GROCER, Aug. 1, 2000, at 57 (stating that "supermarket and mall operators are putting in lamps . . . which are completely inappropriate for the majority of parking lots," and "some businesses – notably restaurants, gas stations and convenience stores – are now lighting their properties with [more light than they need]" for customers to feel safe).

32. *See, e.g.*, Indian Ref. Co. v. Berry, 10 S.W.2d 630 (Ky. 1928) (concerning adjacent filling station's lights shining on plaintiff's property); Sprout v. Levinson, 148 A. 511 (Pa. 1930) (involving a filling station in a commercial area allowed to operate at night with outdoor lights); B.J.'s Wholesale Club, Inc. v. Hutchings, No. CA 99-00732, 2000 Mass. Super. LEXIS 430 (Mass. Super. Ct. Sept. 27, 2000) (involving abutters to a B.J.'s Wholesale Club store who were concerned that the addition of a gas station would bring several problems, including glaring lights).

33. Bower, *supra* note 6, at 96. In the mid-1990s, U.S. commercial buildings used more than 350 billion kilowatt hours per year in electricity for lighting, although the predominant use of light is indoors. *See* Energy Information Administration, U.S. Department of Energy, *At Home and at Work: What Types of Lights are we Using?*, *available at* http://www.eia.doe.gov/emeu/cbecs/lit-type.html (last visited Apr. 19, 2002).

34. "Lighting accounts for 20% to 25% of all electricity consumed in the United States. An average household dedicates 5% to 10% of its energy budget to lighting, while commercial establishments consume 20% to 30% of their total energy just for lighting." *See Energy Efficient Lighting*, *supra* note 20.

35. *See id.*

ity.[36] Furthermore, since many lighting structures are either poorly constructed or give off light that is aimed in the wrong direction, much of what the United States spends on outdoor lighting is wasted.[37] Estimates show that one-third to one-half of all light emitted in outdoor use shines somewhere other than on its intended target.[38] Because this light is being cast into areas where it is not needed, the United States is expending approximately one to two billion dollars a year on wasted energy.[39] Additionally, it not only costs the United States actual dollars to keep the lights on, but doing so also requires depletion of precious natural resources.[40] The Environmental Protection Agency[41] indicates that the majority of lights are illuminated through the burning of coal and oil.[42] Not only are these non-renewable energy sources, they also create atmospheric pollution.[43]

III. IMPACTS OF LIGHT POLLUTION

A. Development of Light Pollution as a New Environmental Concern

Since the 1970s, there has been a growing concern about the amount of

---

36. *See id.*

37. *See Let There Be Less Light*, CAPITAL TIMES, Oct. 15, 1999, at 10A (stating that lights are either too bright or the light is not directed where it is needed).

38. *See* Reese, *supra* note 5, at 23 (estimating that "50 percent of the light emanating from street lamps misses its intended target"); *see also* Bower, *supra* note 6, at 96 (reporting that one-third of light is wasted due to shining upward or sideways rather than down where it is needed).

39. *See* Reese, *supra* note 5, at 24 (citing International Dark Sky Association's statistics of dollars wasted due to inefficient lighting); *see also* Mary Giunca, *Glaring: We're Being Blinded By Our Own Light*, WINSTON-SALEM JOURNAL, Oct. 22, 2000, at 1.

40. *See* Bower, *supra* note 6, at 96.

41. "The EPA is the primary enforcer of all federal environmental statutes." ENVIRONMENTAL LAW HANDBOOK 55 (Thomas F. P. Sullivan ed., 14th ed. 1997).

42. *See* Bower, *supra* note 6, at 96. One of the primary uses of electricity for residential, commercial, and industrial end-users was for lighting, and in order to generate this electricity, United States utilities consumed fossil fuels. *See* U.S. EMISSIONS INVENTORY – 1999: INVENTORY OF U.S. GREENHOUSE GAS EMISSIONS AND SINKS: 1990-1997, at 2-8 (EPA 236-R-99-003, April 1999). Electric utilities generally rely on carbon intensive coal for the majority of their primary energy output, and used 87% of all coal consumed in 1997. *See id.* at 2-10.

43. *See* Bower, *supra* note 6, at 96; *see also* CHARLES E. KUPCHELLA & MARGARET C. HYLAND, ENVIRONMENTAL SCIENCE: *Living Within the System of Nature*, 143-55 (2d ed. 1989). *See generally* U.S. EMISSIONS INVENTORY – 1999, *supra* note 42, at 2-1 to 2-39 (reporting on the pollution created by energy production in the United States).

light emitted from various sources.[44] Astronomers and others have coined the phrase "light pollution" to describe the phenomenon of wasted light being cast into the atmosphere,[45] which ultimately hinders astronomical observations. This relatively new and lesser known environmental concern has been described as "the artificial light that illuminates more than its intended target area,"[46] "sky glow caused by the scattering of artificial light in the atmosphere"[47] or, in more descriptive words, "sky glow, [or] the eerie radiance that emanates from settled areas."[48] In fact, the United States generates so much wasted and upward shining light that the nation's borders and major metropolitan areas appear visible on satellite images taken at night.[49]

Light pollution is not limited to the United States.[50] Normally, 2,000

---

44. *See* Bower, *supra* note 6, at 94. "By the accounts of local and state legislators, light pollution has gained prominence in recent years and has become affixed to the larger debate over sprawl." Graeme Zielinski, *Astronomers Try to Illuminate Region on Threats to Night Skies*, WASH. POST, Sept. 25, 2000, at B04.

45. *See* Bower, *supra* note 6, at 94. *But see* James Geary, *A Thousand Points of Blight*, TIME (International Ed.), Dec. 16, 1996, at 38 (stating that, at least in Europe, excessive light from urban and industrial centers is not yet considered an official pollutant).

46. Reese, *supra* note 5, at 22.

47. H.B. 3990, 181st Gen. Ct., Reg. Sess. (Mass. 1999).

48. Bower, *supra* note 6, at 94. Referring to the night skies over the Washington, D.C. area, a Northern Virginia Astronomy Club member describes it as a "dull blush that . . . suggest[s] the aftermath of a nuclear bomb." Zielinski, *supra* note 44, at B01.

49. *See* Reese, *supra* note 5, at 23. Not only are Las Vegas' lights easily detected from above, but so are those of other cities including Boston, New York, and Los Angeles. *See id.* When NASA mapped the permanent lights around the world using satellite technology, "[t]he United States interstate highways appear as a grid system connecting the brighter dots of the major city centres including the brightly lit New York." Mark Daly, *Who Turned the Light On In Outer Mongolia?*, SCOT. DAILY REC., Oct. 25, 2000, at 18. *See also* Annette Reynolds, *Dark Crusade: Flower Mound Urged to Curb Light Pollution*, DALLAS MORNING NEWS, Aug. 10, 2000, at 1H (areas in north Texas visible from satellite photos); Reese, *supra* note 5, at 23 (quoting an observatory superintendent who noted: "When you're up in an airplane, all that light you see on the ground from the city is wasted. It's going up into the night sky. That's why you can see it."); Brown, *supra* note 7, at 58 ("any city with a population exceeding 10,000 can be identified on [a satellite] photograph"). To view various satellite images demonstrating the effect of misdirected light, visit the International Dark Sky Association's website at http://www.darksky.org/ida/sat.html.

50. For a detailed analysis of British law as it relates to light pollution, see Penny Jewkes, *Light Pollution: A Review of the Law*, J. PLAN. & ENV'T. L., Jan. 1998, 10. "[A]part from the Netherlands, England's conglomeration of brightly lit areas is more extensive than anywhere else in Europe." *Id.* In 1996, Belgium was considered the most light-polluted country in Europe, with England and the Netherlands not far behind. *See* Geary, *supra* note 45, at 38; *see also* Ross Clark, *Let's Keep the Country in the Dark*, DAILY TELEGRAPH, Oct. 14, 2000, *available at* 2000 WL 28449284 (provid-

stars would be visible in northern Europe if it were not for light pollution; unfortunately, most individuals can only see ten percent of those stars.[51] Scientists in northern Europe warn that within twenty-five years, no stars will be visible at all.[52] Similar to satellite photographs taken of the United States, there are many other areas around the globe that appear as bright white dots, revealing that light pollution is a worldwide phenomenon.[53]

In addition to astronomers, others around the United States have begun to recognize light pollution as an environmental concern.[54] Organizations within the United States and abroad, including the International Dark-Sky Association[55] and the Fatal Light Awareness Program,[56] have officially organized to study light pollution, educate others about its impact, and push for increased legislation that would reduce the "eerie radiance."[57]

---

ing additional information about light pollution in England); *Chinese Environment News: Weekly Highlights*, XINHUA ENGLISH NEWSWIRE, Feb. 11, 2001, *available at* 2001 WL 10871629 (noting that some Chinese doctors now believe light pollution may cause near-sightedness).

51. *See* Geary, *supra* note 45, at 38.

52. *See id.*

53. *See* Daly, *supra* note 49, at 18-19. Lights in Japan, England, Scotland, France and Spain appear clearly on satellite photographs, while some areas such as Antarctica and interior South American forests do not reflect any light. *See id.*

54. *See, e.g.*, Doug Irving, *Mayor Wants to Tone Down Glow*, PORTLAND OREGONIAN, Nov. 21, 2000, at D02 (mayor of rural community has recognized light pollution as a growing problem).

55. *See generally International Dark-Sky Association*, *at* http://www.darksky.org/ida/index.html (last visited Apr. 19, 2002). The mission of the International Dark-Sky Association (IDA), a membership-based nonprofit organization, is "[t]o preserve and protect the nighttime environment and our heritage of dark skies through quality outdoor lighting." *Id.* On its website, the IDA maintains a large supply of information sheets to educate the public about light pollution, as well as a resource of links to other websites that can provide additional information. *See Information Sheets*, *at* http://www.darksky.org/ida/infoshts.html.

56. *See generally Fatal Light Awareness Program*, *at* http://www.flap.org/home2.htm (last visited Apr. 19, 2002) (stating the mission of the Fatal Light Awareness Program (FLAP) is "[w]orking to preserve the lives of migratory birds in urban areas"). Information and links on this website include a listing of birds at risk, as well as mortality reports and photos from recent collisions due to illuminated structures. *See id.*

57. *See Local & Regional IDA Sections*, *at* http://www.darksky.org/ida/ida_2/info140.html (last visited Apr. 19, 2002) (providing a list of dark sky protection and light pollution prevention advisory groups and councils). Formal and informal groups have organized to reduce light pollution in countries such as Australia, Canada, Japan and several European countries. *See id.* More than seventeen state and regional chapters of IDA exist in the United States. *See id.*

B. Harmful Effects of Artificial Light and Light Pollution[58]

Scientists are just now beginning to realize the impacts of light and light pollution in various contexts.[59] Over the past thirty years, astronomers were among the first to recognize that light pollution has impaired our visibility of the heavens.[60] Because many, if not most, astronomical observatories are relatively near areas emitting some light, almost all astronomers have been affected by light pollution.[61] The nature of an astronomer's work – observing the night skies – requires a relatively light-free environment.[62] Otherwise, stars and other astronomical phenomena cannot be seen by earth-bound devices, even with the most powerful instruments.[63]

Under normal light-free conditions, there are approximately 2,500 stars

---

58. For purposes of this Note, the various negative impacts of light will focus only on how light harms humans, wildlife and the dark sky as a natural resource. However, when drafting new legislation to reduce light use, all factors should be considered by the legislature. Pollution and economic waste generated as a result of electricity used to keep lights on are also significant side effects of light pollution, but these factors will only be considered incidentally in this Note. *See, e.g.*, Murphy, *supra* note 8, at 627. However, it is important to note that a significant amount of the gases that cause global warming and acid rain are released during the production of electricity for light. *See id.* Billions of tons of these gases—including carbon dioxide, sulfur dioxide, and nitrous oxides—are released each year. *See id.* Furthermore, the equivalent of over eight million tons of coal or thirty million barrels of oil are wasted each year during the production of light that is not directed properly for its intended use. *See* Bower, *supra* note 6, at 96; *see also supra* notes 42-43 and accompanying text.

59. *See* Murphy, *supra* note 8, at 627. "The adverse effect of light pollution on stargazing is clear, but environmental impacts are only now coming into focus." *Id.*

60. *See* Bower, *supra* note 6, at 94. For example, in areas near Washington, D.C., "[t]here is no major astronomical research left in the Washington area, and the US Naval Observatory moved its heavy equipment and operations to Arizona more than 40 years ago." Triplett, *supra* note 27, at 988.

61. *See* M. Mitchell Waldrop, *Taking Back the Night: Astronomical Research Being Affected by City Lights & Satellites*, SCIENCE, Sept. 9, 1988, at 1288 (noting that the sky near an observatory some 100 kilometers from Tucson, Arizona was 6.5% brighter than normal, and an observatory 80 kilometers from San Diego, California had a sky that was 100% brighter than normal); *see also* Arthur H. Rotstein, *Stargazers Call for Lights Out*, THE DETROIT NEWS, Sept. 1, 2000, at A1 (growth near Tucson, Arizona has increased the amount of light by up to 20% in the last 10 years).

62. *See* Rotstein, *supra* note 61, at A1 (noting that "[f]or astronomers . . . keeping the sky dark is a matter of their survival").

63. *See* Johannes Andersen, *Astronomy and the Degrading Environment*, SCIENCE, Apr. 21, 2000, at 443; *see also* William C. Burton & Peter S. Gural, *Measuring the Night Sky*, SKY & TELESCOPE, June 1996, at 82.

that are visible in the night sky in the United States.[64] Unfortunately, only ten percent of Americans today can actually see the majority of these stars from where they live.[65] In suburban areas, only ten percent of the stars in the Milky Way are visible, and even less can be seen from more urban areas.[66]

Many consider the night sky and its contents a natural resource.[67] For the National Park Service (NPS)[68], this natural resource is a vital part of the activities sought by visitors to its parks.[69] To determine whether and how light pollution was impacting star gazing within the NPS parks, the

---

64. *See* Bower, *supra* note 6, at 94. Another source indicates that "[u]nder ideal conditions, one might view a night sky with over 15,000 visible stars plus the Milky Way (which includes billions of stars)." *NPCA Survey Finds Light Pollution Threatens National Park System*, U.S. NEWSWIRE, Mar. 24, 1999, *available at* 1999 WL 4635973.

65. *See* Bower, *supra* note 6, at 94.

66. *See* Reese, *supra* note 5, at 22-23 (attributing reduced visibility of stars in suburbs to factors such as overlit shopping center parking lots); *see also* Graeme Zielinski, *Astronomers Try to Illuminate Region on Threats to Night Skies*, WASH. POST, Sept. 25, 2000, at B1 (in an area forty miles from Washington, D.C., at least fifty degrees of constellations in the night skies are obscured by light pollution).

67. *See* Apperson, *supra* note 5, at 1A. Both amateur and professional astronomers are now realizing the importance of the dark sky to their night observations, and according to one amateur astronomer, "[t]he stars are as much a part of nature as the trees are." Dana Tofig, *Roswell May Dim Lights to Brighten Stars*, THE ATLANTA JOURNAL, Oct. 15, 2000, at C1. Believing this natural resource to be important enough to protect with legislation, New Mexico has passed legislation to protect the dark skies. *See* John Buting, *The Starry Night: Santa Fe's Summer Milky Way Among World's Best*, THE SANTA FE NEW MEXICAN, July 28, 2000, at C2. Even rural Vermont is trying to protect its dark skies that have yet to be impacted by light pollution. *See* Clair Wood, *Light Pollution Has Unknown Consequences to Nature*, BANGOR DAILY NEWS, Aug. 3, 2000, *available at* 2000 WL 22131087. "[Local opponents] are vigorously opposed to a proposed new prison in Springfield[, Vermont] for fear that its lights will diminish stargazing from nearby Breezy Hill. Stellafane, a celebrated festival of the stars that attracts thousands, has been held at Breezy Hill since 1926." *Id.*

68. The National Park Service is a bureau of the U.S. Department of the Interior, and its mission has been codified at 16 U.S.C. §§ 1-4 (1994). *See also* Richard J. Ansson, Jr., *Funding Our National Parks in the 21st Century: Will We be Able to Preserve and Protect Our Embattled National Parks?*, 11 FORDHAM ENVTL. L.J. 1, 5-7 (1999) (providing a brief description of the NPS mission and some of the problems it faces today, including noise pollution).

69. *See NPCA Survey Finds Light Pollution Threatens National Park System*, *supra* note 64. "Star gazing is a connection to humanity's earliest curiosity about our place in the universe. It is practically impossible to see the stars from most cities, but now, clear night vistas in our national parks are an important resource that is literally fading from sight." *Id.* (quoting NPCA President, Thomas Kiernan).

National Parks and Conservation Association (NPCA)[70] conducted a survey of parks within the NPS and found that of those responding to the survey, ninety-four percent stated that "a dark night sky is important to that park's purpose and visitor experience."[71] More importantly, approximately seventy percent of those responding to the survey indicated that light pollution is a problem in four of the five United States regions within the NPS.[72] In the Northeast, Southeast and Midwest regions of the United States, light pollution is an even greater problem due to the higher concentration of urban areas.[73]

In addition to diminishing the dark sky as a natural resource for amateur and professional astronomers alike, light pollution harms other aspects of our natural world as well. Since many species of migrating birds depend on the constellations to guide them during their migrations,[74] "artificial light will cause them to fly off course, often with disastrous results."[75] The

---

70. *See id.* The National Parks and Conservation Association, the United States' only private nonprofit citizen organization, is the leading park advocacy group in the United States that is "dedicated solely to preserving, protecting and enhancing the U.S. National Park System." *Id.*

71. *Id.*

72. *See NPCA Survey Finds Light Pollution Threatens National Park System*, *supra* note 64. "One-third . . . of these parks consider light pollution a moderately serious or very serious problem." *Id.*

73. *See id.* The NPS concentrates more of its efforts to curb light pollution in the Pacific and Rocky Mountain areas so that it can take a proactive approach in these areas where light pollution is not quite as prevalent. *See id.* As a result of the NPCA studies on light pollution, the NPS is currently collecting data that will measure the brightness of light that is interfering with stargazing at five national parks. *See* William A. Updike, *Agency Begins Light Pollution Monitoring: NPS Initiates a Program to Address the Loss of Dark Skies*, NAT'L PARKS CONSERVATION ASS'N MAG., Sept.-Oct. 2000, at 11. This and other data resulting from the survey may help park officials to work with local communities to develop legislation that will reduce light pollution. *See id. See also* Chad Moore, *Light Pollution to be Studied in Parks*, IDAHO STATESMAN, Oct. 20, 2000, at 10 (reporting that the NPS will take its studies nationwide, and will conduct research in at least eight parks in order to determine how best to deal with light pollution that is affecting the view of many stars).

74. *See* JOHN CLOUDSLEY-THOMPSON ET AL., NIGHTWATCH: THE NATURAL WORLD FROM DUSK TO DAWN, 134 (1983).

> It seems unlikely that a young bird could recognize the star pattern of the whole sky, and indeed, research suggests that only one small part of the sky is being used. Experiments with North American indigo buntings have shown that young birds need to watch the starry sky and to find out which part apparently rotates the least.

Id. at 132.

75. Wood, *supra* note 67. In 1954, 50,000 birds were killed after following an Air Force beacon and ultimately flying into the ground; in 1981, 10,000 birds similarly died when they were guided by the floodlights of a smokestack in Canada. *See id. See also* Bower, *supra* note 6, at 94-95.

Fatal Light Awareness Program (FLAP), an organization based in Toronto, Canada estimates that "at least 100 million birds are killed annually by flying into manmade structures."[76] It is believed that the birds are using the manmade light sources as a guide rather than the constellations upon which they would ordinarily rely.[77]

Migrating birds are not the only wildlife that rely on natural nighttime light sources for specific life functions[78] or that are otherwise negatively affected by artificial light.[79] Sea turtles rely on "visual brightness cues to find the sea."[80] Along a large portion of Florida's coast, a region of the continental United States where sea turtles occur in the greatest numbers,[81] there is prolific artificial beachfront lighting.[82] This artificial lighting and an accompanying "urban skyglow" from bright and concentrated inland light sources misleads nesting females from the ocean and often fatally misguides hatchlings trying to make their way back to the ocean once born.[83]

Still other wildlife can be affected by artificial light sources.[84] Moths,

---

76. Bower, *supra* note 6, at 95. Of the 1,500 birds that have crashed into Chicago's McCormick Place Exposition Center, there were 141 different species identified. *See* Wood, *supra* note 67.

77. *See* Bower, *supra* note 6, at 94.

> Night-migrating birds may also be attracted to large, bright lights and in North America huge flocks of wood warblers are sometimes killed as they swarm about lighthouses or airport searchlights, apparently unable to escape from the beam. . . . Why they should be attracted to the lights is not known, unless they too navigate by means of light-compass orientation.

CLOUDSLEY-THOMPSON ET AL., *supra* note 74, at 135.

78. *See generally* CLOUDSLEY-THOMPSON ET AL., *supra* note 74. For example, some fish species are also believed to use astral or star navigation as a means to migrate. *See id.* at 24.

79. In one of the few cases where a court has used the phrase "light pollution," the court in *City of Chula Vista* considers the least tern bird species, which might be impacted by "[n]oise and light pollution from development," as it decides whether a development project was appropriately denied. City of Chula Vista v. California Coastal Commission, 183 Cal. Rptr. 909, 921-22 (Cal. Ct. App. 1982).

80. Katherine R. Butler, Comment, *Coastal Protection of Sea Turtles in Florida*, 13 J. LAND USE & ENVTL. L. 399, 413 (1998).

81. *See id.* at 400.

82. *See id.* at 412.

83. *See id.* at 412-13. When sea turtle hatchlings do not make it to the ocean due to their reliance on the artificial light, they can die from exhaustion, dehydration and predation. *See id.* *See also* Bower, *supra* note 6, at 96 (quoting a marine scientist who stated, "[the sea turtles'] reliance on light is so strong that they'll continue heading to a light source, even if it's an abandoned fire that burns them alive").

84. *See, e.g.*, Steve LaRue, *Biologists Survey How Well Local Reptiles and Amphibians Are*, THE SAN DIEGO UNION-TRIB., Dec. 27, 2000, at F1. The California

one of the great nighttime pollinators in nature,[85] are also attracted to lights.[86] Wasted time at these light sources creates a loss in energy that would otherwise be used to attract a mate.[87] For other animals, disrupted foraging, reproduction, circadian rhythms, and hormone levels may all find their causes to be in excessive artificial light in their natural environment.[88] As the awareness of light pollution grows, the need for additional research

---

glossy snake and the western long-nosed snake, both nocturnal species, have been found to be "missing from coastal areas where there is a lot of light pollution from cars and houses." *Id.* Scientists believe that predators may have been better able to catch these prey, once very common in the region during the 1920s and 1930s, due to the abundance of light in the snakes' natural habitats. *See id.*

85. *See* CLOUDSLEY-THOMPSON ET AL., *supra* note 74 at 140-41. "Moths are the main agents of nocturnal pollination in the temperate regions." *Id.* at 140.

86. *See id.* at 135. Some animals use "light-compass orientation" as a way to navigate at night. *Id.* By keeping a light source, such as the sun or moon, "in the same part of their field of vision, [animals] maintain a constant angle to it as they move." *Id.*

> It is this which causes moths to collide with street lamps and lighted windows at night, for they tend to apply light-compass navigation to any source of light . . . . [If the moth] orientates to a nearby source of light, it must correct its angle as it goes past the light and consequently begins to move in a circle around it, then in a steadily diminishing spiral until it finally strikes the light.

Id. See also Geary, *supra* note 45, at 38. Although the death's-head hawk moth's normal nocturnal migratory route stretches from Africa to Scandanavia, this species has been observed to fly off course over France and Belgium. *See id.* This diversion, believed to be caused by the moths' attraction to brightly lit sports complexes and shopping centers, often results in death to those moths flying too close to the heat generated by the intense lights. *See id. See generally*, Kenneth D. Frank, *Impact of Outdoor Lighting on Moths*, 42 J. LEOPIDOPTERISTS' SOC'Y 63-93 (1988) *excerpt available at*, http://www.darksky.org/ida/info109.html (last visited Apr. 19, 2002).

87. *See* Bower, *supra* note 6, at 96. This attraction to light may be what is causing some species' populations to decline in the United States. *See id.* Moths may lay their eggs in areas where they will not survive. *See* Wood, *supra* note 67. *See also* Kathleen Daminger, *Shedding Light on Moth Behavior*, LANCASTER NEW ERA, July 12, 2001, at 11 (noting that male moths waste precious time when attempting to find female mates due to their erroneous reliance on street lights rather than the moon or bright stars).

88. For example, firefly populations are facing reduction due to the interference of light pollution with their mating patterns. *See* Susan Okie, *A Spotlight on Fireflies*, ORLANDO SENTINEL, July 22, 2001, at G6. Photoperiodic behavior, abnormal responses to artificial light, can affect not only migrating patterns, but also reproduction and foraging behavior. *See* Wood, *supra* note 67. Hormonal levels of animals may be altered by exposure to light and "'anything that alters the hormonal levels will bring enormous changes.'" *Id. See also* Geary, *supra* note 45, at 38 (stating that research conducted by the French Astronomical Society has indicated that "exposure to powerful floodlights can interfere with the circadian rhythms of plants and disorient migratory birds and insects"). *See generally* CLOUDSLEY-THOMPSON ET AL., *supra* note 74, at 12-22 (discussing the role and mechanics of circadian rhythms in wildlife).

and findings indicating the potentially adverse effects on wildlife will also increase.[89]

Scientists also believe that humans are adversely affected by excessive light as well.[90] Drivers may be temporarily blinded by glare given off by lights and thus prone to accidents.[91] Studies also indicate that light present during sleeping hours may disrupt internal clocks and hormone levels.[92] Sleeping patterns themselves may have changed over the course of the post-Edison era due to the increase of artificial light.[93] Because this is a relatively new area, the research available on human effects of light pollution is minimal, however it appears persuasive enough that some negative effects may result.[94]

---

89. *See* Aili Petersen, *Night Lights*, AM. SCIENTIST, Jan. 1, 2001, at 24.

> For all the buzz, little research has been done on other ecological effects of light pollution – for instance, its impact under water. One aquatic ecologist, however, is currently investigating the effects of light pollution in lakes . . . . [T]hese areas, along with coastal waters, are at higher risk than other habitats in developed areas because they are unshielded and openly exposed to light.

Id. One research study conducted at five lakes in urban and rural areas indicated that the migration of Daphnia, a species of freshwater zooplankton, is significantly impacted by light pollution. *See id.* at 24-25.

90. *See* Wood, *supra* note 67. Two separate scientific journals have recently published articles that identify humans as being more sensitive to light than what was previously believed, and that human circadian rhythms can be disrupted by exposure to light. *See id.*

91. *See* Reese, *supra* note 5, at 23.

92. *See* Bower, *supra* note 6, at 95. Disruption of circadian rhythms can create changes in blood and urine chemistry, behavior and melatonin production. *See id.* Furthermore, some scientists believe that "chronic disruptions in melatonin production – such as those caused by sleeping in a room that's bathed in a streetlight's glow – might contribute to the development of 'hormone-related' cancers, including breast cancer." *Id. See also* N.A. Kerenyi et al., *Why the Incidence of Cancer is Increasing: The Role of 'Light Pollution,'* 33 MED. HYPOTHESES 75 (1990) (hypothesizing that the rapid growth rate of cancer incidences correlates to the increase in exposure to light, which reduces melatonin production, over the last 100 years).

