Denver Law Review

Volume 84 | Issue 2 Article 11

January 2006

Vol. 84, no. 2: Full Issue

Denver University Law Review

Vol. 84, no. 2: Full Issue

Published by the
University of Denver
Sturm College of Law

2006 Volume 84 Issue 2

DENVER UNIVERSITY LAW REVIEW

CONTENTS

ARTICLES

COMMENTS

PARENTS INVOLVED & MEREDITH: A PREDICTION REGARDING THE (UN)CONSTITUTIONALITY OF RACE-CONSCIOUS STUDENT ASSIGNMENT PLANS

EBONI S. NELSON†

ABSTRACT

During the October 2006 Term, the United States Supreme Court will consider the constitutionality of voluntary race-conscious student assignment plans as employed in Parents Involved in Community Schools v. Seattle School District No. 1 *and* Meredith v. Jefferson County Board of Education. *These cases will mark the Court's first inquiry regarding the use of race to combat de facto segregation in public education. This article examines the constitutionality of such plans and provides a prediction regarding the Court's decisions.*

This article begins with an analysis of the resegregation trend currently plaguing American educational institutions and identifies two causes for the occurrence: (1) the shift in the Supreme Court's jurisprudence regarding desegregation; and (2) school officials' adherence to the "neighborhood school concept" when making student assignment decisions. This article then examines the challenged plans, specifically their attempts to create and maintain racially diverse student bodies through the use of racial tiebreakers and guidelines. After considering the Supreme Court's prior decisions and rationale regarding the use of race in education, this article predicts that the Supreme Court will strike down both plans as violative of the Equal Protection Clause. In light of this probable outcome, this article urges school officials to consider race-neutral methods to achieve diversity and to improve the quality of education provided to disadvantaged, minority students.

TABLE OF CONTENTS

† Assistant Professor of Law, Thurgood Marshall School of Law; enelson@tsulaw.edu; J.D., Harvard Law School, 2001; B.A., Wake Forest University, 1998. My thanks to L. Darnell Weeden and the faculties of Thurgood Marshall School of Law, University of South Carolina School of Law, and Valparaiso University School of Law for comments, suggestions, and discussion on this Article. Also, I would like to thank Victoria Walton for her invaluable research assistance. Most importantly, I thank Scott and Ella Nelson for their love and support.

INTRODUCTION

In 2003, after a twenty-five year hiatus,[1] the United States Supreme Court reentered the passionate and controversial debate surrounding affirmative action in the context of public education. The Court's dual decisions in *Grutter v. Bollinger*[2] and *Gratz v. Bollinger*[3] sanctioned the limited use of race as a factor in higher education admissions decisions. During the October 2006 Term, the Court will revisit the issue of affirmative action, only this time the inquiry will concern the use of race in elementary and secondary education rather than higher education.

In a somewhat surprising announcement, the Court decided to hear the appeals of two cases challenging school districts' use of race in student assignment decisions.[4] Six months prior to the Court's decision to hear arguments in *Parents Involved in Community Schools v. Seattle School District No. 1*[5] and *Meredith v. Jefferson County Board of Education*,[6] the Court declined to grant certiorari in a similar case,[7] thereby

1. *See* Regents of the Univ. of Cal. v. Bakke, 438 U.S. 265 (1978) (analyzing the constitutionality of race-conscious admissions policies in higher education).
2. 539 U.S. 306 (2003).
3. 539 U.S. 244 (2003).
4. *See* Charles Lane, *Justices to Hear Cases of Race-Conscious School Placements*, WASH. POST, June 6, 2006, at A03, *available at* http://www.washingtonpost.com/wp-dyn/content/article/2006/06/05/AR2006060500367.html.
5. *(Parents II)*, 426 F.3d 1162 (9th Cir. 2005), *cert. granted*, 126 S. Ct. 2351 (U.S. June 5, 2006) (No. 05-908).
6. McFarland v. Jefferson County Pub. Sch., 330 F. Supp. 2d 834, 861 (W.D. Ky. 2004), *aff'd*, 416 F.3d 513 (6th Cir. 2005), *cert. granted sub nom.*, Meredith v. Jefferson County Bd. of Educ., 126 S. Ct. 2351 (U.S. June 5, 2006) (No. 05-915). The Supreme Court has set both cases for argument on December 4, 2006. *See* Argument Calendars (October Term 2006) Session Beginning November 27, 2006, *available at* http://www.supremecourtus.gov/oral_arguments/argument_calendars/MonthlyArgumentCalDecember2006.pdf.

prompting speculation as to the reasons for the Court's apparent about-face. One could attribute the Court's decision to its desire to reconcile circuit court splits regarding the constitutionality of race-conscious student assignment plans pre- and post-*Grutter*.[8] While this reason may be plausible, it would not appear to be the primary reason given that such splits existed prior to the Court's certiorari denial in *Comfort ex rel. Neumyer v. Lynn School Committee*.[9] Others have hypothesized that the Court's decision to grant certiorari was precipitated by the change in its composition—a change that some think may prove to be the death knell of desegregation.[10]

The composition of the Court that declined to hear *Lynn* included Justice Sandra Day O'Connor, who wrote the 5 to 4 *Grutter* opinion upholding the use of race in higher education. Often thought of as the "swing vote" in controversial and pivotal cases,[11] Justice O'Connor retired from the Court in 2006.[12] Following the appointment of her replacement, Justice Samuel A. Alito, Jr., who is commonly thought to be a conservative Justice,[13] the newly constituted Court agreed to hear *Par-*

---

7. *See* Comfort *ex rel.* Neumyer v. Lynn Sch. Comm., 283 F. Supp. 2d 328 (D. Mass. 2003), *aff'd*, 418 F.3d 1 (1st Cir. 2005), *cert. denied*, 126 S. Ct. 798 (2005) (upholding the use of race in elementary and secondary education student transfer policies).

8. *See* Petition for a Writ of Certiorari at 8-15, *Parents II*, 126 S. Ct. 2351 (No. 05-908), 2006 WL 1579631 (detailing circuits' conflicting holdings regarding the constitutionality of race-conscious assignment plans and urging the Supreme Court to grant certiorari "to remove this uncertainty and confusion" regarding "how *Grutter* and *Gratz* affect the Equal Protection rights of students in public high schools"); Petition for a Writ of Certiorari at 8-11, *Meredith*, 126 S. Ct. 2351 (No. 05-915), 2006 WL 165912 (arguing that the Court should grant certiorari because "[t]he decision of the Sixth Circuit directly conflicts with decisions of the Fourth, Fifth and Ninth Circuits concerning voluntarily-adopted race-based student assignment plans designed to advance racial diversity"); *see also* Lane, *supra* note 4.

9. 126 S. Ct. 798 (2005); *see* Petition for a Writ of Certiorari at 7-10, *Lynn*, 126 S. Ct. 798 (No. 05-348), 2005 WL 2275949 (noting conflicts between the First, Second, Fourth, Fifth, Sixth and Ninth Circuits regarding the constitutionality of public schools voluntarily adopting race-conscious student assignment plans to achieve racial diversity).

10. *See* Linda Greenhouse, *Court to Weigh Race as Factor in School Rolls*, N.Y