93. *See* Wolkomir & Wolkomir, *supra* note 19, at 38. While most people would assume that prior to the invention of artificial light people slept continuously throughout the night, the contrary might be true. *See id.* A study done by a National Institute of Mental Health psychiatrist indicated that when "[d]eprived of artificial light, [research volunteers] reverted to the preindustrial pattern" of segmented sleep. *Id.* It is believed that "[w]ithout the stimulus of artificial light, people secrete more prolactin, a pituitary hormone that seems to promote a state of quiet restfulness," and prior to the industrial age, "people slept differently because they had less artificial light recalibrating their hormone production." *Id.*

94. *See* Bower, *supra* note 6, at 95.

IV. AN OVERVIEW OF COMMON LAW CLAIMS INVOLVING LIGHT

Although scientists are just now discovering the harmful physiological and psychological impacts of light on humans and wildlife, other unwanted and disturbing aspects of light existed well before the modern problem of light pollution existed.[95] In fact, despite the much-awaited discovery of electric illumination and its benefits, individuals were relatively quick to react with lawsuits to enjoin light cast by others upon their property,[96] long before anyone recognized "light pollution" as a distinct environmental problem.[97]

Using common law claims in all types of pollution cases has been a tool for many landowners both past and present.[98] Perhaps the two most frequent common law claims brought in environmental litigation cases have been nuisance and trespass.[99] While there are differences between these

---

95. *See, e.g.*, Akers v. Marsh, 19 App. D.C. 28 (1901) (involving one of the first litigated complaints about a non-electric source of light, torch lamps, shining onto plaintiff's property).

96. *See* Shelburne, Inc. v. Crossan Corp., 95 N.J. Eq. 188 (1923) (involving an early instance of electric light shining on to an adjoining property). *See also* 58 AM. JUR. 2D *Nuisances* § 153 (1989) (listing several cases where light cast upon another's property were brought into court under the theory of nuisance). *See generally* K.J. Roberts, Annotation, *Casting of Light on Another's Premises as Constituting Actionable Wrong*, 5 A.L.R. 2d 705 (1949) (providing a comprehensive overview of early twentieth-century cases involving various sources of light that were complained about in court).

97. *See supra* notes 44-94 and accompanying text for a discussion on the problem of light pollution.

98. *See* Frona M. Powell, *Trespass, Nuisance and the Evolution of Common Law in Modern Pollution Cases*, 21 REAL ESTATE L.J. 182, 183 (1992). Despite all the current statutory environmental law, "the common law provides the only legal recourse to individual plaintiffs in the environmental pollution cases." *Id.* at 183 n.2. "Common-law theories protecting interests in real property focus on two essential questions: First, the *nature* of the legal interest protected, and second, the kind of *invasion* of interest which gives an owner or possessor the right to a legal remedy." *Id.* at 184 (emphasis added). *See also* Arnold W. Reitze, Jr., *A Century of Air Pollution Control Laws: What's Worked; What's Failed; What Might Work*, 21 ENVTL. L. 1549, 1554-57 (1991) (providing a brief overview of the common law claims of nuisance and trespass generally in air pollution cases); Bradley v. American Smelting and Refining Co., 709 P.2d 782 (Wash. 1985) (cause of action in trespass due to copper smelter's deposits of particulate arsenic and cadmium falling on another's land); Borland v. Sanders Lead Co., 369 So. 2d 523, 529 (Ala. 1979) (cause of action in trespass against lead smelter for emissions of lead particulates and sulfoxide deposits on another's land); Martin v. Reynolds Metals Co., 342 P.2d 790 (1959) (cause of action in trespass against aluminum manufacturer causing fluoride gases and particulates to fall on another's property).

99. *See* Roger Meiners & Bruce Yandle, *Common Law and the Conceit of Mod-*

two causes of action,[100] the theories behind both are considered related,[101] and, particularly in pollution cases, these common law theories stem from property and tort law.[102] Despite the predominance of legislation as a

---

ern Environmental Policy, 7 GEO. MASON L. REV. 923, 926 (1999). "Before federal environmental statutes became dominant, the common law doctrines most frequently relied upon in environmental litigation were nuisance and trespass." *Id.* "The legal theories of nuisance and trespass were developed long before the birth of [the United States]" and can be traced back to twelfth century England. G. Nelson Smith, III, *Nuisance and Trespass Claims in Environmental Litigation: Legislative Inaction and Common Law Confusion*, 36 SANTA CLARA L. REV. 39, 41 (1995). Since that time, however, the theories have evolved and been applied by the courts in a much different way. *See* PATTON, BOGGS & BLOW, ENVIRONMENTAL LAW HANDBOOK 25-26 (BNA 1994). *See also* Powell, *supra* note 98, at 183 (stating that in addition to trespass and nuisance, the common law claims of negligence and strict liability have also been used in environmental pollution cases); DANIEL P. SELMI & KENNETH A. MANASTER, STATE ENVIRONMENTAL LAW 2-1 (Release #8, Nov. 1997) (noting that the public trust doctrine is an additional common law doctrine that is frequently used in environmental law). *But see* THOMAS M. HOBAN & RICHARD O. BROOKS, GREEN JUSTICE: THE ENVIRONMENT AND THE COURTS 11 (2d ed. 1996) (stating that "today a purely common-law cause of action is rare").

100. "Nuisance law traditionally protected the right of a landowner or occupier to the use and enjoyment of property, while trespass provided compensation for unpermitted physical intrusion upon property rights." SELMI & MANASTER, *supra* note 99, at 2-2 to 2-3. Furthermore, "[a] claim of trespass contemplates actual physical entry or invasion, whereas nuisance liability arises merely by virtue of an activity which falls short of tangible, concrete invasion but interferes with the use and enjoyment of land." Smith, *supra* note 99, at 54. Just because an individual may have a claim under nuisance, it does not follow that he or she will simultaneously have a claim under trespass. *See id.* at 55.

101. *See* Meiners & Yandle, *supra* note 99, at 935. "Trespass, when invoked in pollution cases, is a common law theory closely related to nuisance." *Id.* "A primary distinction between the actions of trespass and nuisance is the difference in the legal right protected." Powell, *supra* note 98, at 185. "[Trespass and nuisance] can be distinguished by comparing the interest invaded; an actionable invasion of a possessor's interest in the exclusive possession of land is a trespass; an actionable invasion of a possessor's interest in the use and enjoyment of his land is a nuisance." Martin v. Reynolds Metals Co., 342 P.2d 790, 792 (Or. 1959).

102. *See* BRUCE YANDLE, COMMON SENSE AND COMMON LAW FOR THE ENVIRONMENT: CREATING WEALTH IN HUMMINGBIRD ECONOMIES 88 (1997).

> The common law-theory that applies to pollution is a part of the law of property and torts, a body of law that protects life and property from harm caused by others. Based on rights, the common law emerges in rulings announced by judges on a case-by-case basis. The law is formed from specific controversies, claims for actual damage, and requests for injunctions against the threat of damage. The rules of tort law that relate to the environment are found in a component of common law that deals with nuisance and trespass. The latter property right violation is associated with uninvited physical invasion of property, while the former relates to harms, like odors . . . , that do not reflect a physical crossing of a

mode of dealing with environmental protection, these common law doctrines continue to have a place in environmental protection.[103] Before the phrase "light pollution" was coined, plaintiffs complaining about intruding light reached the courts by bringing a claim against a neighboring individual or entity in one of two ways: light as nuisance or light as trespass.[104]

As Part IV will reveal, neither of these claims were very successful for aggrieved plaintiffs seeking to dim a neighbor's light. Given the modern problem of light pollution and its *many* sources of light, as well as the nature of the pollutant itself, there may need to be another more reliable and effective source of relief for both humans *and* wildlife.

A. Nuisance in Environmental Pollution and Light as Nuisance Cases

Landowners seeking damages for pollution to their property may bring a nuisance claim[105] since a property owner's right to use and enjoy his property is protected by the law of nuisance.[106] While there are two types of nuisance claims, private and public,[107] both "relate to the unreasonable

---

property boundary.

Id.

103. *See* SELMI & MANASTER, *supra* note 99, at 2-2. "Statutes have not totally eclipsed the common law doctrines for several reasons," including the incorporation of common law concepts into the statutory scheme, the use of common law analysis by courts in interpreting a statute, the inadequacy of statutes to deal with a particular environmental problem, insufficiency of statutory remedies, and the tendency of courts to be receptive to common law actions. *Id.* at 2-6 to 2-9. *See generally* THE COMMON LAW AND THE ENVIRONMENT: RETHINKING THE STATUTORY BASIS FOR MODERN ENVIRONMENTAL LAW (Roger E. Meiners & Andrew P. Morriss eds., 2000) (providing several articles on the use and benefits of the common law over statutory legislation in protecting the environment).

104. *See* Amphitheaters, Inc. v. Portland Meadows, 198 P.2d 847, 850 (Or. 1948) (light as trespass claim); Shelburne, Inc. v. Crossan Corp., 95 N.J. Eq. 188, 191 (1923) (light as nuisance claim). *See also infra* notes 124-40, 182-84 and accompanying text for a discussion of *Amphitheaters, Inc.*

105. *See* Powell, *supra* note 98, at 183. *See also Amphitheaters, Inc.*, 198 P.2d. at 851 (plaintiff's second assignment of error in his appeal was that the trial court erred in denying his claim of nuisance to go forward).

106. *See* Powell, *supra* note 98, at 188.

107. *See* Ronald J. Rychlak, *Common-Law Remedies for Environmental Wrongs: The Role of Private Nuisance*, 59 MISS. L.J. 657, 658-61 (1989). Private nuisance "is an unreasonable interference with the plaintiff's enjoyment of his property," while public nuisance is an "unreasonable interference with the right of, or a threat to, the general public." *Id.* at 658, 660. In many jurisdictions, public nuisances are statutory. *See id.* at 658. "[P]rivate nuisance is an invasion of another's interest in the private use and enjoyment of land." PATTON, BOGGS & BLOW, *supra* note 99, at 26. "The invasion will be actionable only if (1) the conduct giving rise to the invasion is tortious and (2) an interest associated with the use and enjoyment of the plaintiff's

interference with the use and enjoyment of land."[108] In a private nuisance case, the plaintiff must show that the defendant, typically a neighboring individual or entity,[109] invaded the plaintiff's interest in his property either intentionally and unreasonably, negligently, recklessly or, in the case of abnormally hazardous activities, under the theory of strict liability.[110] Furthermore, the invasion must be substantial,[111] but this has not been interpreted to mean that the invasion must have a physical or tangible impact.[112] Once the plaintiff has met his or her burden, "the burden shifts to the defendant to establish that its use was reasonable or the interference inconsequential," but the defendant's conduct will not be excused merely because it was "necessary, modern or efficient."[113] The courts essentially use a balancing test to determine whether the defendant's activities are sufficient

---

land has been invaded." *Id.* (relying on the RESTATEMENT (SECOND) OF TORTS §§ 821, 822). "A public nuisance . . . is an interference with a right of the public at large, which is not limited to the use and enjoyment of land." *Id.* "[T]he plaintiff must prove both tortious conduct and an actionable invasion resulting therefrom." *Id.* at 27. *See also* Smith, *supra* note 99, at 50-53 (describing the difference between private and public nuisance).

108. Smith, *supra* note 99, at 50.

109. "Often the environmental plaintiff is an individual or small group, and the defendant is a large corporate entity." Rychlak, *supra* note 107, at 661.

110. *See* Smith, *supra* note 99, at 50. "A private nuisance cause of action arises when the injury inflicted either diminishes the value of [plaintiff's] property, continually interferes with the power or control of that property, or causes a material disturbance or annoyance to the person in the use or occupation of that property." *Id.* "A private nuisance is a nontrespassory invasion of another's interest in the private use and enjoyment of land." RESTATEMENT (SECOND) TORTS § 821D (1979).

> One is subject to liability for a private nuisance if, but only if, his conduct is a legal cause of an invasion of another's interest in the private use and enjoyment of land, and the invasion is either
>
> (a) intentional and unreasonable, or
>
> (b) unintentional and otherwise actionable under the rules controlling liability for negligent or reckless conduct, or for abnormally dangerous conditions or activities.

RESTATEMENT (SECOND) TORTS § 822 (1965). *See also* Rychlak, *supra* note 107, at 674 (stating that a defendant's mere knowledge that an interference with plaintiff's enjoyment of the land may be sufficient to establish intent).

111. *See* PATTON, BOGGS & BLOW, *supra* note 99, at 27. "An invasion will not be actionable in nuisance unless it is substantial, i.e., if it offends normal persons in that particular locality." *Id. See also* Rychlak, *supra* note 107, at 676-77. "[M]inor interferences of a short duration will not amount to a nuisance." *Id.* at 676.

112. *See* PATTON, BOGGS & BLOW, *supra* note 99, at 27. "Unlike trespass, nuisance does not require a physical invasion of the property but occurs if a condition is maintained on the defendant's land that interferes with the plaintiff's use and enjoyment of the plaintiff's property." Powell, *supra* note 98, at 188-89.

113. *See* Rychlak, *supra* note 107, at 678.

to constitute a nuisance to the plaintiff's enjoyment of his or her property.[114] The most common forms of environmental nuisance have included noise pollution, dust, smoke, vibrations and odors,[115] and remedies for the plaintiff can include monetary and equitable injunctive relief.[116]

It was not until the early twentieth century that courts began hearing nuisance claims from individuals complaining of light being emitted from an adjacent property onto theirs.[117] In *Shelburne, Inc. v. Crossan*,[118] some of the plaintiff's hotel rooms were subjected to light shining in from the defendant's property upon which a large sign emitting the light was af-

---

114. *See* Reitze, *supra* note 98, at 1555. *See also* Andrew Jackson Heimert, *Keeping Pigs Out of Parlors: Using Nuisance Law to Affect the Location of Pollution*, 27 ENVTL. L. 403, 410-12 (1997) (describing the shifting view of the courts over the years to at least recognize some sort of relief for plaintiffs even if the defendant's pollution-creating activity was considered more socially valuable).

115. *See* Rychlak, *supra* note 107, at 660-61 & nn. 17-21. Arguably, some of these types of pollution, and perhaps adding light to the list, would require a showing of intentional interference by the defendant because in some cases they do not physically damage the land. *See id.* at 674-75 (stating that "when there is physical damage to the plaintiff's land, water or improvements, even unintentional or accidental interferences are generally actionable"). *But see id.* at 676 (stating that the harm caused by defendant "is not limited to physical harm or injury" and the standard is whether a reasonable person would find the harm to be substantial).

116. *See* Reitze, *supra* note 98, at 1555-56.

> Nuisance, along with the doctrine of trespass, allows not only the awarding of monetary damages, but also the use of equitable relief such as injunctions or abatement orders. Because equitable relief is available, the private nuisance doctrine is the common law remedy that can be used to directly abate pollution.

Id. at 1555. In *Essick v. Shillam*, residents of a community sought to enjoin the defendant from building a supermarket and adjoining parking lot because they feared that both would constitute a nuisance despite being in conformity with the zoning and building guidelines. *See* Essick v. Shillam, 32 A.2d 416, 417 (Pa. 1943). Despite the defendant's stated intent to "illuminate the premises brightly, inside and out, and the parking lot as well," the court held that there was not yet a nuisance. *Id.* at 419. However, if the resulting illumination did, in fact, become objectionable or produce a sleep-disturbing glare, the court stated "it can be adjusted to satisfy plaintiffs' objections, or, if necessary, enjoined at that time." *Id.* Today, such plaintiffs might possibly still oppose such building projects, but they would probably have more local regulations to ensure that specific lighting requirements are met by the developer. *See* B.J.'s Wholesale Club, Inc. v. Hutchings, No. CA 99-00732, 2000 Mass. Super. LEXIS 430 (Mass. Super. Ct. Sept. 27, 2000).

117. *See* Shelburne, Inc. v. Crossan Corp., 95 N.J. Eq. 188 (1923). The plaintiff in the case was the owner of a hotel in Atlantic City, New Jersey. *See id.* The defendants were an adjacent property owner, and a company renting rooftop space on a building on the property for the purpose of erecting an advertising sign. *See id.* The sign consisted of 1,084 15-watt lights, six 100-watt lights and twenty-eight 75-watt lights. *See id.* at 189.

118. *See id.*

fixed.[119] Because the lights were turned on from dusk until midnight, and on at least one occasion until after midnight, the illumination disturbed guests of the hotel and allegedly lowered the value of the rooms.[120] In its holding, the court reasoned "that the amount of light radiating from the sign does illuminate or 'light up' many of the rooms facing it, and in some instances, at least, to such an extent as to be objectionable to the guests."[121] The court found that the defendant's light is better considered a nuisance,[122] and provided relief for the plaintiff.[123]

There are few cases aside from *Shelburne, Inc.* that have light emissions as the *exclusive* factor declared as a nuisance by the plaintiff.[124] The court in *Amphitheaters, Inc.* reasoned that although light emitted from a neighboring race track onto a drive-in theater's property might not constitute trespass as the plaintiff had claimed,[125] there may be a legitimate claim in nuisance.[126] In *Amphitheaters, Inc.*, the plaintiff was a drive-in theater operator whose theater abutted a parcel of land that included a one-mile outdoor racing track.[127] That the track would be lighted for night racing

---

119. *See id.* at 189.

120. *See id.* at 189-90.

121. *Id.* at 190.

122. *See id.* at 191. "There can be little, if any, doubt that light radiating from lamps of the intensity, and, when placed in the position of those in the sign in question, may become a nuisance, if it (the light) materially interferes with the ordinary comfort, physically, of human existence." *Id.* (citation omitted).

123. *See id.* at 192. The defendant was ordered to limit the sign's hours of operation to the hours before midnight. *See id.*

124. *See, e.g.*, Amphitheaters, Inc. v. Portland Meadows, 198 P.2d 847 (Or. 1948) (discussed *infra* at notes 126-40, 182-84); Shepler v. Kansas Milling Co., 278 P. 757 (Kan. 1929) (reflected sunlight from defendant's grain storage tanks). *See also* Dean N. Alterman, Comment, *Reflected Sunlight Is a Nuisance*, 18 ENVTL. L. 321, 330-31 (1988) (noting a few cases, including *Shepler*, that have dealt with reflected sunlight as a nuisance). For cases that have light as at least one factor in a nuisance claim, see 58 AM. JUR. 2D *Nuisances* § 153 (1989) (listing several cases where light cast upon another's property was brought into court under a nuisance claim); K.J. Roberts, Annotation, *Casting of Light on Another's Premises as Constituting Actionable Wrong*, 5 A.L.R.2d 705 (1949) (providing a comprehensive overview of early twentieth-century cases involving various sources of light that were complained about in court).

125. *See Amphitheaters, Inc.*, 198 P.2d at 850.

126. *See infra* notes 182-84 and accompanying text. *See also* John-Mark Stensvaag, *State Regulation of Nuclear Generating Plants Under the Clean Air Act Amendments of 1977*, 55 S. CAL. L. REV. 511, 536 n.131 (1982) (stating that *Amphitheaters, Inc.* "address[es] a nuisance which might today be called 'light pollution'").

127. *See Amphitheaters, Inc.*, 198 P.2d at 848. Amphitheaters, Inc.'s lease on the property allowed the construction of a drive-in outdoor movie theater, but also provided that the theater's activities could not interfere with the race track which was

was featured extensively in local newspapers.[128] The floodlights were generally aimed at the track, but there was substantial evidence that the emitted light spilled onto the theater's premises and had "a serious effect on the quality of pictures shown on the screen."[129] As a result, the theater operator brought a claim of trespass against the race track operator.[130]

The court, while acknowledging that the line between trespass and nuisance claims was blurry, held that the theater operator's claim was governed by the law of nuisance and not trespass.[131] In analyzing this case of first impression, the court found that nuisance cases tended to fall in one of four categories,[132] and that the plaintiff's claim may possibly be analogous to only the first category.[133] However, the court found itself in a difficult position when it was time to analyze why light should be considered parallel to the cases involving smoke, odors and flies; "no one can contend that light is inherently harmful to persons in the ordinary enjoyment of prop-

---

located on the same property. *See id.* Some of the theater promoters knew that the neighboring race track would be illuminated to some extent, but regardless, the construction of the theater was completed fifteen days before the completion of the race track. *See id.*

128. *See id.* The racetrack was equipped with "approximately 350 1500-watt lights" mounted and clustered on 80-foot poles around the track. *See id.* at 850.

129. *Id.* at 850. Not only did the light reduce the quality of movie pictures, but the theater suffered a financial loss due to this invasion of light. *See id.* There was evidence that the light cast upon the theater's property was equivalent to that of a full moon. *See id.*

130. *See infra* notes 182-84 and accompanying text. On his appeal from the circuit court ruling, the plaintiff contended that the defendant's light was, in fact a trespass, and "error [should be] assigned by reason of the failure of the court to submit to the jury the question of trespass." Amphitheaters, Inc. v. Portland Meadows, 198 P. 2d 847, 850 (Or. 1948).

131. *See Amphitheaters, Inc.*, 198 P. 2d at 850. "The mere suggestion that the casting of light upon the premises of a plaintiff would render a defendant liable without proof of any actual damage, carries its own refutation. Actions for damages on account of smoke, noxious odors and the like have been universally classified as falling within the law of nuisance." *Id.*

132. *See Amphitheaters, Inc.*, 198 P.2d at 851. The court delineated the four classes of nuisance as follows:

> (1) Cases involving harm to human comfort, safety or health by reason of the maintenance by a defendant upon his land of noxious or dangerous instrumentalities causing damage to the plaintiff in respect to legally protected interests of the plaintiff in his land. (2) Cases involving illegal or immoral practices, most of them being public as distinct from private nuisances [i.e. gambling, abortions, lotteries]. (3) Cases involving obstructions to streets, public ways, common rights, access to property and the like. (4) Cases involving damage to the land itself, as by flooding.

Id. The court also listed several Oregon cases that fell into the first class. *See id.*

133. *See id.* The court likened previous cases involving noxious odors, ashes, smoke, dynamite and stream pollution to the type of case being brought by the theater operator. *See id.*

erty."[134] Instead, the court decided to use a balancing test of weighing the plaintiff's complaint against the defendant's activities[135] and looked to other jurisdictions to determine whether the race track's lights did indeed constitute a nuisance.[136]

After much analysis of other courts' decisions on similar issues, the Oregon Supreme Court ultimately held that the defendant's race track lights cast upon the plaintiffs did not constitute a nuisance.[137] The court reasoned that a drive-in theater is of a sensitive nature, not one of the ordinary person,[138] and in analogous cases finding light as nuisance, light was but one of many other factors comprising the nuisance.[139] In its dicta, the court indicated its willingness to classify light as nuisance in other more appropriate cases.[140]

Hildebrand v. Watts is another, more recent case that also indicates a

---

134. *Id.* at 851-52. In previous cases, the factor that was held to be a nuisance was held to be both inherently harmful and an unreasonable and substantial interference with the plaintiff's ordinary use and enjoyment of his property. *See id.* In fact, in another part of the opinion the court states that the "case [at bar] differs fundamentally from other cases, all typical cases of nuisance, in that light is not a noxious, but is, in general, a *highly beneficial element*." *Id.* at 858 (emphasis added).

135. *See id.* at 852. The interference with the use and enjoyment of land must be substantial and unreasonable, *and* "whether a particular annoyance or inconvenience is sufficient to constitute a nuisance depends on its effect upon an ordinarily reasonable man, that is, a normal person of ordinary habits and sensibilities." *Id.* "In determining whether the conduct of the defendant was unreasonable under the facts of this case, the court found that neither party could claim any greater social utility in its conduct than the other." Powell, *supra* note 98, at 191.

136. *See* Powell, *supra* note 98, at 189-91. "The Oregon Supreme Court noted that no Oregon decision had ever held that the casting of light in any quantity or form upon the land of another gave rise to a cause of action under any legal theory." *Id.* at 190. The court, not finding any relevant caselaw in its own jurisdiction, looked to other decisions to see whether light had ever constituted a nuisance. *See Amphitheaters, Inc.*, 198 P.2d at 854-58.

137. *See* Amphitheaters, Inc. v. Portland Meadows, 198 P.2d 847, 858.

138. *See id.* at 857.

139. *See id.* The court did note that in *Shelburne, Inc.*, light was the only factor in the nuisance, but differentiated that case from the one at bar because unlike the theater operator, a hotel dweller trying to sleep at night "was with the normal and ordinary sensibilities of dwellers in the hotel, and with the ordinary use of property." *Id.*

140. *See id.* at 858.

> We do not say that the shedding of light upon another's property may never under any conditions become a nuisance, but we do say that extreme caution must be employed in applying such legal theory. The conditions of modern city life impose upon the city dweller and his property many burdens more severe than that of light reflected upon him or it.

Id.

court's unwillingness to classify light as nuisance in certain instances.[141] In this case, the plaintiffs and defendants were neighbors in a residential area of Connecticut.[142] The plaintiffs objected to the defendants' placement of a security and recreational light on the roof of their house because when it was in use at night, it illuminated the plaintiff's driveway and parts of the house.[143] The court, not convinced by the plaintiff's arguments, found that they had not met their burden of proving by a preponderance of the evidence that the light was a sufficiently impermissible interference so as to be enjoined.[144] The court reiterated that a "balancing of equities is involved in determining whether an interference with one's use of property is such an impermissible interference that it should be enjoined as a nuisance."[145] In this case the court held,

> absent stronger evidence than has been presented here – for example, that sleep is interfered with – the plaintiffs have not made out a case, by a preponderance of the evidence, that the intrusion of light onto their property unreasonably outweighs the legitimate interests of the [defendants] in lighting their own property for purposes of safety and for the recreation of their children.[146]

---

141. *See* Hildebrand v. Watts, No. 13988, 1997 Del. Ch. LEXIS 32, at *20-21 (Del. Ch. Feb. 18, 1997).

142. *See id.* at *2.

143. *See id.* at *4. The defendants "testified that there is crime in their neighborhood and that they placed the light on their house in an effort to protect themselves against vandalism and theft." *Id.* The plaintiffs contended that the light shined into their kitchen, bathroom and screened-in porch during the period from dusk until around 9:30 p.m. or later and interfered with their enjoyment of these areas. *See id.* at *12-13. They also alleged that the light was frequently turned on and off by defendant in order to harass them. *See id.* at *13.

144. *See id.* at *14-21. The court, upon review of the videotaped evidence, was not persuaded that the light was unreasonable or too obtrusive in the kitchen and bathroom. *See id.* at *15. Furthermore, the court did not find that there was any injury to the plaintiffs from the light being cast upon the driveway and front lawn since "they do not use these places for relaxation." *Id.* at *16. In terms of the back porch, the court stated that the plaintiffs were also not entitled to relief because no claim as to the specific light fixture that was causing such light was before the court (only the roof light was before the court), and because there was not sufficient evidence to determine whether the defendants were really turning the light on and off purposely. *See id.* at *16-18.

145. *Id.* at *19. In referring to a previous case that cited the *Shelburne, Inc.* decision, the court noted that "generally courts have been hesitant to enjoin the use of outside lights." *Id.* at *19.

146. *Id.* at *20-21. *But see* Rhudy v. Fairfield Univ., No. CV-990368012S, 2000 WL 1269296, at *1, *5 (Conn. Super. Ct. Aug. 18, 2000) (court found that the activities being conducted on University Field were sufficient to constitute a nuisance that warranted injunctive relief). In *Rhudy*, the plaintiffs, four families living adjacent

While *Shelburne, Inc.*, *Amphitheaters, Inc.*, and *Hildebrand* dealt with light being cast upon a plaintiff's property by one other defendant, the problem of light pollution is created by *many* sources of light, some of which are not located on adjacent property.[147] To date, there have been no cases brought under the theory of private nuisance to combat light pollution, but predictably, such plaintiffs would not prevail. First, the nature of the interest protected by nuisance is "the private use and enjoyment of land."[148] It is hard to imagine that, given the unsuccessful claims of the drive-in theater operator combating a significant source of adjacent light, or the neighbor complaining of a security light shining in an area not traditionally considered an area used for relaxation, a particular individual or group would be successful in combating multiple defendants who are creating a more diffuse "sky glow." Further, it is likely that the glow of light pollution that would be complained about would be produced by multiple and distant defendants, thus creating more legal difficulties in bringing a lawsuit.[149] Courts have indicated that "[t]he hypersensitive user will not recover for the actual harm he suffers, but only (if at all) for the harm that a normal person would suffer from the same invasion."[150] For example, astronomers, who require the darkest of skies, may find themselves in the same position as drive-in theater owners.[151] Finally, since light pollution can be found without other nuisances such as noise, it is uncertain whether

---

to and abutting University Field at Fairfield University, claimed noise and light from the field were creating a nuisance. *See id.* at *2-5. The lighting along University Field consisted of "eight light towers, four on each side of the field, with clusters of floodlights . . . [with each floodlight consisting of] a 1500-watt metal halide lamp" which, according to one expert, is the biggest lamp source of this kind available in commerce. *Id.* at *4. It created a light on some plaintiffs' property up to 30 times brighter than a full moon and reached various rooms, including bedrooms. *See id.* The court issued an injunction that lights be turned off by 7:00 p.m. during the week, by 5:00 p.m. on weekends and light fixtures could not directly shine towards plaintiffs' property. *See id.* at *6.

147. *See supra* Part III.A.

148. Falcone, III & Utain, *infra* note 162, at 68.

149. Even if there were one significant building creating light pollution, it is unlikely that courts would find this to be a nuisance since adjacent property owners fighting one building have had difficulty prevailing in lawsuits. *See* Dean N. Alterman, Comment, *Reflected Sunlight Is A Nuisance*, 18 ENVTL. L. 321, 323 (1988). "Buildings in urban areas often invade the property rights of adjacent landowners (i.e., by causing noise or emitting artificial light). However, courts usually consider these invasions to be reasonable because buildings are socially useful." *Id.*

150. *Id.* at 325 (citing the example of *Amphitheaters, Inc.*).

151. *But see infra* notes 189-90, 196 and accompanying text (indicating that in at least some areas statutes are being drafted to protect the night sky for astronomical observatories).

light alone will be sufficient for a plaintiff to prevail in the majority of cases.[152]

In contrast to the private nuisance claims, "a public nuisance is an unreasonable interference with a right common to the general public."[153] Public nuisance claims are often used to help abate environmental pollution that may harm public health or the general quality of life for the community.[154] To prevail, "a plaintiff must show that the defendant's conduct constitutes substantial and unreasonable interference with the public or with public property."[155] Usually plaintiffs in public nuisance claims are seeking injunctive relief to prevent defendant's continued nuisance activ-

---

152. The lack of cases where light was the sole nuisance factor does not help a plaintiff in determining whether his claim will prevail. *See, e.g., supra* note 124 and accompanying text. Most of the cases that have involved light as a nuisance have also had other factors such as noise, odors, or smoke. *See also* sources cited *supra* note 96 (indicating that light is usually one of many factors in a nuisance suit).

153. Smith, *supra* note 99, at 52. A private nuisance, therefore, is not necessarily created from a public nuisance and, in order for an individual to have standing to recover under a private nuisance claim, he or she must also have an injury distinct from that of the public. *See id. See also* RESTATEMENT (SECOND) OF TORTS § 821B (1977) (defining of public nuisance).

> (1) A public nuisance is an unreasonable interference with a right common to the general public.
> (2) Circumstances that may sustain a holding that an interference with a public right is unreasonable include the following:
> (a) Whether the conduct involves a significant interference with the public health, the public safety, the public peace, the public comfort or the public convenience, or
> (b) whether the conduct is proscribed by a statute, ordinance or administrative regulation, or
> (c) whether the conduct is of a continuing nature or has produced a permanent or long-lasting effect, and, as the actor knows or has reason to know, has a significant effect upon the public right.

Id.

154. *See* John L. Giesser, Comment, *The National Park Service and External Development: Addressing Park Boundary-Area Threats Through Public Nuisance*, 20 B.C. ENVTL. AFF. L. REV. 761, 761 & n.4, 775 (1993) (arguing that the public nuisance doctrine could possibly help the National Park Service in preventing development activities, which would cause increased traffic, smog, *noise*, and *artificial light*, near park boundaries).

155. *Id.* at 776.

> Thus, in order to find a public nuisance, a court must conclude that the defendant's conduct substantially damages public property, or significantly disturbs, offends, or endangers the health of members of the public with ordinary physical stature and "sensibilities." Furthermore, the court must find that the defendant's conduct decreases the environmental quality of the affected property to a level lower than that which the public reasonably expects.

Id. at 776-77.

ity.[156] Bright lights, in combination with other activity, have constituted a public nuisance in some jurisdictions.[157]

Whether a public nuisance claim would be more successful than a private nuisance claim in a light pollution case is uncertain. For example, given the relatively recent and few research findings that light pollution may be harmful to the public health, it may be unlikely that light pollution is yet to be accepted as harming public health.[158] In addition, usually only attorneys general or district attorneys have standing to advance public nuisance claims.[159] In either type of nuisance claim, however, the inconsistency and insufficiency of court holdings across jurisdictions does not seem adequate to deal with the modern problem of light pollution.[160]

---

156. *See id.* at 777. The plaintiff must show that the requested injunction would be the only adequate remedy, and granting of the injunction is of greater benefit than the impact of the injunction on the defendant. *See id.* Most courts that do issue injunctions usually require only a modification in the defendant's behavior, and not a complete cessation of activity. *See id.* at 783.

157. *See id.* at 791 & n.298 (citing Brough v. Ute Stampede Ass'n, 142 P.2d 670 (Utah 1943) and Hansen v. Indep. School Dist. No. 1, 98 P.2d 959 (Idaho 1939)). In *Hansen*, the court considered both how an injunction would impact the defendant's night baseball games, as well as the fact that the field was in a residential area. *See Hansen*, 98 P.2d at 962. The court required an injunction to prevent any lights from further shining onto the plaintiffs' premises. *See id.* The court noted that while night baseball games are not nuisances *per se*, they "become such under circumstances such as here where they are conducted in such a manner as to greatly interfere with legitimate and necessary use and enjoyment of the property of others." *Id.* The dissent argued that whether the baseball games constituted a nuisance in this particular case was a question of fact that should have been determined at trial. *See id.* at 963.

158. See *supra* notes 90-94 and accompanying text for a discussion of the developing data surrounding public health impacts of light pollution.

159. *See* Meiners & Yandle, *supra* note 99, at 927.

160. *See, e.g.*, PATTON, BOGGS & BLOW, *supra* note 99, at 28 (listing some drawbacks to common law nuisance; such as, the courts' unwillingness to allow "trifling inconveniences" to advance as claims, claims barred by statutes of limitations where the defendant's activity is not continuing, the inability of plaintiffs to prevail in court's balancing test, and the defendant's compliance with laws and permits); David Schoenbrod, *Protecting the Environment in the Spirit of the Common Law*, *in* THE COMMON LAW AND THE ENVIRONMENT 4-6 (Roger E. Meiners & Andrew P. Morriss, eds., 2000). While this article, and in fact the entire book, suggests that the common law plays an important role in environmental law, there are some notable drawbacks to the common law: liability in modern day pollution cases is more difficult to define, unsatisfactory remedies and unenforcement. "Science has progressed so that we think about pollution as a matter of degree as well as of kind." *Id.* at 5. "Even if the court can adjudicate liability, it will have trouble providing satisfactory remedies for many kinds of modern pollution problems" since typically only harms that are proven will be given remedies. *Id.*

B. Trespass in Environmental Pollution and Light as Trespass Cases

Trespass can generally be described as "an [intentional][161] intrusion or invasion of tangible property, either real or personal, which interferes with the possessor's interest in the right of *exclusive possession* of the property."[162] Nuisance and trespass, because they are so closely related, often result in confusion for plaintiffs as to which cause of action is best advanced in a particular case.[163]

In pollution as trespass cases, one of the main issues that courts have struggled with has been to what degree the invasion must be tangible or physical.[164] There are three cases that have served as a model for courts

---

161. The requisite intent can be shown by a plaintiff in one of two ways: proving that "the defendant acted for the purpose of causing the invasion," or "the defendant [knew] that there [was] a substantial certainty that its actions [would] result in the invasion." PATTON, BOGGS & BLOW, *supra* note 99, at 29-30. However, there may also be liability shown in cases where the defendant has acted negligently or recklessly. *See id.* at 29.

162. Powell, *supra* note 98, at 185.

> One is subject to liability to another for trespass, irrespective of whether he thereby causes harm to any legally protected interest of the other, if he intentionally
>
> (a) enters the land in the possession of the other, or causes a thing or a third person to do so, or
>
> (b) remains on the land, or
>
> (c) fails to remove from the land a thing which he is under a duty to remove.

Id. at 185 n.10 (quoting RESTATEMENT (SECOND) OF TORTS § 158 (1965)). The harm suffered by the plaintiff in a trespass claim must be substantial. *See* PATTON, BOGGS & BLOW, *supra* note 99, at 30. "The degree of substantiality will depend, as in nuisance, on the facts of the case and the outlook of the court that is examining those facts." *Id.* In addition, "courts have held that the definition of ["thing"] requires something larger and more substantial than smoke, dust, gas, or fumes." Meiners & Yandle, *supra* note 97, at 936. *See also* Joseph F. Falcone, III & Daniel Utain, Comment, *You Can Teach an Old Dog New Tricks: The Application of Common Law in Present-Day Environmental Disputes*, 11 VILL. ENVTL. L.J. 59, 71-72 (2000) (providing a historical overview of how harmful an intrusion had to be in order for a plaintiff to prevail in trespass cases).

163. *See* Smith, *supra* note 99, at 57-67 (providing a general discussion of some considerations for plaintiffs seeking to recover under either theory: statute of limitations, whether the defendant's actions are of a continuing or permanent nature, burden of proof, discovery rule, and damages that can be recovered).

164. Several tests have been developed by courts to determine whether an invasion should be classified as a trespass or instead as a nuisance. *See infra* notes 169-71 and accompanying text. *See also* Falcone & Utain, *supra* note 162, at 71-72. "That the traditional rule of trespass required some kind of physical invasion was undisputed; the real debate was grounded in how *tangible* or *visible* the invasion need be." *Id.* Advances in science, and subsequently a better understanding of potential environmental hazards rendered the "dimensional" test unworkable. *See id.* at 72. Today,

requiring this analysis and may be helpful in analyzing light as a pollutant: *Martin v. Reynolds Metals Co.*,[165] *Borland v. Sanders Lead Co.*,[166] and *Bradley v. American Smelting and Refining Co.*[167] Briefly, these cases represent the courts' development of an appropriate analysis for cases involving something less than the physical touching that is typically seen in trespass cases.[168] Initially, some courts used a "dimensional" test,[169] but ultimately more workable tests were created in order to deal with modern pollution problems.[170] Today, most courts use a "balancing test" similar to

---

most courts utilize a "balancing" test in pollution trespass cases. *See id.* at 73.

165. 342 P.2d 790, 791 (Or. 1959) (landowners brought action against defendant aluminum reduction plant, alleging that defendant caused airborne fluoride gases and particulates to settle upon landowners' property and thus constituted a trespass).

166. 369 So. 2d 523, 526 (Ala. 1979) (landowners brought action against defendant that recovered lead from used automobile batteries, alleging that lead particulates and sulfoxide deposits were falling on landowners' property and thus constituted a trespass).

167. 709 P.2d 782 (Wash. 1985) (landowners brought claims against defendant copper smelter, alleging that "microscopic, airborne particles of heavy metals" were falling on landowners' property and thus constituted either a trespass or nuisance).

168. *See* Falcone, III & Utain, *supra* note 162, at 70-71. "[T]he tort of trespass requires an 'intrusion;' some physical, tangible invasion of the plaintiff's land. This intrusion can occur 'on, beneath or above the surface' of the land." *Id.*

169. *See id.* at 72 & nn.58-59. The intrusion under the "dimensional" test was classified based on its visibility such that any invasions not detectable by the unaided eye were nuisances and visible invasions were trespasses. *See id.* This dimensional test was rejected by the *Martin* court. *See Martin*, 342 P.2d at 794.

> If, then, we must look to the character of the instrumentality which is used in making an intrusion upon another's land we prefer to emphasize the object's energy or force rather than its size. Viewed in this way we may define trespass as any intrusion which invades the possessor's protected interest in exclusive possession, whether that intrusion is by visible or invisible pieces of matter or by energy which can be measured only by the mathematical language of the physicist.

Id. In *Martin*, the court ultimately held that intrusion of fluoride particles qualified as a trespass, and distinguished the case from *Amphitheaters, Inc.* where it had earlier indicated that light could not be considered a trespass when it noted the difference between "a cannon ball and a ray of light." *See id.* The court explained that its statement in *Amphitheaters, Inc.* was *not* "a pronouncement that a trespass can *never* be caused by the intrusion of light rays or other intangible forces," and rather it meant that "the conduct of the defendant in a particular case may not be actionable if it does not violate a legally protected interest of the plaintiff." *Id.* For further analysis of the *Amphitheaters, Inc.* decision in *Martin*, see *id.* at 794-97. *See also* Powell, *supra* note 98, at 197-201 (providing a further analysis of the *Martin* decision).

170. *See* Falcone & Utain, *supra* note 162, at 71-74 & nn.58-62. "[W]ith the evolution of science, and consequently a deeper understanding of potential environmental hazards, the continued application of the 'dimensional' test by courts became infeasible." *Id.* at 72. "In *Borland v. Sanders Lead Co.*, the Alabama Supreme Court expressly acknowledged the problem of limiting the scope of liability under the *Mar-*

that of nuisance cases.[171]

The most recent case that may provide some insight as to how these types of pollution cases will be handled is found in *Bradley*.[172] In *Bradley*, the court accepted the elements of trespass that were set forth by the *Borland* court.[173] However, as this decision relates to light as trespass, the court appeared to indicate that light should appropriately be considered a nuisance.[174] It appears that this court and others will require an actual and substantial damage to be present in trespass claims – perhaps something that a ray of light cannot achieve.[175] These cases may guide plaintiffs who advance light as trespass claims in jurisdictions where there have been few cases where light is the sole factor claimed in the trespass.

Again, defined as "an intrusion or invasion of tangible property, either real or personal, which interferes with the possessor's interest in the right of *exclusive possession* of the property[,]"[176] claims in trespass have been

---

tin test." Powell, *supra* note 98, at 201. In *Borland*, the court found there was a trespass by the defendant's lead reduction activities which caused harm to the plaintiff's farm. *See* Borland v. Sanders Lead Co., 369 So. 2d 523, 529 (Ala. 1979). The court stated that:

> [w]hether an invasion of a property interest is a trespass or a nuisance does not depend upon whether the intruding agent is 'tangible' or 'intangible.' Instead, an analysis must be made to determine the interest interfered with. If the intrusion interferes with the right to exclusive possession of property, the law of trespass applies. If the intrusion is to the interest in use and enjoyment of property, the law of nuisance applies.
>
> . . . .
>
> Under the modern theory of trespass, the law presently allows an action to be maintained in trespass for invasions that, at one time, were considered indirect and, hence, only a nuisance. In order to recover in trespass for this type of invasion . . . a plaintiff must show 1) an invasion affecting an interest in the exclusive possession of his property; 2) an intentional doing of the act which results in the invasion; 3) reasonable foreseeability that the act done could result in an invasion of plaintiff's possessory interest; and 4) substantial damages to the *res*.

Id. at 529. It is perhaps this last requirement that may halt light as trespass cases, and more likely light pollution as trespass cases, in their tracks: how much does light really damage the res? *See* Powell, *supra* note 98, at 202 (positing "[a]t what point does the injury become so 'substantial' that a cause of action arises?").

171. *See* Falcone & Utain, *supra* note 162, at 73. The weight of the defendant's activity is measured against the character and substantiality of the harm. *See id.*

172. Bradley v. American Smelting and Refining Co., 709 P.2d 782 (1985).

173. *See id.* at 790.

174. *See id.* at 791. "When airborne particles are transitory or quickly dissipate, they do not interfere with a property owner's possessory rights and, therefore, are properly denominated as nuisances." *Id.* (citing several cases, including Amphitheaters, Inc. v. Portland Meadows).

175. *See id.*

176. Powell, *supra* note 98, at 185. "One is subject to liability to another for trespass, irrespective of whether he thereby causes harm to any legally protected interest of the other, if he intentionally (a) enters land in the possession of the other, or

used by plaintiffs disturbed by a neighbor's source of electric illumination.[177] In trespass cases, not only must a plaintiff show the defendant was acting intentionally, negligently or recklessly,[178] he "must also show that the harm suffered is substantial."[179] Although most courts initially required the intruding object to be visible to the unaided eye in order for it to be considered a trespass rather than a nuisance,[180] many jurisdictions may have begun to abandon this rule in modern air pollution cases.[181]

Perhaps the leading and only model case for light as trespass, albeit an unsuccessful one for the plaintiff, is *Amphitheaters, Inc. v. Portland Meadows*.[182] The court was not convinced that the cases relied upon by

---

causes a thing or a third person to do so, or (b) remains on the land, or (c) fails to remove from the land a thing which he is under a duty to remove." RESTATEMENT (SECOND) OF TORTS § 158 (1965). An actor's intent does not necessarily have to be hostile or possess a harmful motive, but "is an intent to bring about a result which will invade the interests of another in a way that the law will not sanction." *Bradley*, 709 P.2d at 786 (citing W. PROSSER, TORTS, § 8, at 31-32 (4th ed. 1971)). "Intent may be established if the defendant knows that there is a substantial certainty that its actions will result in the invasion." PATTON, BOGGS & BLOW, *supra* note 99, at 29-30.

177. *See Amphitheaters, Inc.*, 198 P.2d at 850.

178. *See* Powell, *supra* note 98, at n.22. "[I]nterests in land are protected under three theories: The defendant intentionally invaded a legally protected interest; the defendant negligently or recklessly brought about an invasion of the plaintiff's interest; or the defendant accidentally caused an invasion in the course of engaging in an activity for which strict liability is imposed." *Id.*

179. PATTON, BOGGS & BLOW, *supra* note 99, at 30. "The degree of substantiality will depend, as in nuisance, on the facts of the case and the outlook of the court that is examining those facts." *Id.* In many cases, a court will consider direct invasions with less scrutiny as they would indirect invasions in trespass claims. *See id.* "The distinction between injuries which were direct and substantial and those which were considered indirect and less substantial eventually evolved into a fictitious 'dimensional' test: '[I]f the intruding agent could be seen by the naked eye, the intrusion was considered a trespass.'" Powell, *supra* note 98, at 186. Otherwise, invisible objects were considered a nuisance as they were indirect and less substantial. *See id. See also Borland*, 369 So.2d at 527 (stating that "[t]he modern action for trespass to land stemmed inexorably from the common law action for trespass which lay when the injury was both direct and substantial").

180. *See* Powell, *supra* note 98, at 186.

181. *See id.* "The traditional rule limiting trespass to invasion of things that can be seen with the naked eye is an arbitrary one that has been discarded in some modern airborne pollution cases." *Id. But see* Michael C. Anibogu, *The Future of Electromagnetic Field Litigation*, 15 PACE ENVTL. L. REV. 527, 593 n.443 (1998) (noting the rule of a 1982 California Supreme Court decision that "actionable trespass may not be predicated upon non-damaging noise, odor or light"). *See also* Wilson v. Interlake Steel Co., 32 Cal.3d 229, 233 (1982) ("All intangible intrusions, such as noise, odor, or light alone, are dealt with as nuisance cases, not trespass.").

182. 198 P.2d 851 (Or. 1948). For facts of this case see *supra* notes 124-29 and

the plaintiff in his trespass claim were analogous to the case at bar,[183] and held that nuisance was the more appropriate claim for the plaintiff since the race track's operations were not substantial enough to constitute a trespass.[184] Given the relatively unsuccessful holding for the plaintiff in *Amphitheaters, Inc.*, as well as in the traditional pollution cases mentioned above, it is unlikely that light pollution, let alone light crossing one property owner's boundary to another, will ever constitute a trespass. As such, statutes and regulations providing for such types of claims may be required in order to address the growing problem of light pollution.

V. MODERN DEVELOPMENTS IN LIGHT POLLUTION LEGISLATION

Given the increased levels of light use around the country, and the relatively minimal and unpredictable success of common law claims for plaintiffs seeking to curb their neighbors' light use,[185] additional sources of law governing light use are needed if we are to curb light pollution as a larger problem.[186] To date, there is no federal, and relatively little state, legislation in place to reduce light pollution and its impacts,[187] and the majority of legislation for reducing this relatively new type of pollution largely ex-

---

accompanying text.

183. *See Amphitheaters, Inc.*, 198 P.2d at 850-51. The plaintiff relied upon one case which involved a trespass claim where the United States was "continuously firing artillery over the petitioners' land[,]" to which the Oregon Supreme Court replied, "[w]e need not argue the distinction between a cannon ball and a ray of light." *Id.* at 851. Furthermore, the plaintiff's reliance on *Shelburne, Inc. v. Crossan Corp.*, 95 N.J. Eq. 188 (1923), and another case involving light cast upon an individual's property resulted in the court pointing out that those cases were decided under the theory of nuisance, not trespass. *See id.* at 851.

184. *See id.* at 850-51.

185. *See supra* notes 95-184 and accompanying text.

186. *See, e.g., infra* notes 189-239, 242-51 and accompanying text for examples of state and local regulations that are geared towards further reducing light pollution that is different from the typical property owner versus property owner dispute over misdirected light.

187. *See* Brown, *supra* note 7, at 62. The United States Environmental Protection Agency's Green Lights Program is designed to promote voluntary energy conservation, but largely as it relates to *indoor* lighting and not *outdoor* lighting. *See id.* There is nothing currently in the federal Clean Air Act or other federal legislation to regulate light pollution. *See* 42 U.S.C. §§ 7401-7671 (Clean Air Act provisions). *But see* John-Mark Stensvaag, *State Regulation of Nuclear Generating Plants Under the Clean Air Act Amendments of 1977*, 55 S. CAL. L. REV. 511, 535 n.131 (1982) (suggesting that if "air pollution agent" under the Clean Air Act is interpreted broadly, then the Clean Air Act could conceivably cover nonionizing electromagnetic radiation which includes, among other things, *visible light*).

ists at the county and municipal levels.[188]

There are a few pioneer states that have enacted statewide or partially statewide light pollution or light reduction statutes.[189] The goal, scope and purpose of these statutes varies widely. In Arizona, the legislature found that "the continued existence of astronomical observatories in Arizona is in the best interests of the state," thus requiring a tighter control on nearby outdoor lighting.[190] Connecticut, through its statute, tries to "maximize

---

188. *See* Brown, *supra* note 7, at 49. "[L]egislation, regulations, or government policies to control the adverse effects of lighting are being implemented at the federal, state, and especially the local levels." *Id.*

189. Arizona, Connecticut, Maine, Michigan, New Mexico and Texas have state statutes that in some way regulate the use of light within all or a portion of the state. *See* ARIZ. REV. STAT. ANN. §§ 49-1101-06 (2001) (specifying what types of light fixtures can be used in outdoor lighting, such as fully or partially shielded fixtures, automatic shutoff devices and prohibition of mercury vapor fixtures); CONN. GEN. STAT. ANN. §§ 13a-110-110a (2001) (providing the types of roadway light fixtures that may be paid for with state funds); CONN. ACTS. 01-134 (Reg.) (repealing and amending § 13a-110a to include municipal roads and municipal funds in addition to state highways and state funds as within the statute's reach); ME. REV. STAT. ANN. tit. 23, § 708 (West 2001); ME. REV. STAT. ANN. tit. 5, § 1769 (West 1999) (requiring that state funds may only be used to install or replace a permanent outdoor luminaire if it is a full cutoff luminaire falling within the minimum illuminance recommended by the Illuminating Engineering Society of America or the federal Department of Transportation) ; MICH. COMP. LAWS §§ 324.75101-.75106 (2001) (stating factors that determine when and what kind of lighting can be installed within the dark sky preserve area); N.M. STAT. ANN. § 74-12-1 to -10 (Michie 2000) (mandating that after January 1, 2000 all outdoor lighting fixtures that are installed must be shielded or else extinguished automatically between 11:00 p.m. and sunrise); TEX. HEALTH & SAFETY CODE ANN. §§ 425.001-.002 (Vernon 2002) (mandating what kinds of lights can be "installed, replaced, maintained or operated using state funds" and that consideration be given to several factors, including light pollution, beforehand); TEX. LOC. GOV'T CODE ANN. § 240.032 (Vernon 2002) (granting permission to the commissioners court of a county within fifty-seven miles of a major astronomical observatory, acting upon the request of the director of the McDonald Observatory, to adopt orders to regulate the use of outdoor lighting so as to reduce interference with astronomical research). *See also* IND. CODE § 36-7-4-1403 (2000) (specifying that light pollution in Indiana may be included in specifications for development requirements); MONT. CODE ANN. § 23-1-126 (2001) (indicating that Montana's "good neighbor policy of public land use . . . seeks a goal of no impact upon adjoining private and public lands by preventing impact upon those adjoining lands from . . . light pollution"); VA. CODE ANN. § 15.2-920 (Michie 1997) (granting any Virginia locality the power by ordinance to "regulate outdoor lighting within an area one-half mile around planetariums, astronomical observatories and meteorological laboratories"); VA. CODE ANN. § 18.2-121.2 (Michie 1996) (creating a Class 3 misdemeanor in Virginia for those who willfully cast a spotlight upon private property used for agriculture or livestock).

190. ARIZ. SESS. LAWS CH. 236 §1 (West 1999). "[T]he legislature requests the cooperation of public and private utilities, billboard owners, counties, municipalities

energy conservation and to minimize light pollution, glare and light trespass" from roadway lighting.[191] Similarly, a Maine statute was drafted to minimize "glare and light trespass" from publicly funded lights along roadways.[192] A Michigan statute governs a limited region near Lake Hudson[193] where specific sections of Lenawee County in Michigan have been designated as a "dark sky preserve."[194] Some of the activities the legislature intended to protect in those areas includes those that require darkness, such as "enjoyment of the night sky, nighttime photography, and wildlife photography."[195] One Texas statute clearly is aimed at protecting the livelihood of the state's astronomers[196] and gives a wide range of options for county commissioners,[197] while the other statute appears to consider certain factors similar to Maine and Connecticut when allowing the use of

---

and others owning or operating outdoor lights to reduce light pollution which interferes with the successful operation of such observatories." *Id.*

191. CONN. GEN. STAT. § 13a-110a (West Supp. 2001). The title of this section, which falls under the Highway and Bridges title, is "Highway lighting designed to maximize energy conservation and minimize light pollution." *Id.*

192. *See* ME. REV. STAT. ANN. tit. 23, § 708 (West Supp. 2001).

193. *See* MICH. COMP. LAWS ANN. §§ 324.75102-.75105 (West 1999).

194. *See id.* §§ 324.75101-.75102. "'Dark sky preserve' means the area designated in section .75102." *Id.* § 324.75101. Section 324.75102 states that "[t]he state owned land at lake Hudson, legally described as: All state owned land located in [various sections]—Lenawee County, Michigan[,] is designated a dark sky preserve." *Id.* § 324.75012.

195. *Id.* § 324.75103.

196. *See* TEX. LOC. GOV'T CODE ANN. § 240.032(a), (b) (Vernon Supp. 2002). "[T]he commissioners court of a county, any part of which is located within 57 miles of a major astronomical observatory . . . may adopt orders regulating the installation and use of outdoor lighting . . . . The orders must be designed to protect against the use of outdoor lighting in a way that interferes with scientific astronomical research." *Id.* §240.032(a), (c). "[A] facility that is established to conduct scientific observations of astronomical phenomena and is equipped with one or more telescopes that (A) have objective diameters that total 69 inches or more; and (B) are permanently mounted in enclosed buildings" is considered a "major astronomical observatory." *Id.* § 240.031(1).

197. *See* TEX. LOC. GOV'T CODE ANN. § 240.002(c) (Vernon Supp. 2002). The commissioners may

> (1) require that a permit be obtained from the county before the installation and use of certain types of outdoor lighting in a regulated area;
> (2) establish a fee for the issuance of the permit;
> (3) prohibit the use of a type of outdoor lighting that is incompatible with the effective use of the observatory;
> (4) establish requirements for the shielding of outdoor lighting; and
> (5) regulate the times during which certain types of outdoor lighting may be used.

§ 240.032(c).

state funds for lighting.[198]

Perhaps the most comprehensive state statute belongs to New Mexico.[199] Titled the "Night Sky Protection Act,"[200] the purpose of this statute "is to regulate outdoor night lighting fixtures to preserve and enhance the state's dark sky while promoting safety, conserving energy and preserving the environment for astronomy."[201] By enacting this statute, New Mexico became the only other southwestern state to join Arizona in actively trying to preserve the night sky for its astronomers.[202] While Texas also has legislation designed to protect astronomical observations, the Texas statute is more limited than the New Mexico statute in that commissioners can only enact orders if the area is within fifty-seven miles of a major astronomical observatory.[203] Despite being enacted to protect the astronomers in each state, the positive result of these statutes – reduced nighttime light – potentially benefits wildlife and people as well.

Of the six state statutes regulating light pollution within at least a portion of the state's boundaries, only Texas actually defines light pollution.[204] Instead, most of these statutes define and regulate types of acceptable or unacceptable light fixtures installed or repaired with state funds, or other factors that contribute to light pollution.[205] The comprehensive and statewide New Mexico statute defines "outdoor lighting fixture" as "an outdoor artificial illuminating device, whether permanent or portable, used

---

198. *See* TEX. HEALTH & SAFETY CODE ANN. § 425.002. Consideration must be given to "energy conservation, reducing glare, minimizing light pollution, and preserving the natural night environment" when appropriating state funds for lighting. *Id.* at § 425.002(a)(4).

199. *See* N.M. STAT. ANN. §§ 74-12-1 - 10 (Michie 2000).

200. *See id.* § 74-12-1.

201. *Id.* § 74-12-2.

202. *See supra* note 190 and accompanying text.

203. *See* TEX. LOC. GOV'T CODE ANN. § 240.032(a). *See also* VA. CODE ANN. § 15.2-920, *supra* note 189 (allowing localities to regulate outdoor lighting only in areas within one-half mile around such observatories).

204. *See* TEX. HEALTH & SAFETY CODE ANN. § 425.001(2). "'Light pollution' means the night sky glow caused by the scattering of artificial light in the atmosphere." *Id.*

205. *See, e.g.*, ARIZ. REV. STAT. ANN. § 49-1101 (providing definitions only for types of lighting fixtures); CONN. GEN. STAT. ANN. § 13a-110a(6) (providing only a definition of "light trespass" to mean "light emitted by a luminaire that shines beyond the boundaries of the property on which the luminaire is located"); ME. REV. STAT. ANN. tit. 23, § 708(1) (defining only "Commissioner" and "roadway lighting"); ME. REV. STAT. ANN. tit. 5, § 1769 (defining "light trespass" identical to Connecticut's definition); MICH. COMP. LAWS ANN. § 324.75101 (defining only "dark sky preserve" and "fully shielded"); N.M. STAT. ANN. § 74-12-3 (defining only "outdoor lighting fixture" and "shielded").

for illumination or advertisement, including searchlights, spotlights and floodlights, whether for architectural lighting, parking lot lighting, landscape lighting, billboards or street lighting."[206] Beginning January 1, 2000, the New Mexico Night Sky Protection Act took effect and required virtually all newly installed outdoor lighting fixtures to be shielded,[207] and no mercury vapor outdoor lighting fixtures can be sold or installed.[208] Furthermore, "an outdoor lighting fixture not meeting these provisions shall be allowed, if the fixture is extinguished by an automatic shutoff device between the hours of 11:00 p.m. and sunrise."[209] The broad scope of the New Mexico statute becomes quite apparent when compared to other statutes – the regulations are not restricted to state funded activities or lighting fixtures.[210]

Arizona's statute also applies equally to most public and private lighting and is quite similar to the New Mexico statute. One difference between the Arizona and New Mexico statutes is that Arizona exempts streetlight fixtures from being shielded if a shielding device is not available from the light manufacturer.[211] Additionally, the Arizona statute, while allowing automatic shutoff devices for nonconforming light fixtures, does not mention the requirements for outdoor recreational facilities.[212] In this way, the Arizona statute appears to be slightly more lenient in regulating light use outdoors.

---

206. N.M. STAT. ANN. § 74-12-3(A) (Michie 2000). Arizona defines "outdoor lighting fixture" in almost identical terms. *See* ARIZ. REV. STAT. ANN. § 49-1101(2) (West 1999) (adding lighting for recreational areas to the list). *See also* TEX. HEALTH & SAFETY CODE ANN. § 425.001(3) (excluding from its list -- but otherwise similar to Arizona and New Mexico -- "lighting equipment that is required by law to be installed on motor vehicles or lighting required for the safe operation of aircraft").

207. N.M. STAT. ANN. § 74-12-4 (Michie 2000). The Act does not require that incandescent fixtures of 150 watts or less and all other sources of light up to 75 watts be shielded. *See id.*

208. *See id.* § 74-12-6.

209. *Id.* § 74-12-5(A).

210. *See id.* § 74-12-5(B). "No outdoor recreational facility, whether public or private, shall be illuminated after 11:00 p.m. except for a national or international tournament or to conclude any recreational or sporting event or other activity conducted, which is in progress prior to 11:00 p.m. at a ballpark, outdoor amphitheater, arena or similar facility." *Id.*

211. *See* ARIZ. REV. STAT. ANN. § 49-1102. There are also restrictions on the use and installation of mercury vapor lights. *See id.* § 49-1104.

212. *See id.* § 49-1103. Despite the statute, the issue of light pollution continues to be a pressing issue of concern for residents near areas where commercial developments are taking place. *See* Jennifer Sterba, *Some Love Development Idea, Other Residents Cite Concerns*, ARIZ. DAILY STAR, Jan. 31, 2001, at B3 (quoting one resident's concerns about light pollution that would come with a proposed 320,000 square foot shopping center near Tucson).

In contrast, Maine merely regulates state-funded permanent outdoor "luminaires," defined as "the complete lighting system, including the lamp and the fixture,"[213] on highways, or "roadway lighting."[214] Similarly, Connecticut also regulates which state-funded "luminaires" can be used on roadways.[215] Two feasible limitations to these two states' statutes are that they only pertain to state funded lighting, and the lighting is limited to roadway lighting.[216] Furthermore, Michigan's statute is even more restricted in that it regulates very localized lighting activity near Lake Hudson.[217] Although these statutes may be a step in the right direction, the proliferation of light use leading to light pollution has not been limited to state highway lighting or to small regions of a state.[218]

In addition to defining what types of fixtures are to be regulated, the statutes in the six states – Arizona, Connecticut, Maine, Michigan, New Mexico and Texas – all generally provide a minimum indication of how much light should be emitted and where, unless an exemption exists.[219] For example, in its dark sky preserve areas, Michigan requires that "lighting shall be directed downward,"[220] and when possible and/or appropriate, "fully shielded fixtures" or "motion sensor fixtures" should be used.[221] In

---

213. *See* ME. REV. STAT. ANN. tit. 5, § 1769(1)(J) (West 1999). Section 1769, which falls under Maine's Energy Conservation in Buildings Act, also provides the same definition for Maine Revised Statutes Annotated, title 23, section 708 which regulates highway lighting under the state highway law.

214. *See* ME. REV. STAT. ANN. tit. 23, § 708(1)(B). "'Roadway lighting' means lighting that is specifically intended to illuminate roadways for automobiles but does not mean lighting intended to illuminate roadways only for pedestrian purposes." *Id.* State funds may not be used to install or replace any permanent outdoor luminaire unless it is a "full cutoff luminaire when the rated output of the luminaire is greater than 1,800 lumens; . . . does not exceed the minimum illuminance recommended by the federal Department of Transportation for that purpose;" safety in the area cannot be achieved through other passive means, and consideration has been given to minimizing glare and light trespass. *Id.* § 708(2).

215. *See* CONN. GEN. STAT. ANN. § 13a-110a(a)(8), (b) (Supp. 2001).

216. *See id.* § 13a-110(a), (b). *See also* ME. REV. STAT. ANN. tit. 23, § 708(1)(B), *supra* note 214 and accompanying text.

217. *See* MICH. COMP. LAWS ANN. § 324.75102 (West 1999).

218. *See, e.g., supra* notes 27-33, 44-57 and accompanying text.

219. *See infra* notes 220-29 and accompanying text.

220. MICH. COMP. LAWS ANN. § 324.75104(a) (West 1999). The county commission is responsible for ensuring that only lighting for safety, security or the reasonable use and enjoyment of the property is used within the preserve, and that it does not reasonably interfere with activities requiring darkness. *See id.* § 324.75103.

221. *See id.* § 324.75104(b), (c). "Fully shielded" fixtures must be "shielded or constructed so that no light rays are emitted by the installed fixture at angles above 15 degrees below the horizontal plane and also constructed so that the filament or light source is not visible to the naked eye when viewed from a point higher than 15 de-

Connecticut, in addition to being required to "maximize energy conservation and minimize light pollution, glare and light trespass,"[222] no state funds will be used to install or replace permanent outdoor lights unless the light emitted from the luminaire is adequate for its intended purpose,[223] and is a "full cutoff" luminaire on secondary roads or primary roads where safety will not be compromised.[224] Connecticut also requires that state-funded light fixtures be allowed only in areas where "the purpose of the lighting installation or replacement cannot be achieved by reducing the speed limit in the area to be lighted or by installing reflectorized roadway markers, lines, warnings, informational signs or other means of passive or reflective lighting."[225] Similarly in Maine, no state funds will be used to replace or install any outdoor luminaire unless it is a "full cutoff luminaire," and the "maximum illuminance does not exceed the minimum illuminance recommended for that purpose."[226]

In Arizona, all outdoor lighting fixtures are required to be fully or partially shielded if over a certain wattage.[227] If outdoor lights do not meet this requirement or fall under other exemptions, they may use automatic shutoff devices to keep lights off from midnight through sunrise.[228] The New Mexico statute, while similar to the Arizona statute, provides further explanation for permissible use of lighting for private and public outdoor recreational facilities.[229]

In some states, certain exemptions from the statute may apply thus limiting the extent of how much light is being reduced or regulated.[230] For instance, in the Texas statute regulating light use within fifty-seven miles of observatories, there is an exemption for "outdoor lighting in existence or

---

grees below the horizontal plane." *Id.* at § 324.75101(b).

222. CONN. GEN. STAT. ANN. § 13a-110a(b)(1) (Supp. 2001).

223. *See id.* § 13a-110a(b)(2).

224. *See id.* § 13a-110a(b)(3)-(4). "Full cutoff luminaire" is defined as a "luminaire that allows no direct light emissions above a horizontal plane through the luminaire's lowest light-emitting part." *Id.* § 13a-110a(a)(2).

225. *Id.* § 13a-110a(b)(7)

226. *See* ME. REV. STAT. ANN. tit. 23, § 708(2) (West 1999); ME. REV. STAT. ANN. tit. 5, § 1769(2) (West 2002). Consideration should also be made to minimize glare and light trespass. *See id.*

227. *See* ARIZ. REV. STAT. ANN. § 49-1102 (West 1999). Streetlights are exempt if the manufacturer does not provide a shielding device. *See id.* In addition, airport navigational lighting is also exempt from this statute. *See id.* § 49-1105.

228. *See* ARIZ. REV. STAT. ANN. § 49-1103 (West 2001).

229. *See* N.M. STAT. ANN. § 74-12-5(B) (Michie 2000). "No outdoor recreational facility, whether public or private, shall be illuminated after 11:00 p.m. except for a national or international tournament or to conclude any recreational or sporting event which is in progress prior to 11:00 p.m. at a ballpark, outdoor amphitheater, arena or similar facility." *Id.*

230. *See infra* notes 231-36 and accompanying text.

under construction on September 1, 1975."[231] In its other light pollution statute, Texas exempts temporary emergency lighting, temporary lighting necessary for nighttime work, special events requiring additional lighting, lighting used solely to enhance the "aesthetic beauty of an object" or any compelling safety interest not addressable by other means.[232] Despite New Mexico's extensive legislation in this area, there are many exemptions from the statute, including "fixtures on advertisement signs on interstates and federal primary highways,"[233] "fixtures existing and legally installed prior to the effective date of the [statute],"[234] lighting required for "worker safety at farms, ranches, dairies, feedlots or industrial, mining or oil and gas facilities,"[235] and lights required for navigation at airports and aircraft safety.[236] Ironically, these types of exemptions are also some of the larger and more harmful sources of light pollution.[237]

Only two states provide for any kind of penalty for violations in their statutes—New Mexico and Texas.[238] In New Mexico, "[a]ny person, firm or corporation violating the provisions of the Night Sky Protection Act shall be punished as follows: A. for a first offense, the offender may be issued a warning; and B. for a second offense or offense that continues for thirty days from the date of the warning, [the offender will be charged $25.00] minus the replacement cost for each offending fixture."[239] In Texas, a violator may be sued by the county or district attorney for an injunction, and is found to have committed a misdemeanor.[240]

Over the past few years, several state bills have been introduced to state legislatures to enact statutes that will reduce light pollution or prompt stud-

---

231. TEX. LOC. GOV'T CODE ANN. § 240.034 (Vernon 2002).

232. *See* TEX. HEALTH & SAFETY CODE ANN. § 425.002(c)(2)-(5) (Vernon 2002).

233. N.M. STAT. ANN. § 74-12-7(1) (Michie 2000).

234. *Id.* at § 74-12-7(2). "[H]owever, when existing lighting fixtures become unrepairable, their replacements are subject to all the provisions of the Night Sky Protection Act." *Id.*

235. *Id.* at § 74-12-7(4).

236. *See id.* at § 74-12-7(3). *See also* ARIZ. REV. STAT. ANN. § 49-1105 (West 1999) (stating that "[n]othing in this article shall apply to navigational lighting systems at airports").

237. *See, e.g.*, *supra* note 28 and accompanying text.

238. *See* N.M. STAT. ANN. § 74-12-10 (Michie 2000); TEX. LOC. GOV'T CODE ANN. § 240.035 (Vernon 2002).

239. N.M. STAT. ANN. at § 74-12-10 (Michie 2000).

240. *See* TEX. LOC. GOV'T CODE ANN. § 240.035 (Vernon 2002). "(a) A county or district attorney may sue in the district court to enjoin a violation of this chapter. (b) A person who violates an order adopted under this chapter commits an offense. An offense under this section is a Class C misdemeanor. (c) Both civil and criminal enforcement may be used against the same conduct." *Id.*

ies of light pollution.[241] Perhaps the increasing introduction of these bills is due in large part to organizations such as the International Dark Sky Association and its local chapters that provide information for citizens seeking to introduce legislation at the state or local level.[242] While it is encouraging that some states are becoming increasingly aware that the night sky is a viable natural resource[243] and are providing for the study of

---

241. Alabama, Colorado, Connecticut, Iowa, Maryland, Massachusetts, Minnesota, New Hampshire, New York, Oregon, Pennsylvania, Rhode Island, Vermont and Wyoming have recently seen the introduction of bills and resolutions related to statewide study or reduction of light pollution. *See generally* H.R.J. Res. 316, Reg. Sess. (Ala. 2000) (failed to be carried over from 2000 regular session); H.B. 1160, 63d Gen. Assem., 1st Reg. Sess. (Colo. 2001) (pending); H.B. 5636, Gen. Assem., 2002 Reg. Sess. (Conn. 2002) (failed Joint favorable deadline); H.B. 265, 78th Gen. Assem., 1st Sess. (Iowa 1999) (failed to be carried over from 78th Second Session); H.B. 2141, 78th Gen. Assem., 2d Sess. (Iowa 1999) (failed to be carried over from 78th Second Session); S.S.B. 3035, 78th Gen. Assem., 2d Sess. (Iowa 1999) (senate study bill listing the findings of light pollution and its effects); H.S.B. 574, 78th Gen. Assem., 2d Sess. (Iowa 1999) (house study bill listing the findings of light pollution and its effects); H.J.R. 14, Gen. Assem., 415th Sess. (Md. 2001) (enacted: creates a task force to study the causes and consequences of light pollution); H.B. 3990, 181st Gen. Court., Reg. Sess. (Mass. 1999) (failed to move from House Committee on Science and Technology); H.B. 3528, 182d Gen. Court, Reg. Sess. (Mass. 2001) (pending: to Senate Committee on Ways and Means); H.R. 3368, 81st Leg., Reg. Sess. (Minn. 1999) (failed to be carried over from 81st Regular Session); H.B. 2276, 82d Leg., Reg. Sess. (Minn. 2001) (pending); S.B. 2204, 82d Leg., Reg. Sess. (Minn. 2001) (pending); H.R. 727, Gen. Court, 156th Sess. (N.H. 1999) (enacted: established a committee to study the effects of and possible solutions to light pollution); H.B. 222, Gen. Court, 157th Sess. (N.H. 2001) (pending); S.B. 6799, 223rd Leg. Sess. (N.Y. 1999) (attempted to amend the environmental conservation law); A.B. 6357, 223rd Leg. Sess. (N.Y. 1999) (attempted to amend the environmental conservation law); A.B. 5352, 224th Leg. Sess. (N.Y. 2001) (passed by Senate, but subsequently vetoed by Governor); H.B. 3973, 71st Leg. Assem. (Or. 2001) (pending); H.R. 175, 185th Gen. Assem., Reg. Sess. (Pa. 2001) (pending: attempting to establish a light pollution study committee); H.B. 300, 185th Gen. Assem., Reg. Sess. (Pa. 2001) (pending: includes penalty provision for individuals committing light trespass); S.B. 2399, 2001-2002 Leg. Sess. (R.I. 2002) (pending); H.R. 698, Adjourned Sess. of 1999-2000 Biennium (Vt. 1999) (failed to be carried over); S.B. 41, 55th Leg., 2000 Budget Sess. (Wyo. 2000) (failed to be carried over); S.B. 73, 56th Leg. (Wyo. 2001) (failed to be carried over); and H.B. 308, 56th Leg. (Wyo. 2001) (failed to be carried over). *See also* Lynn Bartels, *Lawmaker Has Bright Idea: Dim the Lights*, DENVER ROCKY MOUNTAIN NEWS, Jan. 16, 2001, at 10A (reporting that Colorado Representative Andrew Romanoff is proposing statewide legislation to reduce light pollution).

242. *See A Guide to Outdoor Lighting Regulations*, *at* http://www.skykeepers.org/ordguide.htm (last visited Apr. 19, 2002) (providing a long list of links to general reference on how to draft ordinances and other regulations for introduction to legislature, as well as a fairly comprehensive listing of currently enacted and proposed state and local regulations in the United States).

243. *See, e.g.*, H.S.B. 574, 78th Gen. Assem., 2d Sess. (Iowa 1999). In this

light pollution,[244] many states repeatedly fail to pass statewide light pollution laws.[245] Whether or not these bills pass now or in the future may depend on increased awareness and research findings regarding the effects of light pollution, and the determination that this kind of pollution should be regulated.

Municipal and county legislation is currently the most abundant source of light regulation.[246] Traditionally, these bylaws have mandated that emitted light does not extend visibly beyond the property line or create significant glare, and they represent the codification of light as trespass or nuisance.[247] More recently however, as awareness and appreciation of the harmful effects of light pollution has grown, municipalities have been enacting regulations specifically to reduce light pollution.[248] Regions within

---

House Study Bill, the General Assembly found that

> [t]he night sky is an important part of the natural heritage of the citizens of [Iowa] and steps should be taken to minimize the amount of terrestrial light that shines up into the night sky. Uplight from terrestrial light sources is wasteful and has made it increasingly difficult for the citizens of the state to enjoy the night sky because celestial objects are obscured.

Id. "Celestial objects" includes "stars, constellations, the milky way, meteors, comets, the northern lights, star clusters, nebulae, galaxies, the moon, and the planets." *Id.*

244. *See, e.g.*, H.J.R. 14, Gen. Assem., 415th Sess. (Md. 2001); H.R. 727, Gen. Court, 156th Sess. (N.H. 1999).

245. *See supra* note 241 and accompanying text.

246. "Most towns have zoning bylaws with some provisions related to lighting, intended to avoid nuisance conditions such as bright light from commercial properties spilling onto residential properties." Brown, *supra* note 7, at 59. For purposes of this Note, however, only a small and general sampling of the myriad of lighting regulations will be highlighted.

247. *See id.* at 60. *See, e.g.*, NEWPORT, R.I., CODIFIED ORDINANCES ch. 17.96, § 17.96.020(H), *available at* http://ordlink.com/codes/newportr/maintoc.htm (last visited Apr. 19, 2002) ("All external illumination shall be directed or shielded in such a manner that the illuminated areas are confined essentially to the property on which the illumination originates").

248. *See* Brown, *supra* note 7, at 61. For example, Bisbee, Arizona has entitled one of its ordinances the "Light Pollution Code." BISBEE COUNTY, ARIZ., ORDINANCE 0-98-10, art. 7.10 (1998), *available at* http://c3po.cochise.cc.a3.us/astro/pollution05.htm (last visited Apr. 19, 2002). The purpose of this ordinance is to:

> A. Minimize light pollution and light trespass for the enjoyment and use of property and the night environment by the citizens of the City of Bisbee, and
> B. Encourage the use of types, kinds, construction, installation and uses of outdoor light fixtures, lighting practices and systems which will reduce light pollution and light trespass, and
> C. Benefit astronomical research and observations, and
> D. Conserve energy while increasing nighttime visibility, utility, security and productivity.

Id. at art. 7.10.1. In the city of Ketchum, Idaho, there exists an ordinance known as the "Dark Sky Ordinance." *See* KETCHUM, ID., ORDINANCE 743 (June 21, 1999), *available at* http://darksky.org/~pai/vaIDA/ords/ketadrlo.html (last visited Apr. 19,

the United States where astronomical observatories are more common have been some of the first communities to enact light pollution prevention legislation.[249] However, areas that do not have local astronomical research stations have also taken into account the night sky when enacting ordinances.[250] Some communities specify design requirements for certain lighting activities,[251] the maximum output of light from a fixture,[252] or sug-

---

2002). Recognizing that Ketchum's dark nighttime sky is a natural resource to be enjoyed by both residents and tourists to this resort area, the purpose of the ordinance is worded more extensively than the Bisbee ordinance:

> 1.2 Purposes - The general purpose of this Ordinance is to protect and promote the public health, safety and welfare, the quality of life, and the ability to view the night sky, by establishing regulations and a process of review for exterior lighting. This Ordinance establishes standards for exterior lighting in order to accomplish the following:
>
> a. To protect against direct glare and excessive lighting;
> b. To provide safe roadways for motorists, cyclists and pedestrians;
> c. To protect and reclaim the ability to view the night sky, and thereby help preserve the quality of life and the tourist experience;
> d. To prevent light trespass in all areas of the City;
> e. To promote efficient and cost effective lighting;
> f. To ensure that sufficient lighting can be provided where needed to promote safety and security;
> g. To allow for flexibility in the style of lighting fixtures;
> h. To provide lighting guidelines;
> i. To provide assistance to property owners and occupants in bringing nonconforming lighting into conformance with this Ordinance; and,
> j. To work with other jurisdictions within Blaine County to meet the purposes of this Ordinance.

Id. at § 1.2.

249. *See* Brown, *supra* note 7, at 60. "A small number of municipalities developed more focused outdoor lighting regulations as early as the 1970s, primarily to reduce adverse effects on astronomical research." *Id.* In Arizona, one of the states that also has a statewide light pollution prevention statute, there are two major observatories in the Tucson area. *See id.* Because of its local regulations, "Tuscon is reputed to be the only city in the United States with a population exceeding 500,000 where the Milky Way is visible from the city center . . . ." *Id.* Similar legislation was later passed in Boxborough and Harvard Massachusetts, two towns close to the Oak Ridge Observatory, the largest telescope east of the Mississippi River. *See id.*

250. *See id.* at 60-61. For some localities, the incentive to reduce light use may be purely economic and not environmental. *See* Keri Buscaglia, *Light Pollution is a Glowing Concern*, CHIC. TRIB., Jan. 31, 2001, at 4 (noting that some dark sky advocates often point out that with a few simple modifications to current lighting, a savings of $200,000 a year could be attained in a city of 600,000 residents); *Let There Be Less Light*, CAPITAL TIMES, Oct. 15, 1999, at 10A (turning off 1,200 street lights in Madison, Wisconsin would have saved $50,000 annually). While many of these towns also see the economic incentive in reducing the number of lights or increasing the energy efficiency in the bulbs that are used, the aesthetic benefits of reducing light pollution may also be a driving factor. *See* Brown, *supra* note 7, at 61-62.

251. *See* Brown, *supra* note 7, at 62. Atlanta, Georgia requires billboards to be

gest or require certain periods during which exterior lights should be turned off.[253] Other communities have enacted ordinances with the local habitat and wildlife in mind.[254] Penalties for violations can also be found

---

lit from the top of the sign facing down rather than lights on the bottom of the sign facing up. *See id.* In Limerick, Pennsylvania, the town ordinance requires, among other things, that "[d]irectional fixtures such as flood lights, spot lights and sign lights shall be installed or aimed so that they do not shine directly into the window of a neighboring residence, directly into a roadway, or skyward." LIMERICK TOWNSHIP, PA, ORDINANCE ch. 180 § (C)(3)(b) (Nov. 6, 1996). Limerick also has a provision for illuminated signs as well: "Externally illuminated signs shall be lighted by fixtures mounted at the top of the sign and aimed down rather than by fixtures mounted at the bottom of the sign and aimed up." *Id.* at § 3(f). In Moab, Utah the requirement is more of a performance standard rather than a specific design standard. *See* MOAB, UTAH, CITY CODE § 15.44.070 (2001). "Lighted signs shall have stationary and constant lighting . . . . A lighted sign on a property that abuts a residential zone shall be subdued and shall not be allowed to penetrate beyond the property in such a manner as to annoy or interfere with the adjacent residential properties." *Id.* at § 15.44.070 (A)-(B). If a person complains about any light that reaches a residential property, the city council can either dismiss the complaint as unreasonable or require the offending party to use shielded lights. *See id.* at 15.44.070 (B). In Bisbee, Arizona, outdoor lights of greater than 150 watts are required to be either fully or partially shielded, with an exemption created for streetlights if such lights are not available from the manufacturer, and a general exemption for outdoor sporting events (which are required to use shielded fixtures) that have started prior to 11:00 p.m. *See* BISBEE COUNTY, ARIZ., ORDINANCE 0-98-10, art. 7.10, §§ 7.10.3, 7.10.7 (May 5, 1998).

252. *See, e.g.*, KETCHUM, ID., ORDINANCE 743 §§ 3.2, 3.6 (June 21, 1999) (providing certain exemptions from using shielded fixtures where the wattage or lumen output is below a specified amount); LIMERICK TOWNSHIP, PA., ORDINANCE 180 § 3(C)(1)(a) (Nov. 6, 1996) (providing a chart for the intensities of light allowed in certain areas such as parking lots and roadways). Pima County, Arizona is one of the first communities to establish a lumens per acre restriction on lighting in a metropolitan area, it is only the third law in the country to limit lumens per acre. *See* Rotstein, *supra* note 61, at A2 (noting also that Flagstaff and Cottonwood, Arizona are the other two communities with such stringent requirements for lighting output).

253. *See, e.g.*, KETCHUM, ID., ORDINANCE 743 § 3.5 (June 21, 1999) (encouraging community members to turn off "[a]ll non-essential exterior commercial and residential lighting . . . after business hours and/or when not in use. Lights on a timer are encouraged. Sensor activated lights are encouraged to replace existing lighting that is desired for security purposes"); BOULDER, COLO., BOULDER REV. CODE, tit. 10, ch. 11, § 10-11-3(c)(5)(D)(1981) (requiring indirectly or internally lit signs in a residential area to be turned off between 11:00 p.m. and 7:00 a.m., unless required for safety); BOULDER, COLO., BOULDER REV. CODE, tit. 10, ch. 11, § 10-11-3(c)(5)(E) (1981) (requiring any illuminated sign visible from and located within 300 feet of residentially zoned property to be turned off between 11:00 p.m., or one-half hour after the use of which is completed, and 7:00 a.m.).

254. *See, e.g.*, MAUI, HAW., COUNTY CODE, ch. 19.7, § 19.70.100(O) (2001) (providing in its zoning statute for the Lanai Project District of Maui, Hawaii that

in some ordinances.[255]

Notwithstanding the benefits of the state and local regulations that have been passed, there are also many impediments that have been encountered.[256] One of the problems with many regulations is that they are often too vague, difficult to enforce or remain completely unenforced by law enforcement officials.[257] In other cases, they do not appear to be stringently[258] or sufficiently[259] worded to prevent light pollution. Fears of

---

"lighting shall be established in a manner so as to not adversely impact the surrounding areas including the shoreline and ocean"); FLA. ADMIN. CODE ANN. r. 62B-55.001 (2000) (requiring the Florida Department of Environmental Protection to "designate coastal areas utilized, or likely to be utilized, by sea turtles for nesting, and to establish guidelines for local government regulations that control beachfront lighting"). For a recent case involving the Florida ordinance, see Loggerhead Turtle v. County Council of Volusia County, 148 F.3d 1231 (11th Cir. 1998), *cert. denied*, 526 U.S. 1081 (1999). *See also* Katherine R. Butler, Comment, *Coastal Protection of Sea Turtles in Florida*, 13 J. LAND USE & ENVTL. L. 399, 423-26 & nn.236-54 (providing a detailed overview of the state of Florida's and its several coastal counties' attempts to protect sea turtles through lighting ordinances); Cece Von Kolnitz, *Oak Island/Providing an Improved Habitat*, MORNING STAR, Sept. 28, 2000 (describing a Wilmington, North Carolina project to create a sea turtle nesting area and the recent approval of a "lighting ordinance meant to reduce light pollution for the turtles").

255. *See, e.g.*, BISBEE COUNTY, ARIZ., ORDINANCE 0-98-10, art. 7.10, § 7.10.9 (1998) (imposing a minimum fine of $100 to a maximum of $1,000 for each day there is a violation of the ordinance).

256. *See infra* notes 257-62 and accompanying text.

257. *See* Brown, *supra* note 7, at 59. Many light fixtures, even those that are shielded, often emit some amount of light that may cross a property line, and often there is not an adequate definition of what constitutes glare so it is difficult to prohibit such activity. *See id.* "Because of the imprecision and unintended stringency of such bylaws, they are rarely enforced." *Id.* Furthermore, residents have expressed fears that individual freedoms may be compromised by the "light police." *See id.* at 61. In Deschutes County, Oregon, one of the region's first dark sky preservation ordinances was enacted six years ago, imposing maximum allowable lighting intensities. *See* Doug Irving, *Mayor Wants to Tone Down Glow*, PORTLAND OREGONIAN, Nov. 21, 2000, at D02. Despite its presence on the books, the county had only recently begun enforcing this ordinance in September of 2000, and has yet to issue the $250 fine for violations. *See id. But see* Peter H. Lehner, *Act Locally: Municipal Enforcement of Environmental Law*, 12 STAN. ENVTL. L.J. 50, 55-66 (1993) (arguing that municipalities are perhaps in the best position to enforce environmental law against polluters because they "have great incentive to act, knowledge about which of the numerous local violators present the greatest threat, flexibility to pursue those violators under a variety of laws, [standing and credibility,] and the ability to respond quickly").

258. *See, e.g.*, ANN ARBOR, MICH., CITY CODE, ch. 61, § 5:510 (1998). This zoning ordinance provides that signs may be illuminated by artificial lights, provided they are white, but does not require any specific output or placement of the lights. *See id.* In residential areas of Quincy, Massachusetts, "no outdoor decorative or flood-lighting shall be permitted except lighting primarily designed to illuminate walks,

crime and reduced safety also seem to inhibit cities and towns from enacting more rigorous regulations.[260] Certain enacted or proposed regulations, denounced as illegal, unfair or inapplicable in some cases, have been the center of controversy in lawsuits or town meetings.[261] Finally, convincing

---

driveways, parking areas, doorways, outdoor living areas or outdoor recreational facilities, . . . except temporary holiday lighting, and except decorative floodlighting of institutions, public or historic buildings." QUINCY, MASS., CITY CODE tit. 17, § 17.36.010 (1999).

259. *See, e.g.*, LAS VEGAS, NEV., CITY CODE § 18.12.260 (2001). In a city generally known for its thousands of outdoor electric lights, Las Vegas' street lighting ordinance only vaguely requires that "[s]treet lighting for public streets shall be designed, installed or upgraded in accordance with City standards." *Id.*

260. *See, e.g.*, Terri Williams, *Astronomer Fights for 'Good Sky,'* DALLAS MORNING NEWS, Dec. 22, 2000, at 1M (quoting Mesquite, Texas Mayor Mike Anderson as saying that, when asked to consider changing the types of bulbs in the retail and business areas of the city, it was not feasible because "[i]t's important we provide enough lights for our citizens for their public safety"); *Let There Be Less Light*, CAPITAL TIMES, Oct. 15, 1999, at 10A (safety concerns of residents were seen to outweigh the economic savings that would have resulted from proposed legislation to turn off over 1,000 lights in residential neighborhoods and along busy roads). *But see supra* note 26 and accompanying text (showing, for example, that a U.S. Department of Justice study indicates that crime does not increase at night when there is less light than during the day).

261. For recent litigation involving the Tucson, Arizona Outdoor Lighting Code (OLC), which designates how billboards must be illuminated, see Whiteco Outdoor Advertising v. City of Tucson, 972 P.2d 647, 648 (Ariz. Ct. App. 1998) (involving the ability of a charter city in using its police powers to ban light fixtures mounted on the bottom of billboards and whether the nonconforming use statute in an outdoor lighting code precludes it from doing so) and Eller Media Co. v. City of Tucson, 7 P.3d 136, 139 (Ariz. Ct. App. 2000) (subsequent case related to *Whiteco Outdoor Advertising* [business name changed to Eller] that heard constitutional issues of the OLC). The controversy in *Whiteco Outdoor Advertising* involved the changing of the original OLC enacted in 1985 (which required billboards to be illuminated from the top, but exempted those that were already in existence and designed with lights elsewhere) to a more recent version of the OLC enacted in 1994 which required *all* lighting to be mounted on top of billboards without exception. *See Whiteco Outdoor Advertising*, 972 P.2d at 648. Whiteco was notified by the City of Tucson that its lights were in violation of the 1994 OLC, and subsequently Whiteco argued that the City's ability to regulate such lighting was granted under the state's zoning statutes which had a nonconforming use provision (a provision that stated no new ordinance could affect property that was already in existence). *See id.* at 649. The court ultimately decided that the City did in fact have police powers to regulate billboard lighting, and is not restricted by the state zoning laws in regulating illumination of signs. *See id.* at 650. Subsequently in *Eller Media Co.*, the court considered Eller Media Company's claims that the injunction granted by the lower court to stop illuminating signs from the bottom violates substantive due process and equal protection rights. *See Eller Media Co.*, 7 P.3d at 139. Because the OLC requirement does not affect free speech or any other

the general public that light pollution is even an environmental problem, and more importantly, that it can be solved, is yet another hurdle in advancing legislation.[262]

---

fundamental right (there was no claim by Eller that its billboard messages were inhibited by top rather than bottom mounted lights), and Eller Media Company is not a member of suspect class, the court held the appropriate test is the rational basis test. *See id.* at 139-40. Under this test, the court held that the OLC's stated purpose of reducing light shining upwards to better astronomical observations is rationally related to the requirement of top-mounted lights. *See id.* at 140. Furthermore, the court also disputed that Eller Media Company was denied equal protection under the law simply because other light sources, such as parking lots, are permitted to use other lights. *See id.* at 140-41. "This suggestion would preclude the City from regulating all but the brightest bulb in the night sky . . . it presumes that all light sources should and must emit the same amount of light." *Id.* at 141. *See also* Howard Fischer, *Cities the Boss on Lighting Can Regulate Businesses*, ARIZ. BUS. GAZETTE, Jan. 18, 2001, at 1 (reporting on the most recent activity in the *Whiteco Outdoor Advertising* and *Eller Media* cases). "Arizona cities have broad rights to regulate outdoor lighting by businesses, even for purely aesthetic reasons . . . the Arizona Supreme Court upheld a 16-year old Tucson ordinance that requires billboards to be lighted from the top." *Id.* *See also* Daniel M. Monte, *Summerfield Decides to Turn Down Lights, Keep Small-Town Feel*, GREENSBORO NEWS & RECORD, Sept. 6, 2000, at B7 (some residents believed that the proposed ordinance, which was ultimately passed, would be costly for some to be in compliance, difficult to enforce or create an unfair burden for certain activities that require nighttime lighting such as ball fields); Irving, *supra* note 5, at D2 (reporting that Sandy, Oregon Mayor Linda Malone expects to hear opposition to her proposed lighting law from car dealerships, another commercial entity that tends to use bright outdoor lights); Isabel Sanchez, *Sportsplex Loses Light Decision*, ALBUQUERQUE J., Sept. 23, 2000, at E1 (reporting that district judge held that a Bernalillo County law requiring light poles to be no more than 16 feet and must not produce light trespass 10 feet beyond a property line is constitutional despite plaintiff sports complex's claim that the law does not have a "rational basis"); Rotstein, *supra* note 5, at 12 (noting that an attorney for a commercial entity argued that the industry standard for outdoor lighting in shopping centers is 900,000 lumens which is three times the amount of the recently enacted lighting code requirements).

262. *See* Zielinski, *supra* note 66, at B1. Despite the growing prominence of light pollution as an environmental concern and something of great importance to astronomers and non-astronomers alike, light pollution appears to remain low on the list of priorities for anti-sprawl and environmental activists. *See id.* Gaithersburg, Virginia Councilmember-at-Large, Ann Somerset, who was elected to a platform that had members who were opposed to light pollution legislation, said, "I think [light pollution] has been below the radar, because I don't think the average person is aware anything can be done. I think you just assume that if you want to see the beauty of the night sky, you have to go to the mountains." *Id.* However, there may be a growing appreciation for light pollution as akin to other more widely accepted forms of pollution such as air and water pollution. *See* Chris Reinolds, *Bright Lights in Need of Dimmer Switch*, ATLANTA JOURNAL & CONSTITUTION, Nov. 9, 2000, at 9 (quoting a local astronomy club member)

VI. THE NOISE POLLUTION MODEL: A RECOMMENDATION FOR FEDERAL ASSISTANCE IN REDUCING LIGHT POLLUTION

Given that the common law,[263] state legislation, and local regulation[264] of light do not appear to be adequately dealing with the growing modern problem of light pollution, there may be a need for some additional federal legislative guidance as to how we deal with this relatively new environmental concern. To date, there is no federal regulation of light pollution as it is not explicitly mentioned in the federal Clean Air Act.[265] Until there is at least some uniformity of laws among the states, there may be difficult hurdles for organizations and individuals attempting to curb light pollution without some federal legislation.[266]

---

263. *See supra* notes 95-184 and accompanying text.

264. *See supra* notes 185-260 and accompanying text.

265. *But see supra* note 187 and accompanying text.

266. For example, standing, as required by Article III of the United States Constitution, may be an issue for individuals, advocacy organizations or astronomical research labs wishing to stop the problem of too much light being emitted from too many sources. *See, e.g.*, Lehner, *supra* note 257, at 62 & n.42 (referring to what the Supreme Court held in 1992 in *Lujan v. Defenders of Wildlife*, this article noted that "[l]ack of standing is a growing problem for non-profit environmental groups and local environmentalists"). *See also* Lujan v. Defenders of Wildlife, 504 U.S. 555 (1992). In *Lujan*, the plaintiffs were members of an environmental advocacy group that were attempting to stop activities abroad that might have been in violation of the Endangered Species Act. *See id.* at 562-63. Under the *Lujan* dicta stating that "of course, the desire to use or observe an animal species, even for purely esthetic purposes, is undeniably a cognizable interest for purpose of standing," it is uncertain whether individuals who are members of astronomy groups, or nonprofit organizations that study and promote particular nocturnal species, would be considered as possessing Article III standing to argue that light pollution activities must be stopped *without some sort of federal legislation*. *See id. See also* Friends of the Earth, Inc. v. Laidlaw Envtl. Serv., 528 U.S. 167, 174-75, 183-85 (2000) (a suit similar to *Lujan* which held that an environmental group had standing to sue under the citizen suit provision of the Clean Water Act and also reaffirmed that recreational and aesthetic reasons are sufficient for standing). *See generally* Peter Van Tuyn, *Thirtieth Anniversary Edition Essays: "Who Do You Think You Are?": Tales from the Trenches of the Environmental Standing Battle*, 30 ENVTL. L. 41 (2000) (offering a pre-*Friends of the Earth* look at standing in environmental suits as well as a short interpretation of the *Friends of the Earth* decision); David R. Hodas, *Standing and Climate Change: Can Anyone Complain About the Weather?*, 15 J. LAND USE & ENVTL. L. 451, 454-55 (2000) (offering an interesting discussion of standing and the required degree of injury from pollutants responsible for global warming).

> Must the plaintiff be directly harmed by the pollutant itself, as is the case in the classic nuisance and pollution cases? Or, may the plaintiff complain about the impact of climate change that will be widespread and suffered by all persons where the threatened impact is only a statistical artifact rather than a particular event or effect that is harmful to the plain-

An analogue to light pollution can be found in noise pollution.[267]

> There are several parallels to be drawn between light pollution and noise [pollution], which occupied a similarly uncertain territory prior to 1960. ... Light has the potential to cause distress and is an equally insidious pollutant. Noise and light are both intangible and ephemeral. ... It is the perception of the relative degree, frequency and effect of the problem which causes noise pollution to be more regulated than light pollution rather than any technological differences.[268]

Under this assumption, we may find potential solutions and ideas on how best to deal with light pollution, and perhaps avoid some of the mistakes that have already been made, by looking to the federal regulation of noise.

Described as unwanted sound,[269] noise is considered an environmental pollutant[270] with several harmful effects—similar, in many aspects, to light pollution.[271] The generally recognized sources of unwanted sound include

---

> tiff? Thus, the climate change standing problem goes to the central question of what is injury, how particularized it must be, and is standing to be essentially a constitutionalization of the special injury rule in public nuisance?

Id. On the other hand perhaps there is some indication that federal legislation is not necessary to give the required standing for light pollution cases. *See, e.g.*, Texas Dep't Transp. v. City of Sunset Valley, 8 S.W.3d 727, 728 (Tx. 1999) (reaffirming a lower court decision that the municipality of Sunset Valley had standing to sue for the physical taking of its streets, had standing derived from injuries that may result from the increased travel on the expanded highway as well as resultant noise and light pollution, and that the Texas Department of Transportation was not immune from such a lawsuit). This Note does not consider the issues of standing that may arise under federal or other legislation, but recognizes that it may, in fact, be another hurdle in preventing light pollution.

267. *See* Jewkes, *supra* note 5, at 10.

268. *Id.* (making this analogy as part of an analysis of light pollution in England).

269. *See* FRANK P. GRAD, TREATISE ON ENVIRONMENTAL LAW, VOLUME 4, 5-2 (2000). *See also* Donald F. Anthrop, *The Noise Crisis*, 20 U. TORONTO L.J. 1 (1970), *reprinted in* NOISE POLLUTION AND THE LAW 5 (James L. Hildebrand ed., 1970) (explaining how sound is measured and produced).

270. *See* GRAD, *supra* note 269, at 5-2. "[Noise] is nonetheless unique among 'pollutants' in that it leaves no residual accumulations. It dissipates very rapidly, and even very loud noises are rather limited in the geographic area they affect. Nonetheless, noise has become a substantial problem, especially in urban areas where three-fourths of the nation's people live. In spite of its rather limited geographic reach, noise has become a national problem in terms of the control strategies that must be applied." *Id.*

271. *See* CLIFFORD R. BRAGDON, *Preface* to NOISE POLLUTION: THE UNQUIET CRISIS xvii (1970). "A threat to physical and psychological well-being, the sounds of our technology follow us through our working, leisure, and sleeping hours." *Id.* Some

airplanes and airports,[272] automobiles,[273] sonic booms,[274] and a myriad of general everyday products.[275] Noise pollution began to receive significant legal and media attention during the 1970s.[276] In terms of environmental priority, like light pollution, noise pollution is not foremost in most peo-

---

of the negative effects of noise include hearing loss, annoyance, psychological problems, physiological stress, and interference with sleep. *See id.* at 63-80. *See also generally* KUPCHELLA & HYLAND, *supra* note 43, at 499-505 (highlighting the various harmful effects of noise, including impacts on wildlife); Jason A. Lief, Note, *Insuring Domestic Tranquility Through Quieter Products: A Proposed Product-Nuisance Tort*, 16 CARDOZO L. REV. 595, 598-600 (1994) (noting that hearing loss to older individuals may be related to noise encountered during younger ages and the psychological stress resulting from noise may be linked to some cancers). Many of the same or similar health impacts stemming from light pollution are discussed *supra* at notes 90-94 and accompanying text.

272. *See generally* Kristen L. Falzone, Comment, *Airport Noise Pollution: Is There a Solution in Sight?*, 26 B.C. ENVTL. AFF. L. REV. 769, 769 (1999). "Since the introduction of commercial jets in 1958, the noise problem generated from airport operation has become increasingly widespread, affecting millions of Americans." *Id.* *See also* Herbert Tenzer, *Jet Aircraft Noise: Problems and Their Solutions*, 13 N.Y.L.F. 465 (1967), *reprinted in* NOISE POLLUTION AND THE LAW 114 (James L. Hildebrand ed., 1970) (noting that, despite the advantages of the commercial jet industry, noise pollution is a primary disadvantage, particularly for those living in proximity to airports).

273. *See generally* Steven N. Brautigam, Note, *Rethinking the Regulation of Car Horn and Car Alarm Noise: An Incentive-Based Proposal to Help Restore Civility to Cities*, 19 COLUM. J. ENVTL. L. 391, 393-95 (1994). Since its debut in the late nineteenth century, the automobile has contributed to noise pollution through engines, horns and most recently, electronic burglar alarms. *See id.* Analogous disturbances can be said to have occurred since the development of the light bulb which was also developed around the same time. *See supra* notes 17-43 and accompanying text.

274. *See* Anthony J. Ortner, *Sonic Boom: Containment or Confrontation*, 34 J. AIR L. & COM. 208 (1968), *reprinted in* NOISE POLLUTION AND THE LAW 240-42 (James L. Hildebrand ed., 1970) (describing the phenomenon of sonic booms, which produce a loud sound when aircraft reach the speed of sound).

275. *See generally* James L. Hildebrand, *Preface* to NOISE POLLUTION AND THE LAW 8-20 (James L. Hildebrand ed. 1970); KUPCHELLA & HYLAND, *supra* note 43, at 506 (noting that several products used in the home can produce a significant amount of noise, as well as city noise generated by traffic). Similar to noise pollution, light pollution also has several notorious sources of unwanted light. *See supra* notes 24, 28-33 and accompanying text.

276. *See* Hildebrand, *supra* note 272, *Preface* at v. "Noise pollution and its legal implications is becoming a relevant, current, and important topic for discussion in our technologically expanding society. This is the first book concerning *noise pollution and the law* to be published in the United States." *Id.* (referring to the book which was published in 1970). Light pollution was also becoming a noticeable problem around the 1970s, and today there is significant media coverage of this phenomenon. *See supra* notes 5-6 and accompanying text.

ple's minds.[277] While the various sources of modern noise control laws include local, state and federal legislation as well as the common law,[278] this was not always the case.

Common law was initially the only source of relief for individuals who wanted to stop nearby noise.[279] Nuisance claims comprised the earliest methods of dealing with noise.[280] Products liability claims have also been used in many instances.[281] In relation to airport noise pollution in particular, inverse condemnation has been another means by which plaintiffs have sought judicial action.[282] Eventually, however, these types of claims became inadequate in addressing the growing public need of dealing with

---

277. *See* Brautigam, *supra* note 273, at 401-02. "With the exception of airport noise, environmental groups have generally ignored noise pollution. This relative neglect of the noise issue may be due to the fact that other, more visible sources of pollution such as air and water pollution have been viewed as more threatening to human and ecological health." *Id.* at 401. In comparison, light too is not very high on environmentalists' priority list either. After the early 1980s when President Ronald Reagan cut funding for the Office of Noise Abatement and Control, noise pollution began to receive less federal attention. *See* Brad Cooper, *Cities Respond to Clamor for Peace and Quiet with Ordinances*, KANSAS CITY STAR, Apr. 15, 2001, at B1 (noting that noise "is getting heightened attention across the country," particularly since the mid-1990s).

278. *See* Albert J. Rosenthal, *Noise Control and the Law*, *in* HANDBOOK OF NOISE CONTROL, 37-3 to 37-5 (Cyril M. Harris, ed., 2d ed., 1979) (outlining the development of various types of noise pollution laws through the mid-1970s); *see generally* GRAD, *supra* note 269, at 5-35 to 5-101 (describing in detail the federal regulation of noise since the enactment of the first federal noise pollution statutes in the 1970s). *See also supra* notes 95-255 and accompanying text as these kinds of pollution control have been utilized to control light pollution.

279. *See* Rosenthal, *supra* note 278, at 37-3. "The earliest forms of noise control law were the creation of the judiciary, well before sufficient public interest was generated in the subject to impel legislatures to enact statutes." *Id.*

280. *See* Lief, *supra* note 271, at 609-10. However, "[t]he need to protect industrial development has justified new limits on nuisance law just as the industrial age is spawning the most powerful instruments of nuisance." *Id.* at 611. "Even when nuisance law provides a remedy against a particular user of a noisy product, it creates no incentive for manufacturers to make a quieter product." *Id.* at 612. *See also* Brautigam, *supra* note 273, at 417-20 (providing a background of nuisance cases as they relate to car horns and alarms).

281. *See, e.g.*, Brautigam, *supra* note 273, at 420-21; Lief, *supra* note 271, at 612-14. This Note will not address products liability claims since they do not appear to have relevance to light pollution.

282. *See* Falzone, *supra* note 272, at 777-79. This type of action is "based on the Fifth Amendment of the U.S. Constitution, which requires compensation for the 'taking' of private property." *Id.* at 777. This Note will not address inverse condemnation claims since most sources indicate that it is not government actors who are generating light pollution, which arguably would constitute a taking.

sources of noise other than a disruptive neighbor.[283]

Similar to current light pollution activities, state and local governments began enacting or strengthening local ordinances to deal with noise.[284] However, these forms of controlling noise have not been without problems. Some hurdles include enforcement, which remains largely intermittent or impossible;[285] vaguely drafted ordinances, which often renders them subjectively enforced;[286] and lack of funding to combat noise pollution.[287] Additionally, there has been some indication from the United States Supreme Court that it may consider anti-noise ordinances to be unconstitutional.[288]

On the federal level, there has been some involvement in alleviating

---

283. *See* Rosenthal, *supra* note 278, at 37-3. "In urban areas particularly, other sources and other victims [of noise] emerged. Common-law remedies proved inadequate." *Id.*

284. *See id.* at 37-3 to 37-4. This is similar to what has been happening with light pollution. *See supra* notes 185-258 and accompanying text. *See also* GRAD, *supra* note 269, at 5-101. "Meaningful governmental regulation aimed at securing a quieter environment is a relatively new development." *Id.* The legal basis for regulating noise originates from the police powers doctrine. This allows state and local governments to act through legislation to protect their constituents' health. *See id.* "If any support for the exercise of the police power were at all necessary, that support is now more readily available" than when state and local governments first addressed noise pollution issues. *Id.* at 5-102.

285. *See* Rosenthal, *supra* note 278, at 37-3. Often ordinances are too vague to enforce, do not provide adequate deterrence to violators since many penalties are misdemeanor penalties, and the authority to control noise has not been granted to a specialized agency. *See* GRAD, *supra* note 269, at 5-104 to 5-105. While street and traffic noise are usually left to police, and health laws that deal with air and water pollution are delegated to some official body, there is rarely a noise-specific agency that can handle noise pollution issues at the state or local level. *See id.* at 5-105. For example, in 1999, the Lexington, Massachusetts, police department had to respond to more than 300 noise-related calls, comprising roughly 2.3 percent of its total calls that year. *See* Alice Hinkle, *Towns Propose Noise Regulations in Quest for Quiet*, BOSTON GLOBE (Northwest Weekly Edition), Mar. 25, 2001, at 1.

286. *See* Falzone, *supra* note 272, at 780. "[M]ost noise ordinances prohibit 'unreasonable' or 'unusual' noise. Due to their subjective nature, ordinances utilizing this language are difficult to enforce." *Id.* With the advent of technology which now can more adequately measure sound, some ordinances were drafted to have maximum decibel outputs. *See id.* However, at least with aircraft, although some sounds may be loud and disruptive, they do not last long enough to violate the ordinances. *See id.*

287. *See* Lief, *supra* note 271, at 616. "[D]espite recent renewed interest in noise, the withdrawal of funding for federal anti-noise efforts may have reduced the incentive to regulate noise at the local level." *Id.*

288. *See id.* at 619-21. (noting that in minority opinions, Justice Frankfurter did not find "aural aggression" to be protected by the Constitution).

noise pollution.[289] The Noise Control Act of 1972[290] was enacted to "promote an environment for all Americans free from noise that jeopardizes their health or welfare."[291] Unlike other environmental statutes, the Noise Control Act did not provide specific abatement goals for the Environmental Protection Agency (EPA), the agency charged with overseeing this Act.[292] In 1978, additional federal grant support was provided for with the addition of the Quiet Communities Act.[293] While the EPA Office of Noise Abatement and Control (ONAC) was eliminated in 1981 due to the permanent termination of funding for the Office,[294] this provision of the Act would provide for a good model to begin understanding and reacting to the problem of light pollution on a national scale.

Perhaps the greatest feature of the Quiet Communities Act was the provision for funding public education about noise pollution,[295] research to discover the effects of noise pollution[296] and technical assistance to state

---

289. *See* Brautigam, *supra* note 273, at 425-26. "The federal role in noise control has generally been modest, and non-existent [in certain areas of noise control]" *Id.* at 425. The Federal Aviation Agency is charged with regulating airport noise. *See id.* at 425-26.

290. 42 U.S.C. §§ 4901-4918 (1994).

291. 42 U.S.C. § 4901(b).

292. *See* Brautigam, *supra* note 273, at 426. Due to the lack of specific guidance, there was no legal reason for states and local governments to address noise pollution and as a result, efforts to curb noise pollution were absent. *See id.*

293. *See* 42 U.S.C. § 4913 (1994). *See also* Brautigam, *supra* note 273, at 426.

294. *See* Brautigam, *supra* note 273, at 426. With the elimination of ONAC, the federal role in noise pollution also ended, however during its active involvement, the federal government had been limited to research and grants. *See id.* Upon closing of ONAC, it is estimated that more than 1,000 community noise abatement programs had to shut down due to the cessation of funding. *See* Falzone, *supra* note 272, at 785-86. Despite the elimination of ONAC, the Noise Control Act remains in effect. *See* Lief, *supra* note 271, at 621-22. "This has created a situation where the federal government is essentially incapable of regulating noise, while the states are preempted from setting their own standards for those products already regulated by the EPA." *Id.* at 622. Within the last ten years, there have been some attempts to re-open ONAC, but not everyone has supported the idea for various reasons. *See id.* In March 2001, a bill was introduced in the 107th Congress aiming to reestablish ONAC and provide for studies and an annual appropriation of $21 million for ONAC activities. *See* Quiet Communities Act of 2001, H.R. 1116, 107th Cong. (2001).

295. *See* 42 U.S.C. § 4913(a) (1994).

> [T]he Administrator [of the Environmental Protection Agency] shall ... (a) develop and disseminate information and educational materials to all segments of the public on the public health and other effects of noise and the most effective means for noise control, through the use of materials for school curricula, volunteer organizations, radio and television programs, publication, and other means ...

Id.

296. *See* 42 U.S.C. § 4913(b) (1994). Funding through this provision could be

and local governments.[297] By replacing "noise" with "light" in most, if not all, provisions of the Quiet Communities Act, one can imagine a "Dark Communities Act." For instance, given the relatively little research about the harmful effects of light pollution on humans, a provision for funding additional research as to what kinds of harm may occur with the increasing problem of light pollution may be necessary.[298] Furthermore, a "Dark Communities Act," if drafted in the same way, would also allow for the study of harmful effects on wildlife.[299] As we have seen, some scientists and bird advocacy groups have already begun to see the harmful effects of light on migratory birds and other animals, but they have relatively little control on stopping those light sources.[300] Arguably, unless one of these animals qualifies for protection under an endangered species statute, there may be no incentive to find out to what degree light pollution is affecting wildlife or what kinds of standards should be set for illumination.

Furthermore, as we have seen in Part V, states and local governments have intermittently and non-uniformly regulated light pollution.[301] By providing grants and other federal actions to these state and local governments, there could be a more coordinated effort to reduce light pollution.[302]

---

used to

> conduct or finance research directly or with any public or private organization or any person on the effects, measurement, and control of noise, including but not limited to –
>
> > (1) investigation of the psychological and physiological effects of noise on humans and the effects of noise on domestic animals, wildlife, and property ...

Id.

297. *See* 42 U.S.C. § 4913(d) (1994). Under this provision, funding could be used to help "develop and implement a national noise environmental assessment program to identify trends in ... ambient levels, and compliance data." *Id.*

298. *See supra* notes 90-94 and accompanying text.

299. *See supra* notes 74-89 and accompanying text.

300. See *supra* notes 74-89 and accompanying text for a discussion on the known harmful effects of light pollution on animals.

301. *See supra* notes 185-258 and accompanying text.

302. *See* 42 U.S.C. § 4913(c) (1994). Administration of a nationwide Quiet Communities Program was to include, but not be limited to

> (1) grants to States, local governments, and authorized regional planning agencies for the purpose of –
>
> > (A) identifying and determining the nature and extent of the noise problem within the subject jurisdiction;
> >
> > (B) planning, developing, and establishing a noise control capacity in such jurisdiction, including purchasing initial equipment;
> >
> > (C) developing abatement plans for areas around major transportation facilities (including airports, highways, and rail yards) and other major stationary sources of noise, and, where appropriate, for the facility or source itself; and

Since many areas of the nation are urban and have always dealt with excessive light or are suburban and have accepted urban sprawl as a way of life, states, cities and towns may continue to enact (or not) legislation that will not deal with the problem adequately so as to solve it nationally.[303] By enacting a "Dark Communities Act," or some degree of federal legislation to deal with this type of pollution in the same way that other federal legislation deals with keeping our air and water clean, we may be able to understand the problem of light pollution before it becomes too great.[304]

VII. CONCLUSION

This Note has revealed that light pollution is a growing problem with growing public awareness.[305] At first glance, it is viewed as more of a problem for astronomers and backyard stargazers than the general public; but as we have seen, light pollution does, in fact, injure many facets of our natural world and general quality of life.[306] The common law does not appear to have the answers for organizations, communities, or those individuals seeking a darker sky.[307] Furthermore, only a handful of states have enacted effective legislation, and local regulations vary widely or do not exist at all.[308] Now is the time for the federal government to become "en-*light*ened" about the potential scope of this growing national problem by funding initial research and providing public education about light pollution.

Kristen M. Ploetz

(D) evaluating techniques for controlling noise (including institutional arrangements) and demonstrating the best available techniques in such jurisdiction;

Id.

303. *See supra* note 258 and accompanying text.

304. *See* GRAD, *supra* note 269, at 5-91. "The Quiet Communities Program appears to be a first step in the direction of a state implementation plan similar to other pollution control legislation." *Id.* Conceivably, the same thing could be said if a similar Act were drafted to deal with light pollution.

305. *See supra* Part III and accompanying notes.

306. *See id.*

307. *See supra* Part IV and accompanying notes.

308. *See supra* Part V and accompanying notes.

Sign on the Dotted Line: Enforceability of Signed Agreements, upon Divorce of the Married Couple, Concerning the Disposition of Their Frozen Preembryos

I. INTRODUCTION

". . . it was a giant step towards the full laboratory control of human reproduction."[1]

Many married couples await the joyous news that they have conceived a baby. Some wait for years to hear the words, but never do. Couples that do not conceive after one year of unprotected intercourse are medically defined as infertile.[2] There are various reasons for infertility,[3] some of which can be treated with hormones and some of which cannot.

In vitro fertilization [hereinafter "IVF"] has been an answer to many of these couples' prayers. Since the first IVF center opened in the United States in 1981,[4] the number of individuals seeking the assistance of medi-

1. SIR ZELMAN COWEN, REFLECTIONS ON MEDICINE, BIOTECHNOLOGY AND THE LAW 32 (1986). The author was discussing the first successful in vitro fertilization technique, which resulted in the birth of baby girl Louise Brown. *Id.* at 32-33.

2. *See* ROBERT K. AUSMAN & DEAN SNYDER, MEDICAL LIBRARY, LAWYER'S EDITION vol. 1 § 1:32(a) (1988).

3. There are many factors that cause infertility. The major cause of infertility is the male factor, which is an inadequate production of normal sperm. This 'male factor' accounts for 30% of the infertility problem. The 'tubal factor' accounts for 25% of the infertility problems. This factor arises when a woman's fallopian tubes do not allow the transports of sperm and oocyte. An abnormal ovulation may also create problems. The 'cervical factor' accounts for 20% of the infertility problems. This factor exists when there are "cervix structure abnormalities of environmental problems" such as bacteria. *See id.* at § 1:32(b). Other considerations may be chronic infections of nutrition deficiencies. In at least 30% of infertile couples, two or more factors exist and are the cause of infertility. *See id.* at § 1:32(b).

4. *See* Alise R. Panitch, Note, *The Davis Dilemma: How to Prevent Battles over*

cal technology to conceive a child has consistently increased. Approximately 80,000 women are artificially inseminated each year in the United States.[5] Few, however, experience the joys of conception and birth through IVF.[6]

IVF allows fertilization of a woman's egg by a man's sperm to occur in a petri dish.[7] The fertilized egg is then implanted in a woman's uterus, which, when successful, results in a child. In 1978, Louise Brown, the first successful child conceived by means of in vitro fertilization, was born in the United Kingdom.[8]

An enhanced method of IVF that involves cryopreservation now exists. Cryopreservation freezes the fertilized egg until it can be implanted into the woman's uterus at the most optimum time.[9] Essentially this procedure allows the actual implantation of the fertilized egg to occur days and even weeks after the eggs have been retrieved and fertilized in the petri dish.[10] Five and one half pound baby girl Zoe Leyland, born in Australia in 1984, was the first baby conceived by this enhanced IVF method.[11]

The lapse of time between egg retrieval and fertilization to the actual implantation into the woman's uterus may be problematic. For example, what if something happened to one or both of the participants after the eggs have been fertilized and stored, awaiting implantation?

Some clinics require that the couple execute a disposition agreement, prior to egg retrieval. Disposition agreements outline various contingencies that may occur to the couple, individually or jointly as a couple, and require that the couple decide how to dispose of the preembryos under the various contingencies listed.[12] The common dispositions are that the preembryos be donated, destroyed, or given to the clinic for research.[13] What happens, however, if what the couple once agreed upon is later cancelled or rescinded by one of them?

This Note examines three recent cases, *A.Z. v. B.Z.*,[14] *J.B. v. M.B.*[15] and

---

Frozen Preembryos, 41 CASE W. RES. 543, 546 (1991).

5. *See* MARTIN GUGGENHEIM ET AL., THE RIGHTS OF FAMILIES 304 (1996).

6. As few as 20 percent of IVF fertilizations result in birth. *See* LAWRENCE J. KAPLAN & ROSEMARIE TONG, CONTROLLING OUR REPRODUCTIVE DESTINY: A TECHNOLOGICAL AND PHILOSOPHICAL PERSPECTIVE 265 (1996).

7. *See infra* Part I.

8. *See* COWEN, *supra* note 1, at 34.

9. *See infra* notes 23-47 and accompanying text.

10. *See* Panitch, *supra* note 4, at 548.

11. *See First Baby Born of Frozen Embryo*, N.Y. TIMES, Apr. 11, 1984, at A16. *See also Australia Disputes Arise on Embryos*, N.Y. TIMES, June 23, 1984, § 1, at 30.

12. *See* John A. Robertson, *Prior Agreements for Disposition of Frozen Embryos*, 51 OHIO ST. L.J. 407, 409-10 (1990).

13. *See id.* at 410.

14. 725 N.E.2d 1051 (Mass. 2000).

Litowitz v. Litowitz,[16] in which married couples who participated in the IVF procedure and executed some form of disposition agreement for their preembryos no longer wished to honor the previous disposition. It examines the reasons given by the courts in holding the agreements contrary to public policy and looks at the possibility of enforcing these agreements under contract law.

Part I of this Note provides a general explanation of the IVF process and the creation of the right to procreate.[17] Part II provides an overview of how preembryos are legally treated and how past cases have handled disposition agreements.[18] Part III focuses on the three recent cases that have addressed the issue of the enforceability of disposition agreements, and outlines the courts' arguments.[19] Part IV focuses on contract law, how various other familial agreements are treated by courts, and finally how applying contract law would have affected the outcomes of these recent cases.[20]

This Note concludes that disposition agreements should generally be enforced under contract law when there is a previous disposition agreed upon by both parties.[21] Enforcement of the agreement under contract law, however, must consider the various defenses available in contract law and apply them accordingly. In situations were the disposition is not clearly ascertainable or when the court is made to decide the disposition, the preembryos should be destroyed.[22]

II. IN VITRO FERTILIZATION AND CRYOPRESERVATION

IVF, combined with the cryopreservation procedure has given many infertile couples an opportunity at parenthood. The IVF process initially stimulates the woman's ovaries so that she may release multiple eggs.[23] These eggs are then surgically removed[24] and placed in a petri dish, where sperm cells collected from the man are added to the petri dish so that fer-

15. 751 A.2d 613 (N.J. 2000).
16. 10 P.3d 1086 (Wash. 2000).
17. *See infra* notes 23-47 and accompanying text.
18. *See infra* notes 48-113 and accompanying text.
19. *See infra* notes 114-211 and accompanying text.
20. *See infra* notes 212-319 and accompanying text.
21. *See infra* notes 320-23 and accompanying text.
22. *See id.*
23. *See* Mark C. Haut, Note, *Divorce and the Disposition of Frozen Embryos*, 28 HOFSTRA L. REV. 493, 495 (1999).
24. There are two procedures used to remove eggs from the woman. One method is by laparoscopy, which uses general anesthesia or local anesthesia with mild sedation. *See* AUSMAN & SNYDER, *supra* note 2, at § 1:40(e). Another possible method to remove the eggs is through ultrasound – directed needle aspiration. *See* Panitch, *supra* note 4, at 547 n.30.

tilization of the eggs may occur.[25] After fertilization, the zygote[26] will develop in the petri dish until it reaches a two-to-eight-celled stage, at which point it is called a "preembryo."[27] This stage is also the normal stage of implantation into the woman's uterus by a cervical catheter.[28] However, if it is determined that implantation would be more successful at a later date, the preembryos are frozen in liquid nitrogen.[29] This procedure is called cryopreservation.[30] It can prolong implantation of the preembryos in the uterus, [31] as well as allow subsequent implantations if the first attempt is unsuccessful.[32] Cryopreservation also lowers the cost of IVF as well as the physical and emotional burden a woman experiences.[33]

Cryopreservation, however, does have one major disadvantage: Prolonging implantation leaves the couple vulnerable to the unanticipated reality of death, separation, or divorce. Some IVF clinics provide a consent form, which also serves as an agreement between the couple regarding the disposition of their frozen preembryos in the event that an unanticipated circumstance occurs. Courts, however, do not always enforce these agreements when conflict arises.

A. Right to Procreate

One of the basic civil right's a person has is the right to procreate.[34]

---

25. *See* Panitch, *supra* note 4, at 547.

26. A zygote is defined as the joining of male and female reproductive germ cells to produce a new organism. *See* THE AMERICAN HERITAGE DICTIONARY 795 (1983).

27. *See* Panitch, *supra* note 4, at 547. It is called a preembryo because implantation has not occurred and advanced cell differentiation has not begun. *See id.* The American Fertility Society also uses the term "preembryo." *See Litowitz v. Litowitz*, 10 P.3d 1086, 1088 n.1. (Wash. App. 2000).

28. *See* KAPLAN & TONG, *supra* note 6, at 264.

29. *See* Haut, *supra* note 23, at 495. This procedure also allows the physician to control the number of preembryos transferred at one time, which can reduce the possibility of multiple births and the risks to both the mother and the fetus. *See* ANDREA L. BONNICKSON, IN VITRO FERTILIZATION: BUILDING POLICY FROM LABORATORIES TO LEGISLATION 30 (1989).

30. See Panitch, *supra* note 4, at 547.

31. *See* Panitch, *supra* note 4, at 548.

32. *See* Jennifer Marigliano Dehmel, Note, *To Have Or Not To Have: Whose Procreative Rights Prevail In Disputes Over Dispositions of Frozen Embryos?* 27 CONN. L. REV. 1377, 1377 (1995). Usually multiple eggs are retrieved from the woman because of the hormones, however, not all are implanted at once, so as to give the woman a second chance at implantation if the first is not successful. *See id.*

33. *See* Haut, *supra* note 23, at 495.

34. *See Skinner v. Oklahoma*, 316 U.S. 535, 541 (1942) (holding that absent a compelling state interest a state can not deprive individuals of their right to procreate by means of sterilization).

This right stems from the fundamental right to privacy established in the Constitution.[35] It is also grounded in the belief that a person's marital bedroom should be free from governmental intrusion.[36]

One of the earliest United States Supreme Court cases to discuss personal rights not explicitly stated in the Constitution or the Bill of Rights is *Meyer v. Nebraska*.[37] In *Meyer*, the Court stated that the liberty guaranteed by the Fourteenth Amendment did consist of "the right of the individual to contract, . . . to marry, . . . [and] to establish a home and bring up children."[38] Thus, although these rights were not necessarily stated in the constitution, they could and would be recognized.

Other cases that followed *Meyer v. Nebraska* further reinforced the right to procreate: *Griswold v. Connecticut*[39] held that married couples had a fundamental right to procreate because it was within a "zone of privacy" expected by the individual and protected by the Constitution.[40] Seven years later, *Eisenstadt v. Baird*[41] extended the *Griswold* holding to include unmarried couples.[42] *Roe v. Wade*,[43] perhaps the most well known case in this area, held that a woman had the right to procreational autonomy, without state interference.[44] In *Planned Parenthood v. Danforth*[45] the Court held that a woman did not need spousal consent before terminating her pregnancy, reasoning that "the woman . . . physically bears the child and . . . is more directly and immediately affected by the pregnancy [, thus] the balance weighs in her favor."[46] Finally, in 1992, the United States Supreme Court "recognized protection accorded to liberty relating to intimate relationships, the family, and decisions about whether or not to beget or bear a child" in *Planned Parenthood v. Casey*.[47]

This judicial precedence has established the right to procreate, as well as the right not to procreate, and granted constitutional protection for these rights. It has given individuals protection from state involvement in their

35. U.S. CONST. amend. IV.
36. *See* Griswold v. Connecticut, 381 U.S. 479, 485 (1965).
37. 262 U.S. 390 (1923).
38. *Id.* at 399.
39. *See supra* note 36.
40. *Griswold*, 381 U.S. at 485.
41. 405 U.S. 438 (1972).
42. *Id.* at 453-54
43. 410 U.S. 113 (1973).
44. *Id.* at 154. The right not to procreate by means of abortion, is however subject to the fetus' viability, at which point the state has a compelling state interest to protect the fetus and thus can infringe upon the woman's right not to procreate. *Id.* at 163.
45. 428 U.S. 52 (1976).
46. *Id.* at 71.
47. 505 U.S. 833, 857 (1992).

intimate relationships, and limited the possibility of government intrusion by establishing the need of a compelling state interest to deprive someone of this constitutional right.

III. LEGAL STATUS OF FROZEN PREEMBRYOS

The legal status of frozen preembryos can be categorized into three distinct views.[48] The first viewpoint, and the most conservative, is that preembryos are human beings and have a right to life because life begins at conception.[49] On the other extreme, preembryos are viewed as the property of the couple and therefore subject to equitable distribution.[50] The intermediate view is that preembryos deserve special respect because of their potential for human life.[51]

A. Preembryos as Human Beings

Proponents of this right to life view consider a preembryo a human being at the time of conception and give it full rights and protections under the law.[52] This view sees natural conception as the same as artificial or assisted conception because it focuses on the ultimate result and not the method by which it was achieved.[53] It also places the interests of the preembryo first, using reasoning similar to the "best interest of the child" doctrine.[54] Finally, this view favors the individual seeking to give the preembryos life through donation or implantation over the destruction of the preembryos.

Some scholars argue against the belief that preembryos are human beings. They claim that preembryos cannot be human beings because they are not autonomous or self-conscious.[55] Still other scholars believe a preembryo *is* human, however, they argue that it does not deserve legal protection like a human being does.[56]

---

48. *See* Panitch, *supra* note 4, at 553. *See also* Haut, *supra* note 23, at 497.
49. *See* Haut, *supra* note 23, at 497.
50. *See* Panitch, *supra* note 4, at 553. *See also* Haut, *supra* note 23, at 497, 499.
51. *See supra* note 48.
52. *See* Panitch, *supra* note 4, at 555.
53. *See* DONALD EVAN, CONCEIVING THE EMBRYO: ETHICS, LAW, AND PRACTICE IN HUMAN EMBRYOLOGY 47 (1996).
54. *Davis v. Davis*, 842 S.W.2d 588, 594 (1992). The "best interest of the child" doctrine is usually used in awarding custody. Section 402 of the Uniform Marriage and Divorce Act outlines the relevant factors the court must consider. *See* UNIF. MARRIAGE AND DIVORCE ACT § 402 (1998).
55. *See* EVAN, *supra* note 53.
56. *See id.* at 48, 152 (emphasis added).

B. Preembryos as Property

Another viewpoint on the status of preembryos is that they are property and should be handled as such. This viewpoint focuses on the rights of the gamete[57] providers and allows preembryos no state protection.[58] This argument rests on the decision of *York v. Jones*,[59] in which the court held that the clinic could not refuse to transfer the frozen preembryos, upon the couple's request, to another clinic, which terminated the bailor-bailee relationship.[60] The couple had a property interest in the preembryos, because they were the gamete providers.[61] *Hecht v. Kane* [62] also supports this property view. In *Hecht*, the California Appeals Court concluded that the decedent had an interest in his sperm that could be distributed as property within his estate.[63]

Many courts however, have been unwilling to classify frozen preembryos as property because of the moral and ethical implication.[64] The Ethics Committee of the American Fertility Society (1990) also has considered this view of preembryos as property unethical.[65] They reconcile this conclusion by pointing out the prohibition in many states of the sale of sperm, eggs and preembryos.[66]

C. Preembryos as Deserving of Special Respect

The third view, which is a compromise between the property view and the human life view, is most widely supported.[67] It grants preembryos a "special respect" because of their potential for human life.[68] Proponents of this view consider preembryos not yet human, because they do not have

57. Gamete is defined as a mature sperm of egg, capable of participating in fertilization. *See* THE AMERICAN HERITAGE DICTIONARY 289 (1983).

58. *See* Haut, *supra* note 23, at 499.

59. 717 F. Supp. 421 (1989).

60. *Id.* at 427.

61. *See id.* at 429.

62. 16 Cal. App. 4th 836 (1993). In *Hecht*, decedent deposited 15 vials of sperm in a sperm bank. In the same month, he committed suicide, leaving a clinic agreement, a will, and a letter, all three of which stated he wanted his long-term girlfriend to use the sperm to conceive after his death. His two adult children contested his request. *See id.* at 840-41.

63. *See id.* at 850.

64. *See* A. Mechele Dickerson, *From Jeans to Gene: The Evolving Nature of Property of the Estate*, 15 BANK. DEV. J. 285, 316 (1999).

65. *See* Howard W. Jones Jr., *Policy Considerations For Cryopreservation in In Vitro Fertilization Programs, in* ISSUES IN REPRODUCTIVE TECHNOLOGY 209, 211 (Helen Bequaert Holmes ed., 1994).

66. *See id.*

67. *See* Dehmel, *supra* note 32, at 1384. *See also* Panitch, *supra* note 4, at 561.

68. *See* Haut, *supra* note 23, at 500.

a differentiated nervous system,[69] yet much more than mere tissue because they are a "powerful symbol or reminder of the unique gift of human existence."[70]

IV. HISTORICAL LEGAL TREATMENT OF DISPOSITION AGREEMENTS

There are only two cases, *Davis v. Davis*[71] and *Kass v. Kass*,[72] that have directly addressed the issue of disposition agreements of frozen preembryos. *Kass* involved a signed prior agreement between the couple,[73] however *Davis* did not. Consequently, both cases held that disposition agreements should be enforced.

A. *Davis v. Davis*[74]

Shortly after their marriage, Mary Sue Davis became pregnant, but because it was an ectopic pregnancy, she had to have a fallopian tube removed.[75] Four other ectopic pregnancies followed and so her remaining fallopian tube was ligated, making it impossible for her to conceive a child naturally.[76] After an unsuccessful attempt at adoption, the couple began IVF treatment.[77]

The Davises attempted six IVF implantations with no success.[78] After a brief period of postponed treatment, the couple once again began IVF treatment, which resulted in the retrieval of nine eggs.[79] The first transfer of frozen preembryos was unsuccessful, and before another attempt was made, Junior Davis filed for divorce,[80]leaving the unimplanted preembryos in storage at the IVF clinic.[81]

The couple did not enter into an agreement concerning the disposition of their preembryos.[82] However, the trial court awarded "custody" to Mrs.

---

69. *See* Dehmel, *supra* note 32, at 1383.

70. John A. Robertson, *Reproductive Technology and Reproductive Rights: In the Beginning: The Legal Status of Embryos*, 76 VA. L. REV. 437, 447 (1990).

71. 842 S.W.2d 588 (Tenn. 1992).

72. 663 N.Y.S.2d 581 (A.D. 2 Dept. 1997).

73. *Se e infra* notes 94-108 and accompanying text.

74. 842 S.W.2d 588 (Tenn. 1992).

75. *See Davis*, 842 S.W.2d at 591.

76. *See id.*

77. *See id.*

78. These six attempts at IVF were done without the technique of cryopreservation. *See id.* at 591-92. The Davis' went through the first six IVF attempts using the original IVF technique that requires implantation within a few days of egg retrieval and fertilization. *See id.*

79. *See Davis*, 842 S.W.2d at 592.

80. *See* Davis v. Davis, 842 S.W.2d 588, 592 (Tenn. 1992).

81. *See id.* at 589.

82. *See id.* at 590.

Davis so as to "bring these children to term through implantation."[83] The Court of Appeals reversed the trial court's holding, stating that Mr. Davis had a "constitutionally protected right not to beget a child where no pregnancy has taken place . . .."[84] It also stated that they shared an interest in the remaining preembryos and gave the couple joint control over the preembryos' disposition.[85]

The Tennessee Supreme Court reviewed the case because the Court of Appeals did not provide adequate guidance as to how to resolve the issue.[86] The Tennessee Supreme Court agreed with the Court of Appeals that the trial court was incorrect in determining that frozen preembryos were human beings.[87] The court held that they were not "human beings" or property, but that they deserved special respect because of their potential for human life.[88] It also stated that "an agreement regarding disposition of any untransferred preembryos in the event of contingencies . . . should be presumed valid and should be enforced as between the progenitors," because as progenitors, they retained decision-making authority.[89] However, since the couple did not have an agreement regarding the disposition of their frozen preembryos, the court applied a balancing test which weighed each provider's interest and determined that Mr. Davis' interest in becoming a parent outweighed Mrs. Davis' interest in donating the frozen preembryos to an infertile couple.[90] The court concluded that the manner to resolve such a dispute is to first look at the preferences of the progenitors.[91] If there is a dispute, then the prior agreement should be carried out.[92] However, if there is no prior agreement, then the interests of each party should be weighed.[93] Thus, a rule was presumptively estab-

---

83. *See id.* at 589. Testimony by Dr. Jerome Lejeune, a French geneticist, influenced the court's conclusion that the fertilized eggs were "children" because he referred to then as "early human beings" and "tiny persons." *Id.* at 593. Other experts disputed Dr. Lejeune's testimony at trial. *Id.*

84. *Id.* at 589.

85. *See id.*

86. *See* Davis v. Davis, 842 S.W.2d 588, 590 (Tenn. 1992). Since to couple obviously did not agree on the disposition of their frozen preembryos joint control would not resolve their dispute.

87. *Id.* at 594.

88. *See id.* at 597. The court addressed the issue as to whether frozen preembryos were property or persons. *See id.* at 594-97.

89. *See id.*

90. *See id.* at 604. By the time the case reached the Tennessee Supreme Court, Mrs. Davis had remarried and decided not to use the frozen preembryos for implantation, but rather for donation to an infertile couple. *See id.* at 590.

91. *See* Davis v. Davis, 842 S.W.2d 588, 604 (Tenn. 1992).

92. *See id.*

93. *See id.*

lished.

B. *Kass v. Kass*[94]

Mrs. Kass had difficulty conceiving because of a medical condition, so the couple decided to attempt conception through IVF.[95] The couple underwent ten unsuccessful attempts.[96] Prior to the last procedure, the couple signed a consent form pertaining to the disposition of the frozen preembryos.[97] The consent form stated, in part,

In the event of divorce, we understand that legal ownership of any stored pre-zygotes must be determined in a property settlement and will be released as directed by order of the court of competent jurisdiction . . ..

> In the event that we no longer wish to initiate a pregnancy or are unable to make a decision regarding the disposition of our stored, frozen pre-zygotes . . .
>
> [o]ur frozen pre-zygotes may be examined by the IVF Program for biological studies and be disposed of by the IVF Program for approved research investigation as determined by the IVF Program.[98]

Days after being notified that the last implantation was unsuccessful, the couple began the process of an uncontested divorce.[99] The divorce documents stated the exact same thing the consent form did. Shortly thereafter, Mrs. Kass changed her mind and wrote a letter to the clinic and her IVF physician stating her opposition to the destruction or release of the remaining preembryos.[100] She then filed a complaint seeking "sole custody" of the preembryos.[101] The trial court made several determinations[102] and ultimately granted Mrs. Kass "the exclusive right to determine the fate of the subject pre-embryos."[103]

On Mr. Kass' appeal, the New York Court Appeals reviewed the case and held that the parties' mutual intent expressed in both the informed consent document and the uncontested divorce document should be given

---

94. 663 N.Y.S.2d 581 (1997).
95. *See id.* at 583.
96. *See id.*
97. *See id.*
98. *Id.* at 584.
99. *See id.*
100. *See* Kass v. Kass, 663 N.Y.S. 2d 581 (1997).
101. *Id.* at 585.
102. In brief, the court addressed the legal status of preembryos; found that the husband's procreative right terminated at fertilization; held the informed consent document provided no guidance in the event of a divorce; and determined that the uncontested divorce agreement did not constitute a waiver by Mrs. Kass. *See id.* 584.
103. *Id.*

effect.[104] Using *Davis*, the only other case to address the same issue, the court stated that the consent form executed prior to the last IVF attempt was a "manifestation of [the] mutual intent [that] exist[ed]"[105] and should be enforced because it was an "unequivocal statement of [their] intent."[106] The court held that:

> . . . where the parties have indicated their mutual intent regarding the disposition of the pre-zygote in the event of the occurrence of a contingency, that decision must be scrupulously honored, and the courts must refrain from any interference with the parties' expresses wishes.[107]

The court also stated that only if there is no disposition agreement setting forth the intent of the parties should a balancing test be used.[108]

C. Court View of Disposition Agreements Based on *Davis* and *Kass*

Based on the holdings of *Davis*[109] and *Kass*,[110] one could conclude that courts were inclined to enforce prior agreements made between the married couple concerning the disposition of their frozen preembryos. In *Davis*, the court stated that a disposition agreement should be presumed valid and enforced because as progenitors the couple retains decision-making authority and their wishes should be enforced;[111] and provided guidance if there was no prior agreement.[112]

In *Kass*, the prior disposition agreement was found valid and enforced by the court because the consent agreement executed prior to the IVF procedure manifested the couple's mutual intent regarding the disposition of their frozen preembryos.[113] So was the issue settled? Was it safe to say that prior agreements between a couple concerning the disposition of their frozen preembryos would be enforced?

V. 2000 - A CONTRARY VIEW BY COURTS

In the year 2000, courts in three states decided whether a prior agreement between a couple concerning the disposition of their frozen preembryos was enforceable.[114] In March, the Massachusetts Supreme Judicial

---

104. *See id.*
105. *Id* at 585.
106. Kass v. Kass, 663 N.Y.S. 2d 581, 587 (1997).
107. *See id.* at 590.
108. *See id.* at 586-87.
109. *See supra* notes 74-93 and accompanying text.
110. *See supra* notes 94-109 and accompanying text.
111. *See supra* notes 87-103 and accompanying text.
112. *See supra* notes 74-93 and accompanying text.
113. *See supra* notes 95-108 and accompanying text.
114. Massachusetts in *A.Z. v. B.Z.*; New Jersey in *J.B v. M.B.*; and Washington in *Litowitz v. Litowitz.*

Court was asked to address the issue; three months later, the Superior Court of New Jersey considered the same issue; and four months after New Jersey addressed the issue, Washington did as well. All three courts held that the prior agreements were unenforceable, thus choosing not to follow *Davis* or *Kass*.

A. *A.Z. v. B.Z.*[115]

Early on, the couple encountered difficulties conceiving and underwent fertility testing.[116] After a miscarriage, which resulted in the removal of one fallopian tube, the couple underwent further fertility treatment.[117] There was a second ectopic pregnancy and the wife's remaining fallopian tube was removed.[118] The couple then underwent IVF treatments for approximately three years, which resulted in the birth of twin daughters and more preembryos than necessary.[119] Two vials of frozen preembryos remained at the clinic.[120]

Approximately three years later the wife had one of the vials thawed and implanted without the husband's knowledge, which did not result in a pregnancy.[121] The couple later separated and the husband filed for divorce; as well as a motion to permanently enjoin his wife from implanting the remaining vial that contained four frozen preembryos.[122]

To participate in the IVF program, the couple had to sign a consent form before each egg was retrieved.[123] The consent form required the couple to decide the disposition of their frozen preembryos under certain contingencies.[124] The couple could opt to donate or destroy the preembryos, as well as write in their own alternative disposition for the frozen preembryos.[125] The form also stated that the couple could change the disposition chosen at any time so long as they both conveyed it in writing.[126]

The couple signed a total of seven consent forms, all of which stated that if they "'[s]hould become separated, [they] both agree[d] to have the embryo(s) . . . return[ed] to [the] wife for implant[ation].'"[127] The first

115. 725 N.E.2d 1051 (Mass. 2000).
116. *See id.* at 1052.
117. *See id.*
118. *See id.* at 1053.
119. *See id.*
120. *See id.*
121. *See* A.Z.v. B.Z., 725 N.E.2d 1051, 1053 (Mass. 2000).
122. *See id.*
123. *See id.*
124. *See id.* at 1054.
125. *See id.*
126. *See* A.Z.v. B.Z., 725 N.E.2d 1051, 1054 (Mass. 2000).
127. *Id.*

agreement, signed in 1988, was completed by the wife and signed by both the husband and the wife.[128] The husband signed six additional consent forms (for six additional procedures) that were blank, and were later filled out by the wife.[129] In each instance, the wife completed the form by filling in the option of "[s]hould we become separated," with the same response as that given on the first form- the preembryos were to be returned to the wife for implantation.[130]

1. Probate Court

The Probate Court held the agreement unenforceable because of changed circumstances since the last disposition agreement signed in 1991.[131] The judge determined that "[n]o agreement should be enforced in equity when intervening events have changed the circumstances such that the agreement which was originally signed did not contemplate the actual situation now facing the parties."[132] Consequently, the judge held the agreement unenforceable and that the issue should be settled by balancing the wife's interest in procreation against the husband's interest in avoiding procreation.[133] The judge found that the husband's interest in avoiding procreation outweighed the wife's interest in procreating and granted the husband's request for a permanent injunction preventing the wife from using the frozen preembryos.[134]

2. Supreme Judicial Court

The wife appealed the decision of the Probate and Family Court enjoining her from using the frozen embryos stored in the clinic.[135] In reviewing the case, the Supreme Judicial Court of Massachusetts questioned the purpose of the agreement, and concluded that the form's purpose was to give the clinic guidance.[136] The court, however, chose not to enforce the agreement for several other "independent reasons."[137]

The court gave five reasons. First, the purpose of the agreement was to explain the benefits and risks of the procedure and to provide the clinic

128. *See id.*

129. *See id.*

130. *Id.*

131. *See id.* The court concluded that the birth of twins, the wife's restraining order against her husband, the husband's filing for divorce, and the wife's attempt to conceive by thawing a vial were changed circumstances. *See id.* at 1054-55.

132. A.Z.v. B.Z., 725 N.E.2d 1051, 1055 (Mass. 2000).

133. *See id.* at 1055.

134. *See id.*

135. *See id.* at 1052

136. *See id.* at 1056

137. *See id.*

with guidance should the couple, as a unit, no longer wish to use the preembryos.[138] Second, the consent form contained no duration provision and the court would not assume the consent form was intended to govern four years after it was executed.[139] Third, the term "'[s]hould we become separated,'" used in the consent form, was ambiguous because legally speaking, "separated" has a different meaning than "divorced" and the form only provided for a disposition in case of separation.[140] Fourth, the court questioned whether the document represented the intentions of both donors because of the spouses' conduct.[141] Finally, the court determined that the agreement was not a separation agreement that was binding in a divorce because it had no custody or support provisions.[142]

Despite the numerous reasons the court gave for not enforcing the agreement, and the various ambiguities it addressed, the court stated that even if the agreement was unambiguous, it still would not have enforced the agreement because it was against public policy to compel a gamete donor to become a parent against his or her will.[143] The court further stated it was reluctant to enforce "prior agreements to enter into familial relationships . . . against individuals who subsequently reconsider their decision."[144] Thus, the Supreme Judicial Court of Massachusetts held that public policy would not allow the court to enforce the disposition agreement before it.[145]

B. *J.B. v. M.B.*[146]

The couple in this case was married for approximately four years.[147] Their initial attempts at conception were unsuccessful.[148] Their reproduction problems were attributed to the wife's endometriosis and the blockage of one fallopian tube.[149] In an attempt to conceive, the couple began IVF treatment.[150] During the IVF treatment the wife became pregnant,[151] and

138. A.Z. v. B.Z., 725 N.E.2d 1051, 1056 (Mass. 2000).
139. *See id.* at 1057
140. *Id.*
141. *See id.* Before signing the first form, the wife inquired about the disposition section of the form; also, the husband signed the forms in blank. *See id.* at 1057.
142. *See id.* at 1057.
143. *See id.*
144. A.Z. v. B.Z., 725 N.E.2d 1051, 1059 (Mass. 2000).
145. *See id.*
146. 751 A.2d 613 (N.J. Super. A.D. 2000).
147. *See id.* at 615.
148. *See id.*
149. *See id.*
150. *See id.*
151. *Id.* at 615. It was unknown whether the pregnancy was the result of the IVF treatment or natural conception. *See id.*

gave birth in March of 1996.[152] The couple separated six months later, and a judgment of divorce was entered in 1998.[153] The divorce decree "'reserved a decision concerning the parties' cryopreserved preembryos/embryos.'"[154] Soon after the divorce decree was entered, the wife moved for summary judgment to destroy the frozen preembryos.[155]

The couple signed a consent form that outlined different contingencies and the disposition of the frozen preembryos under the specific contingencies, one of which was divorce.[156] The consent form stated that "all control, direction and ownership of [the] tissues [would] be relinquished to the IVF Program" if: (1) a dissolution of the marriage by court order occurred, unless the court specifies who is to take control of the preembryos; or (2) death of both individuals, unless provided for otherwise in a will; or (3) the patient is no longer capable to sustain a normal pregnancy, but the couple is given two years to determine a host womb; or (4) by "our/my election," which must be in writing; or (5) the periodic payments for maintenance of the preembryos are not made.[157] The consent form also had a provision stating that "the control and disposition of the embryos belongs to the Patient and her Partner."[158]

1. Trial Court

The trial judge determined that the couple underwent IVF treatment so as to conceive within the context of their marriage.[159] However, since the marriage no longer existed, the reason to preserve the frozen embryos no longer applied.[160] The court's decision also relied on the fact that the husband had no reproductive problems, and thus, could conceive a child with another woman.[161] Ultimately, the husband appealed the trial court's decision that granted summary judgment in favor of the wife who sought the destruction of the frozen preembryos.[162]

2. Superior Court - Appellate Division

The Appellate Division of the Superior Court of New Jersey focused on

152. J.B. v. M.B., 751 A.2d 613, 615 (N.J. Super. A.D. 2000).
153. *See id.*
154. *Id.*
155. *See id.*
156. *See id.* at 616.
157. *Id.*
158. J.B. v. M.B., 751 A.2d 613, 615 (N.J. Super. A.D. 2000).
159. *See id.* at 616
160. *See id.*
161. *See id.*
162. *See id* at 614-15.

the main argument each party raised, which was the right to procreate (husband's argument) and the right not to procreate (wife's argument).[163] The court also stated that "even if the wife were relieved of the financial and custodial responsibility" for the child born from the frozen preembryos, the biological connection would still exist, therefore, her right to not procreate would be impaired.[164] Whereas the husband's right to procreate would not be impaired, because he is able to reproduce.[165]

Interestingly, the court did not use the constitutional grounds of one's right to privacy or one's right to procreate to resolve the case.[166] Instead it used the case of *In the Matter of Baby M* as guidance in its decision.[167] By using the *Baby M* case the court concluded that it could not compel the wife to become a parent because termination of her parental rights would result if there was a successful implantation.[168] Thus, the court did not seem too concerned about the implantation itself violating her rights, but rather that the termination of parental rights to a resulting child, would violate her rights.

The Superior Court also agreed with the holding of the Supreme Judicial Court of Massachusetts in the *A.Z.* case, that a "contract to procreate" was contrary to public policy and unenforceable because it would force parenthood on an individual.[169] As a result, the New Jersey Superior Court held the disposition agreement executed by J.B. and M.B. unenforceable because of the same public policy used in *A.Z.*[170]

C. *Litowitz v. Litowitz*[171]

The *Litowitz* case is factually different from *A.Z. v. B.Z.* and *J.B. v. M.B.* The couples involved in both *A.Z.* and *J.B.* each provided their own eggs and sperm and used the wife's uterus as the host womb. In *Litowitz*, the eggs were donated by another woman, fertilized by Mr. Litowitz's sperm to create the preembryos, and implanted into a surrogate.[172] Mrs. Litowitz had no involvement in the IVF process. She did not provide the eggs or the host womb.[173]

---

163. *See id.* at 618.
164. *See* J.B.v. M.B., 751 A.2d 613, 619 (N.J. Super. A.D. 2000).
165. *See id.*
166. *See id.* at 620.
167. *See id.*
168. *See id.*
169. *See id.* at 619.
170. *See* J.B. v. M.B., 751 A.2d 613, 619-20 (N.J. Super. A.D. 2000).
171. 10 P.3d 1086 (Wash. App. Div. 2000).
172. *See id.* at 1087.
173. Mrs. Litowitz had two biological children of her own. The couple conceived a child in 1980 and married in 1982. After the birth of the their child Mrs. Litowitz

David and Becky Litowitz used an egg donor and a surrogate to attempt to conceive.[174] Of the donated eggs, which were later fertilized by Mr. Litowitz's sperm, three preembryos were implanted into the surrogate and two preembryos were cryogenically preserved for future implantation if necessary.[175] A child was conceived from the three implanted preembryos but by the time the child was born Becky and David had separated.[176]

The couple entered into two separate agreements. The first agreement was with the egg donor.[177] The second agreement the Litowitzes entered into was with the clinic that cryogenically preserved the preembryos.[178] The informed consent agreement with the clinic provided that:

> . . . any decision regarding the disposition of our [preembryos] will be made by mutual consent. In the event we are unable to reach a mutual decision regarding the disposition of our [preembryos], we must petition to a Court of competent jurisdiction for instructions concerning the appropriate disposition of our [preembryos].[179]

The couple also agreed that under certain circumstances the preembryos would not be permitted to develop.[180] There was, however, no provision covering the dissolution of the marriage.[181]

1. Trial Court

At trial, Becky asked to be awarded the preembryos under the egg donor agreement.[182] David, however, wanted to put the preembryos up for

---

had a hysterectomy, rendering her unable to have another natural birth or donate any eggs. After their marriage in 1982, the couple wanted to have another child and looked at surrogacy as their way of conceiving. *See id.* at 1088.

174. *See id.* at 1088.

175. *See id.* at 1089. In total, five eggs were donated and fertilized by Mr. Litowitz. *See id.* at 1089.

176. *Id.* at 1088.

177. *See* Litowitz v. Litowitz, 10 P.3d 1086, 1089 (Wash. App. Div. 2000). The egg donor contract defined the Litowitzes as the "intended parents" and provided that:

> All eggs produced by the Egg Donor pursuant to this Agreement shall be deemed the property of the Intended Parents and as such, the Intended parents shall have the sole right to determine the disposition of said egg(s). In no event may the Intended Parents allow any other party the use of said eggs without express written permission of the Egg Donor.

Id. at 1089.

178. *See id.*

179. *Id.*

180. *See id.* at 1089. One example covered under this provision was the death of both. *See id.*

181. *See id.* at 1089.

182. *See supra* note 178 and accompanying text.

'adoption.'[183] The trial court awarded the frozen preembryos to David "with orders to use his absolute best effort for adoption to a two-person family outside of the state of Washington . . .."[184] In reaching its decision to award the preembryos to Mr. Litowitz, the trial court applied a best interest of the child standard.[185]

2. Washington Court of Appeals

Mrs. Litowitz appealed the trial court's decision, which granted Mr. Litowitz control of the frozen preembryos.[186] The Washington Appeals Court, however, affirmed the lower court's decision.[187] The court used all the previous cases that had addressed the issue of dispositions agreement for frozen preembryos.[188] The court first looked at *Kass* and held that unlike *Kass*, there was no express disposition agreement to enforce.[189] The court then compared the similarities of *Davis* to its present case and found that in both cases there was no evidence that the party seeking to prevent the use of the frozen preembryos "intended to pursue reproduction outside the . . . marital relationship."[190] The court also used *Davis* to determine that Becky had no constitutional right to procreate because she did not contribute any gametes.[191] Finally, when examining the *A.Z.* case with the case before it, the court concluded that it would not enforce an agreement that "compel[ed] one donor to become a parent against his or her will" because of similar public policy arguments in *A.Z.* and *J.B.*[192]

VI. POLICY ARGUMENTS IN THE TWENTY-FIRST CENTURY CASES: FORCED PARENTHOOD

The policy governing the holdings in all three cases discussed above were identical. In *A.Z.*, *J.B.*, and *Litowitz* the courts held that the agree-

183. *See* Litowitz v. Litowitz, 10 P.3d 1086, 1090 (Wash. App. Div. 2000).

184. *Id.* at 1089. The trial court's ruling was later reduced to a court order, pending all appeals. *See id.*

185. *See id.*

186. *See id.* at 1088.

187. *See id.* at 1088, 1093. The Appeals court did, however, state that the best interest of the child test that the trial court used was incorrect. *See id.* at 1088.

188. *See id.* at 1090.

189. *See* Litowitz v. Litowitz, 10 P. 3d 1086, 1091 (Wash. App. Div. 2000).

190. *Id.*

191. *See id. at* 1092. The court used the *Davis* reasoning that stated that "[t]he case would be closer if Mary Sue Davis were seeking to use the preembryos herself, but only if she could not achieve parenthood by any other reasonable means." *Id.*

192. *See id.* at 1091. The Washington Court of Appeals stated that it would not read into the contracts an implied promise to continue family planning after dissolution. *See id.* at 1091.

ments were contrary to public policy because a person could not be compelled to become a parent against his or her will. Each court, however, had a unique analysis in reaching its ultimate decisions.

A. The Supreme Judicial Court's Argument of 'Forced Parenthood' in *A.Z. V. B.Z.*

In *A.Z. v. B.Z.*, the court stated that as a matter of public policy it would not enforce the disposition agreement with the clinic as binding between the couple because it affected the future familial relationship.[193] The court listed various independent reasons for its decision to hold the disposition agreement unenforceable, some of which addressed what the court believed were ambiguities in the agreement.[194] The court stated, however, that "even had the husband and the wife entered into an unambiguous agreement . . . [it] would not enforce an agreement that would compel one donor to become a parent against his or her will,"[195] thus, making it clear that the court decided this cases strictly on public policy grounds and not for the various reasons given.

Because the court did not base its holding on the list of reasons it gave, its rationale is unclear. It is clear, however, that the court did not enforce the prior agreement because it wanted to allow the person time to change his or her mind.[196] The court used *R.R. v. M.H.*, a Massachusetts surrogacy case[197] to explain its reason for allowing a period for reconsideration and to emphasize the need for a "reasonable" waiting period.[198] In a surrogacy situation, it seems as though that time period is shortly after birth, thus after the contract is satisfied.[199]

In cases involving disposition agreements, however, the agreement is to jointly decide the disposition of the frozen preembryos in case a life-changing event occurs to one or both of the parties before implantation of the preembryos. Therefore it only follows, that the 'time period' analogous to the surrogacy 'time' the *A.Z.* court believed was necessary to be able to reconsider one's decision, is a few days after the preembryos are cryogenically preserved and stored in the clinic.

---

193. A.Z. v. B.Z., 725 N.E.2d 1051, 1058 (Mass. 2000).

194. *See id.* at 1057.

195. *See id.* at 1057.

196. *See id.* at 1059.

197. *See id.* at 1059.

198. *See id.*

199. A surrogacy agreement provides for carrying the child nine months and then giving it to the contracting couple. Thus, once both elements are done, the time period begins. This reasoning is based on the language from the *A.Z.* court, which stated there was a "waiting period during which the mother could change her mind." A.Z. v. B.Z., 725 N.E.2d 1051, 1059 (Mass. 2000).

B. The Superior Court's Argument of "Forced Parenthood"

The court in *J.B. v. M.B.* came to the same conclusion, but by a different analysis. Instead of focusing on a reasonable period of time to reconsider one's decision, the *J.B.* court underwent an analysis of one's right to procreate verses one's right not to procreate.[200] It determined that because the husband could still father children without the use of the frozen preembryos his right to procreate was not impaired; whereas the wife's right to not procreate would be denied if the husband was allowed to use the frozen preembryos.[201]

The court also discussed the issue of termination of parental rights.[202] The court, however, stated that it did not decide the case based on constitutional grounds but rather on constitutional principles.[203] The court continued with an analysis of the Baby M. case and concluded that it would not compel her to become a parent against her will to then have to terminate her parental rights.[204]

C. The Washington Court of Appeal's Argument of "Forced Parenthood"

The facts in *Litowitz* involved a different fact pattern from *A.Z.* and *J.B.* The Washington Court of Appeals, however, used *A.Z.* and *J.B.* to conclude that it was against public policy to compel a donor to become a parent against his or her will.[205]

One major factual difference in *Litowitz* was that the wife did not provide any genetic material.[206] The court reconciled this difference by referring to language used in *Davis*, which stated that "decisional authority rests in the gamete-providers alone,"[207] thus concluding that Mrs. Litowitz did not have a right to procreate because she provided no gametes to create the preembryos.[208] This decision settled one major aspect of the 'to procreate verses not to procreate analysis - in situations were only one person provides the gametes necessary to create a preembryo, that person's decision over the disposition of the preembryos supercedes.

However, will this serve as precedence or will this ruling be ignored? Also, what if neither of the two contribute gametes? Suppose the eggs and

200. *See* J.B. v. M.B., 751 A.2d 613, 618-20 (N.J. Super. A.D. 2000).
201. *See id.* at 618-19.
202. *See id.* at 620.
203. *See id.* at 620.
204. *See id.* at 620.
205. *See* Litowitz v. Litowitz, 10 P. 3d 1086, 1090 (Wash. App. Div. 2000).
206. *See supra* notes 172-177 and accompanying text.
207. Davis v. Davis, 842 S.W.2d 588, 602 (Tenn. 1992).
208. *See Litowitz*, 10 P.3d at 1092.

sperm were both donated and used to create a preembryo that a couple 'adopts,' but before the preembryos can be implanted the couple separates and divorces. What should occur when one of the individuals wants to use the preembryos, the other does not?

Litowitz also addressed the issues of express terms in the contract and the intent of the parties.[209] The court compared the Litowitz's agreement to the one in *Kass* to illustrate the differences between the two agreements and to highlight the fact that the Litowitz's never had an express agreement that could be enforced.[210] This lack of an express agreement was also linked to the intent, or lack of intent, of the parties also discussed in Kass. The court stated that, "[a]bsent any evidence that David 'intended to pursue reproduction outside the . . . marital relationship . . . '" it refused to create the obligation.[211] This refusal, however, can only exist if there was no express agreement regarding the disposition of the frozen preembryos, because if there was an express agreement, then the court would have the evidence it needed.

Compiling the holdings of the three courts and looking at the analysis used, it seems that future courts will weigh one party's right to procreate against the other's right not to procreate and disregard any prior disposition agreement. This balance, however, is dependant on whether the use of the frozen preembryos is the person's last chance to conceive, and, whether the individuals both provided the gametes or whether only one did, in which case the person who did provide the gametes would have the final say as to the disposition of the frozen preembryos. However, what about the agreements signed by the parties? Should not contract law apply?

VII. CONTRACT LAW

The principle purpose of contract law is to provide rules for commercial communication as well as non-commercial communication.[212] The right to contract is grounded on the premise that it is in the public's best interest to control and order their own affairs through legally enforceable agreements.[213] A general definition of a contract is, "a promise, or set of promises, for the breach of which the law gives a remedy, or the performance of which the law in someway recognises [sic] as a duty."[214] A contract

209. *See id.* at 1091.

210. *See id.*

211. *Id.*

212. *See* HOWARD O. HUNTER, MODERN LAW OF CONTRACTS: FORMATION, PERFORMANCE, AND RELATIONSHIPS OF PARTIES 22-2 (1986).

213. *See* E. ALLEN FARNSWORTH, CONTRACTS 321 (3rd ed. 1998).

214. P.S. ATIYAH, AN INTRODUCTION TO THE LAW OF CONTRACTS 37 (5th ed. 1995).

usually involves an offer and an acceptance or assent to the offer.[215] Identifying the two different stages of "offer" and "acceptance," however, is nearly impossible in certain situations[216] because contracts can be formed under relatively informal circumstances.

The three basic elements necessary to form an agreement or contract are (1) an offer (2) an acceptance, and (3) consideration.[217] An offer is defined as "a promise by the offeror to do or abstain from doing something."[218] The offer does not have to consist of the express words "I offer," and it may be implied from the offeror's language or conduct.[219] To determine if an actual offer was made,[220] the court determines whether the offeror intended no further negotiations, just an acceptance of the terms.[221] If the offeror only intended an acceptance, then an actual offer was made.[222] An acceptance to the offer is "an assent to the terms of the offer."[223] This stage of forming a contract can also occur under express words or implied contract.[224] An acceptance to an offer[225] is an assent to the exact terms of the offer, however, courts may disregard trivial variations of the terms.[226]

In addition to requiring an offer and an acceptance in order to form a contract, consideration must be exchanged.[227] Consideration exchanged between the parties is more than a promise, it is something of value such as an item, services, or money.[228] This exchange, however, need not be of

---

215. *See* JOHN D. CALAMARI & JOSEPH M. PERILLO, THE LAW OF CONTRACTS 25 (4th ed. 1998).

216. *See* CALAMARI & PERILLO, *supra* note 215, at 25. *See also*, G.H. TREITEL, AN OUTLINE OF THE LAW OF CONTRACT 21 (4th ed. 1989).

217. In addition to the three basic elements mentioned above, the parties must also have "capacity" to contract. *See* HUNTER, *supra* note 212, at 14-1. Capacity necessitates that the parties be of proper age and of sound mind. *See id.* at 14-2. Most states have set the age requirement for entering into a contract at eighteen. *See id.* at 14-2, 14-3. In determining the mental capacity, courts can consider "evidence concerning a person's mental abilities and manifestations of aberrant behavior over a period of time." *Id.* at 14-24.

218. ATIYAH, supra note 214, at 56.

219. *See id.*

220. As opposed to an "invitation to treat," which requires further negotiations. *See* TREITEL, *supra* note 216, at 8; *see also* ATIYAH, *supra* note 214, at 57.

221. *See* ATIYAH, *supra* note 214, at 57.

222. *See id.*

223. TREITEL *supra* note 216, at 10.

224. *See id.*

225. As opposed to a counter-offer, which is a rejection of the original offer. *See* ATIYAH, *supra* note 214, at 66.

226. *See* TREITEL, *supra* note 216, at 10.

227. *See* HUNTER, *supra* note 212, at 13-8.

228. *See id.*

equivalent value to make the contract valid and enforceable.[229]

There are many different types of contracts, but for the purposes of this Note express contracts and implied in fact contracts[230] will only be discussed. An express contract is easily distinguishable from an implied in fact contract. It is defined as an agreement formed orally or by a writing.[231] The terms of the agreement and the intent of the parties are usually explicit and unequivocally expressed.[232] An implied contract, on the other hand, is established by the conduct of the parties.[233] The actions of the person imply that that person has agreed with the terms.

A. Defenses to Contract

A defense to a contract implies that the contract is valid, but should not be enforced for certain reasons. There are many defenses to enforcing a contract, including misrepresentation or fraud. An agreement that consists of misleading statements, which may have induced the other party to enter into the agreement contract is said to be a misrepresentation.[234]

A contract can also be held unenforceable because of policy reasons such as changed circumstances or unconscionability. Unconscionability has been defined as "a provision that 'no fair-minded person would impose on another and that no competent person would freely and knowingly submit to.'"[235] The court, however, usually raises policy reasons, because it can take it upon itself to not enforce the agreement if it deems the agreement is against the public good.[236]

B. The Effects of Public Policy on Contract Law

"A contract is considered unenforceable if its performance would be contrary to public policy."[237] Public policy looks at whether the agree-

229. *See id.* at 13-9 n.34.

230. There are three types of implied contracts; contracts with (1) terms implied in fact; (2) terms implied in law; and (3) terms implied by custom or usage. *See* TREITEL, *supra* note 216, at 63. However, since implied in fact contracts are the only applicable forms for this Note, this is the only one that will be discussed.

231. *See* CALAMARI & PERILLO, *supra* note 215, at 21

232. *See* Mario J. Trespalacios, *Frozen Embryos: Toward an Equitable Solution*, 46 U. MIAMI L. REV. 803, 827 (1992).

233. *See* ATIYAH, *supra* note 214, at 44; *See also* Paula Walter, *His, Hers, and Theirs - Custody, Control and Contracts Allocating Decisional Authority Over Frozen Embryos*, 29 SETON HALL L. REV. 937, 963 (1999).

234. *See* TREITEL, supra note 216, at 122.

235. Haut, *supra* note 23, at 520 (quoting the definition used by other authors).

236. *See* CALAMARI & PERILLO, *supra* note 215, at 819.

237. Barbara Cohen, *Surrogate Mothers: Whose Baby Is It?*, 10 AM. J.L. & MED. 243, 253 (1984).

ment may be injurious to the public.[238] The injury resulting from the performance of the agreement need not be actual to be considered contrary to public policy.[239] If the contract is found void as a matter of public policy, the court leaves the parties as it finds them.[240] This means that neither party will be able to enforce the contract.

An agreement or contract that is considered contrary to public policy creates a situation the law regards as harmful.[241] Policies that are regarded as contrary to public policy can be based on economic notions (such as policies against the restraint of trade),[242] the desire to protect government institutions,[243] or moral values.[244] These policies, however, change over time as society's values change,[245] and can leave many unsuspecting individuals uninformed about agreements they believe are enforceable, but are ultimately not enforced by courts because of the moral implications.

In situations where enforceability of an agreement is questionable because of the public policy implications, the court's decision "must rest on a delicate balancing of factors for and against enforcement".[246] Factors weighed in favor of enforcing the agreement include, "the public interest in protecting the justified expectations of the parties . . . any forfeiture that will result by loss of the reliance interest . . . and a party's excusable ignorance of the contravention of public policy . . .."[247] The main factor used to determine whether not to enforce the agreement, even if public policy considerations exist, depends on the strength of the policy and the furtherance of that policy by not enforcing the agreement.[248]

C. Familial Agreements

Courts generally avoid interfering with family agreements. However, since the dynamics of "a family" and/or family relations are constantly changing, familial agreements have increasingly become issues in the courtroom and are now commonly governed by public policy.

238. *See id.*
239. *See id.*
240. *See* KAPLAN & TONG, *supra* note 6, at 319.
241. *See* TREITEL, *supra* note 216, at 156.
242. *See* FARNSWORTH, *supra* note 213, at 326-27.
243. *See id.* at 327
244. *See id.* at 326
245. *See* ATIYAH, *supra* note214, at 345.
246. FARNSWORTH, *supra* note 213, at 324.
247. *See id.*
248. *See id.* (emphasis added).

1. Antenuptial Agreements[249] - "Preserving the Integrity of Marriage"[250]

A husband and wife can enter into a contract with each other.[251] The common contracts a husband and wife enter into with one another are antenuptial agreements and post-nuptial agreements.[252] Antenuptial agreements are made between prospective spouses in an attempt to resolve issues concerning property rights or support.[253] Post-nuptial agreements serve generally the same purpose but are entered into after marriage.[254] The consideration in an antenuptial agreement is the marriage itself.[255] Since the enforcement of antenuptial agreements are more frequently before the court, the focus of this section will be on these agreements.

Antenuptial agreements have historically been shunned and not enforced by courts because they were considered agreements in anticipation of divorce, and thus encouraging divorce.[256] However, in the 1970's courts began to reconsider their initial positions and started enforcing these agreements,[257] because they can simplify the separation, which can be a stressful and emotional time.[258] Courts still, however, heavily scrutinize antenuptial agreements by applying procedural requirements, which analyze the circumstances surrounding the execution of the agreement, and look at the substantive fairness of the agreement, which examines the actual terms in the agreement.[259]

---

249. These agreements are also called premarital agreements, but for this Note I will refer to these agreements as antenuptial agreements.

250. *See* Katharine B. Silbaugh, *Marriage Contracts and the Family Economy*, 93 NW. U. L. REV. 65, 81 (1998).

251. *See* STEWART MACAULAY ET AL. CONTRACTS: LAW IN ACTION 202 (1995). Contracts between a husband and wife usually arise when one of the individuals is wealthy and it is their second or third marriage. *See id.* Cohabitants can also enter into enforceable contracts as long as the consideration for the contract is not sex. "Courts refuse to enforce agreements where the consideration is illicit sex." HUNTER, *supra* note 212, at 22-31.

252. *See generally* HUNTER, *supra* note 212, at 22-31, 22-34.

253. See SAMUEL GREEN & JOHN V. LONG, MARRIAGE & FAMILY LAW AGREEMENTS 112 (1984).

254. *See id.*

255. *See id.* at 119.

256. *See* HUNTER, *supra* note 212, at 22-31.

257. *See* Silbaugh, *supra* note 250, at 73. "[C]ourts [began to take] the position that a premarital agreement gave marriage stability because it forced a couple to review financial issues before marriage...." *Id.*

258. *See* HUNTER, *supra* note 212, at 22-31. Even the Uniform Marriage and Divorce Act encourages separation agreements. *See* HUNTER, *supra* note 212, at 22-35.

259. *See* Silbaugh, *supra* note 250, at 72-75. The procedural requirements courts look at to determine if the antenuptial agreement is enforceable include at least one of the following requirements: (1) separate counsel; (2) full disclosure of assets; (3)

Antenuptial agreements that are used to distribute property are most favored by courts and usually enforced.[260] However, if these agreements contain non-monetary terms or provisions, the court will generally not enforce it, reasoning that the agreement is contrary to public policy because they dictate the conduct within the marriage.[261] Non-monetary terms may address issues of "care, custody or religious upbringing of children," or they may address more personal and intimate interactions within a marriage such as the allocation of work responsibilities within the marriage or the frequency of sexual relations.[262]

One of the most common reasons courts give for refusing to enforce these 'conduct' provisions in antenuptial agreement is that they violate public policy[263] by debasing the integrity of marriage and/or "altering the incidents of marriage."[264] The first reason is centered on the belief that such provisions would commercialize the institution of marriage.[265] The second public policy argument, which is against altering the incidents of marriage, is grounded in the Restatement of Contracts Section 190 and its Comments, which states that "essential incidents" of marriage that are altered by a contract will not be enforced.[266]

By refusing to enforce these non-monetary terms of an antenuptial agreement, a court seems to be playing a paternalistic role. It is trying to control the relationship between a husband and wife and the issues a married couple faces. What societal interest is the court furthering by entering into these intimate decisions? Is not the interest involved an individual interest?

2. Surrogacy - the "Womb for Hire"[267]

Surrogacy agreements are contractual in nature.[268] The surrogate

disclosure of the rights being waived; (4) sufficient time before the wedding to review the agreement. *See id.*

260. *See* GREEN & LONG, *supra* note 253, at 114.

261. *See* Silbaugh, *supra* note 250, at 78-81.

262. *See id.* at 71.

263. *See id.* at 79. The other common reason the courts give for not enforcing these provisions is lack of consideration. *See id.* 79-81.

264. *See id.* at 81. Incidents of marriage can involve: "finances, living arrangements, education, birth control, children, and any other topic which may be relevant to the individuals involved." GREEN & LONG, *supra* note 253, at 114.

265. *See* Silbaugh, *supra* note 242, at 82.

266. *See id.* at 83.

267. KAPLAN & TONG, *supra* note 6, at 314.

268. *See* GREEN & LONG, *supra* note 253, at 246. There are various reasons why a couple would attempt to conceive a child by these means. Many times the woman is sterile, or her age may make pregnancy risky for her or the child. The couple may also want a child who is biologically connected to the couple; something adoption

mother is artificially inseminated with the husband's sperm of the contracting couple (the sperm donor and his wife), and upon the infant's birth, the surrogate mother gives up all parental rights to the contracting couple.[269] The couple then establishes their parental rights which usually involves the wife of the sperm donor legally adopting the child.[270] The consideration in a surrogacy agreement is a fee that is given to the surrogate mother for compensation.[271] The problem with surrogacy contracts arise when monetary compensation is not given to the surrogate until the child is legally adopted, which then appears to be a purchase price.[272]

One of the most well known surrogacy cases is the New Jersey Supreme Court decision of *In the Matter of Baby M.*[273] In brief, Mr. Stern and Mrs. Whitehead entered into a surrogacy contract where Mrs. Whitehead would be artificially inseminated using Mr. Stern's sperm.[274] Mrs. Whitehead was to carry the child to term, deliver the child to the Sterns, and relinquish any and all parental rights.[275] Mrs. Whitehead was to receive $10,000 dollars after the child's birth and delivery to Mr. Stern.[276] Mrs. Whitehead initially adhered to the contract, however, after a few days she wanted the child permanently returned to her.[277] In its decision, the New Jersey Supreme Court declared the surrogacy contract unenforceable as against public policy.[278] The two general reasons given were that a surrogacy contract may exploit financially needy woman, and second, that these contracts are a form of baby selling.[279]

At least eight states agree with the decision of the New Jersey Supreme Court and hold surrogacy contracts unenforceable as against public policy.[280] Florida, Michigan, Nevada, New Hampshire, New York, Virginia, Washington, and West Virginia all prohibit payments to surrogate mothers

---

cannot offer. *See id.* at 246-47.

269. *See* GREEN & LONG, *supra* note 253, at 246.

270. *See* Tara Rayne Shewchuk, Note, *Regulation of Pre-Conception Agreements: A Synthesis of Individual Rights and Community Values*, 1 HEALTH L.J. 147, 155-56 (1993).

271. *See* GREEN & LONG, *supra* note 253, at 255. The monetary compensation is usually for medical and travel expenses as well as pharmaceutical and therapeutic expenses. *See id.* at 256.

272. *See id.*

273. 537 A.2d 1227 (1988) [hereinafter "*Baby M.*"].

274. *See id.* at 1235.

275. *See id.*

276. *See id.*

277. *See id.* at 1236-37.

278. *Id.* 1246-51.

279. *See* KAPLAN & TONG, *supra* note 6, at 318 (discussing the *Baby M.* case).

280. *See id.* at 319.

beyond compensation for their expenses.[281] The underlying public policy of the statutes is that the statutes will prevent the sale of infants.[282]

Consequently, since there are few statutes completely prohibiting surrogacy, the courts rely on the public policy implications to support their refusal to enforce surrogacy agreements. These policies, the exploitation of poor women and the creation of a baby selling market, are held to be contrary to public policy because of the philosophy that "human beings should not be subject to barter."[283]

D. Applying Contract Law to the Twenty-First Century Cases

Courts seem to apply contract law to familial agreements whenever possible. As seen earlier, in the cases of antenuptial agreements and surrogacy contracts, courts initially applied contract law.[284] It therefore seems only logical to begin the analysis of disposition agreements between gamete providers at the same juncture. Using the contract law framework to analyze the enforceability of the agreements in *A.Z.*, *J.B.*, and *Litowitz*, the court should focus mainly on the rights of the parties involved.[285] Since it has already been established that individuals, including married couples, have the right to enter into a contract,[286] enforcing these contracts would illustrate the right each individual possesses. Applying contract law to determine the enforceability of IVF disposition agreements would also minimize misunderstandings between the parties because an actual consensus was reached.[287]

The two possible contract theories that are applicable when examining disposition agreements under contract law are: (1) that an express contract

281. *See* Lori B Andrews & Nanette Elster, *Regulating Reproductive Technologies*, 21 J. LEGAL MED. 35, 42 (2000). *See* FLA. STAT. ANN. §§ 63.212(1)(i), 742.15(4) (West 1997); MICH. COMP. LAWS ANN. §§ 722.853(a), 722.859 (West 1997); NEV. REV. STAT. ANN. § 126.045(3) (Michie1995); N.H. REV. STAT. ANN. § 168-B25(V) (1996); N.Y. DOM. REL. LAW § 123(1) (McKinney 1997); VA. CODE ANN. §§ 20-156, 20-160(B)(4)-(5) (Michie 1997); WASH. REV. CODE ANN. §§ 26.26.210(1), 26.26.230 (West 1997); W. VA. CODE § 48-4-16 (1997). *See also supra* note 266. Two jurisdictions ban surrogacy contracts completely: the District of Columbia and Arizona. *See* Andrews and Elster, *supra* note 276, at 42; D.C. CODE ANN. § 16-402 (1996); ARIZ. REV. STAT. ANN. § 25-218 (1996).

282. *See* Barbara Cohen, *supra* note 237, at 252.

283. Steven M. Recht, Note, *"M" is for Money: Baby M and the Surrogate Motherhood Controversy* 37 AM. U. L. REV. 1013, 1026 (1996).

284. *See supra* notes 250-289 and accompanying text.

285. *See* Trespalacios, *supra* note 232, at 826.

286. *See supra* note 251.

287. This was one of the reasons the *Kass* court chose to enforce the agreement. *See* Kass v. Kass, 663 N.Y.S. 2d 581, 588 (1997).

existed because there was an express provision as to how the preembryos would be disposed of, and therefore the agreement should be enforced; or (2) that an implied contract to complete the IVF procedure existed and should be enforced. If an express contract theory were used it would focus on the words and terms used in the document(s) signed.[288] For example, if the agreement signed stated that the frozen preembryos are to be destroyed upon divorce or separation then that wish would be honored. This express contract theory would also enforce any unique disposition that was agreed upon, even if not within the normal dispositions of donation, destruction, or giving the preembryos to the clinic for research.

Under the implied in fact contract theory, the couple's actions, and the circumstances surrounding the actions, would determine if a contract existed between them.[289] Thus, as long as the parties' physical, financial and emotional commitments were satisfied then an implied contract to complete the IVF procedure would exist.[290] Since consideration is needed in order to create a legally binding contract,[291] the consideration in these agreements would be the gamete donation process that the husband and wife experience. Although the husband's sperm donation is not comparable to the pain the wife must endure during the egg retrieval procedure, the exchange of consideration does not have to be of equal value.[292] The two possible contract theories of course would be subject to the defenses available in contract law.[293]

One detail of importance that must be addressed is that the disposition agreements involved in *A.Z.*, *J.B.* and *Litowitz* are actually disposition agreements with the clinic. These disposition agreements serve the purpose of giving the clinic guidance as to how to dispose of the frozen preembryos stored at their clinics. However, it can also be construed as an agreement between the couple and the clinic. The couple's mutual participation in the IVF process requires them both to decide on the various contingencies, therefore the agreement may extend to be binding on the couple with each other because the "contingencies that arise [are] between the donors themselves."[294]

1. Contract Law and *A.Z. v. B.Z.*

In *A.Z. v. B.Z.* there was an express contract between the clinic and the

288. *See supra* notes 230-233 and accompanying text.
289. *See id.*
290. *See* Trespalacios, *supra* note 232, at 829.
291. *See supra* notes 217-229 and accompanying text.
292. *See supra* note228 and accompanying text.
293. *See supra* notes 227-229 and accompanying text.
294. Trespalacios, *supra* note 232, at 827-28.

couple,[295] which as stated above, may be seen as binding on the couple. The agreement that had to be completed by the couple prior to egg retrieval provided a blank space after the contingency, "should we become separated."[296] This space was filled in by the wife with the statement that "they both agreed to have the embryo(s) . . . returned to the wife for implantation."[297] The husband and wife then signed the form. This process created a valid contract between the husband and the wife in *A.Z. v. B.Z.*; and had the same sequential fact pattern occurred after each egg retrieval there would be no reason to not enforce the contract. This unfortunately was not the case.

The six egg retrievals that followed were not executed in the same manner, and therefore, enforcement of the agreement is highly unlikely under contract law. The husband signed the six forms that followed in blank.[298] Both husband and wife signed only first agreement *after* the form was completed, including the provision referring to the couple's separation.[299] This fact raises the question of whether the husband was aware of the disposition in the six later forms making the terms unenforceable, and whether he waived his right to contest the contract.

If the Supreme Judicial Court had analyzed *A.Z. v. B.Z.* under contract law, the husband could have had various defenses making the agreement unenforceable. The first defense the husband could have raised is the defense of misrepresentation. Misrepresentation refers to a misleading statement purposely made to induce the other party to enter into the agreement.[300] The husband could have raised this defense because he signed all the forms in blank except the first,[301] and since an agreement was executed for each egg retrieval and implantation of the eggs, the validity of the first agreement expired once the second agreement was signed and egg retrieval completed.

The husband could also have raised the defense of unconscionability.[302] This defense can set aside any unreasonable provision because the bargain was grossly unfair.[303] Enforcement of the agreement in *A.Z.* would be unconscionable because the husband was not aware of the disposition of the frozen preembryos. He signed the agreement in blank, therefore could not and did not consent to the disposition of the preembryos.

---

295. *See* A.Z. v. B.Z., 725 N.E.2d 1051, 1054 (Mass. 2000).
296. *Id.*
297. *Id.*
298. *See id.*
299. *See supra* notes 127-29 and accompany text.
300. *See* TREITEL, *supra* note 216, at 122.
301. *See* A.Z. v. B.Z., 725 N.E.2d 1051, 1054 (Mass. 2000).
302. *See supra* note 235 and accompanying text
303. *See* ATIYAH *supra* note 214, at 300.

2. Contract Law and *J.B. v. M.B.* and *Litowitz v. Litowitz*

Applying contract law to *J.B. v. M.B.* and *Litowitz* is a little more difficult because there was no explicit provision providing for the disposition of the frozen preembryos in either case,[304] each agreement simply allowed a court to decide the disposition of the preembryos. Thus, in both cases the court was left to determine who would get the frozen preembryos.

In *J.B. v. M.B.*, the couple signed an agreement provided by the clinic which stated that the IVF clinic would gain control of the frozen preembryos if there was a dissolution of the marriage, unless a court specified which individual (husband or wife) was to gain control of the frozen preembryos.[305] This was clearly an express agreement. It expressly stated how the situation would be handled: if there was a dissolution of marriage the IVF Program would gain control unless the court gave one of the individuals control of the preembryos.[306] It however, did not provide for an explicit disposition of the preembryos. The provision in the agreement stated that the family court would decide the disposition of the preembryos when the divorce is made final and the divorce decree is entered. The form specifically states: ". . . direction and ownership of our tissues will be relinquished to the IVF Program under the following circumstances: 1. a dissolution of our marriage by court order, unless the court specifies who takes control and direction of the tissue . . .."[307] This is exactly what occurred. The judge made a finding as to who would gain control of the preembryos. Thus, the express provision, that the court would decide the disposition of the frozen preembryos upon the dissolution of the marriage, was enforced.

This case has no valid defense against enforcement of the contract. There was no misrepresentation or fraud because they both signed the agreement with presumptively the same knowledge regarding the agreement. No one inserted clauses without the other's awareness, as in *A.Z. v. B.Z.* There also is no argument of unconscionability because there was nothing so overreaching as to argue that no competent person would have agreed to such a contract.

In *Litowitz*, the facts are different from any of the cases previously discussed, because the eggs did not come from Mrs. Litowitz and there was a surrogate mother involved.[308] However, an express agreement concerning

---

304. In both *J.B. v. M.B.* and *Litowitz v. Litowitz* the disposition agreements stated that the court would decide the disposition of the frozen preembryos if they (the couple) could not mutually agree upon a separation or divorce.

305. J.B. v. M.B., 751 A.2d 613, 616 (N.J. Super. A.D. 2000).

306. *See id.*

307. *Id.*

308. *See* Litowitz v. Litowitz, 10 P. 3d 1086, 1087-88 (Wash. App. Div. 2000).

the disposition of the frozen preembryos did exist.[309] Similar to *J.B. v. M.B.*, the agreement did not specifically provide for the disposition of the preembryos in case of divorce or separation,[310] however, it did state that the couple must petition a court of competent jurisdiction "concerning the appropriate disposition" of the preembryos if they are "unable to reach a mutual decision."[311] Since the couple was unable to mutually decide on the disposition of the frozen preembryos, they petitioned a court to decide who would gain control of the frozen preembryos, and the court did exactly what was required of it. It decided the disposition of the frozen preembryos. Thus, here too, the contract was enforced.

This case also does not provide for any contract law defenses. There was no misrepresentation involved, and there was no unfair surprise.

3. When Should Contract Law Apply?

The best way to handle disputes between gamete providers concerning the disposition of their frozen preembryos is by enforcing the signed IVF disposition agreements under contract law principles when the disposition is clearly ascertainable by an express agreement. By this I mean that if the provision undeniably states that the preembryos should be donated, destroyed, given to the clinic for research, or any other clear disposition, then the provision should be enforced. If there is a defense to the express contract, then the preembryos should just be destroyed.

Contract law, however, should be limited to express agreements with clear dispositions. The theory of implied in fact contracts should not enter the arena of frozen preembryos because no actual words were used to make a decision regarding the disposition of the frozen preembryos. By not verbally expressing the disposition, one or both of the parties involved may not completely understand the ramifications of the disposition. At least when something is brought to the other's attention, verbally or in writing, the other party can raise his or her concerns.

In situations were the express agreement is to allow the court to decide, such as in *J.B.* and *Litowitz*, the court should look at the actions of the parties in order to determine what the couple may have chosen had they been given the opportunity to determine the disposition of their frozen preembryos. Take, for instance, the facts of *J.B. v. M.B.* In this case, the contract stated that the court should decide on the disposition of the preembryos. The facts also stated that there was a possibility that the wife initially had a desire to donate the preembryos to an infertile couple,[312] but

309. *See id.* at 1088.

310. *See id.* at 1089.

311. *Id.*

312. J.B. v. M.B., 751 A.2d 613, 616 (N.J. Super. A.D. 2000). In the husband's

because the agreement stated that the court must decide, it was never an express disposition agreement.

When the court balances what it *believes* is the interests of the parties, the court takes away the gamete providers' decision-making power.[313] The court completely ignores the intent of the parties; this is contrary to the purpose of executing contracts.[314] A contract is meant to provide direction for the parties involved, not to be disregarded and replaced by a decision made by a court.

4. The Effects of Contract Law on the *Davis* and *Kass* Decisions

Davis and *Kass* are the only two cases, prior to the twenty-first century cases, that addressed the enforceability of disposition agreements, and both held that prior agreements concerning the disposition of a couple's frozen preembryos should be presumed valid and enforced. The *Davis* court focused on the fact that these agreements illustrated the wishes of the decision makers,[315] while *Kass* centered on the manifest intent of the parties.[316] These two reasons the courts gave for presuming the validity of the agreements are not affected by the proposal in this Note to use contract law to enforce express disposition agreements.

The reasons *Davis* and *Kass* gave to enforce the disposition agreements depend on the existence of an express disposition agreement. The decision-making authority *Davis* addresses only exists if a decision is made, because an authority can not be enforced if it is not exercised, and the only manner by which to exercise this authority is by actually making a decision, which is creating an express disposition agreement. Similarly a "manifestation of mutual intent"[317] can only exist in an express disposition agreement. Nothing can show the intent of the parties more clearly than an express disposition agreement, which the *Kass* court believed should be given effect.[318]

---

affidavit, he stated that the wife had mentioned that she would like the frozen preembryos donated to an infertile couple if anything should happen, because she knew of many couples who were trying to conceive but could not. *See id.*

313. See John A. Robertson, *Prior Agreements for Disposition of Frozen Embryos*, 51 OHIO ST. L.J. 407, 415 (1990). (emphasis added).

314. *See* FARNSWORTH, *supra* note 213, at 321

315. Davis v. Davis, 842 S.W.2d 588, 602-03 (Tenn. 1992). The court believed that the couple's decision should be respected because their decision affected their lives. The court stated that "because no one else bears the consequences of these decision in the way that the gamete-providers do," their decision should only matter, and they should have the "sole decisional authority." *Id.*

316. Kass v. Kass, 663 N.Y.S. 2d 581, 587 (1997). The court believed it should refrain from interfering with the wishes of the parties. *See id.* at 590.

317. *Kass*, 663 N.Y.S.2d at 587.

318. *See id.*

Consequently, the *Davis* and *Kass* decisions would not be affected by my proposal to use contract law to enforce express disposition agreements. In fact, it provides two useful reasons why courts should enforce express disposition agreement under contract law principles.

VIII. CONCLUSION

Medically assisted reproduction has come a long way. It has changed the lives of many individuals attempting to conceive, as well as the function of the judicial system. What once was considered to be an intimate decision between a man and a woman has entered the courtroom to be analyzed and scrutinized by a court.

Examining the recent cases of *A.Z. v. B.Z.*,[319] *J.B. v. M.B.*[320] and *Litowitz v. Litowitz*[321] has given us a glimpse at one of the most recent dilemmas arising from advances in reproductive technology - the enforceability of signed agreements, upon divorce of the married couple, concerning the disposition of their frozen preembryos. It has also provided some insight as to how the courts have held the way they have, and what the future of assisted reproduction and the law holds.

Since public policy was the focus of the courts, this Note briefly examined the general principles of contract law and how a contract that is considered contrary to public policy is handled[322] in order to understand the effects of 'public policy.' To assist in my understanding of the role of public policy in familial agreements, this Note examined two such agreements before a court, antenuptial agreements and surrogacy contracts. This brief analysis of the familial agreements illustrates that public policy can be arbitrary, because not every court holds the same issues contrary to public policy.

This Note also compares the only two other cases that existed prior to these three 21st Century cases, *Davis v. Davis* and *Kass v. Kass,* to determine if there is any consensus in the judicial system as to how disposition agreements regarding frozen preembryos are to be enforced. The holding in *Davis* and *Kass*, however, seemed to provide no guidance to *A.Z.*, *J.B.* and *Litowitz* because *Davis* and *Kass* both hold that the disposition agreements should be presumed valid and enforced and *A.Z.*, *J.B.* and Litowitz held otherwise.

Applying contract law may assist courts in deciding whether to enforce the express disposition agreements. Contract law simplifies the court's analysis and limits its interference into familial agreements, and well as

319. 725 N.E.2d 1051 (Mass. 2000).
320. 751 A.2d 613 (N.J. Super. A.D. 2000).
321. 10 P.3d 1086 (Wash. 2000).
322. *See supra* Part IV.

provides defenses for unfair or unconscionable dispositions of the frozen preembryos.

Contract law is the best method by which to enforce express disposition agreements. If the agreement contains an express disposition that the two parties agreed upon and signed, and no contract defenses apply, thus there is no reason not to enforce the agreement.

If, however, there is no express provision regarding the exact disposition of the frozen preembryos, then the frozen preembryos should be destroyed. Since the parties did not consider how the frozen preembryos should be disposed of at the start of IVF treatment, the court should not have to have that responsibility. As the court does not know the background into the lives of the parties, they are not equipped to determine the most equitable solution.

Public policy will always exist as a method by which not to enforce a contract. However, courts must consider the fact that individuals have a right to contract. Simply because the court itself does not agree with the disposition agreement previously entered into by the couple regarding their frozen preembryos does not allow it to hold them contrary to public policy. Courts must realize that technological advances in human reproduction are moving at a rate that the judicial system has failed to maintain.

Marysol Rosado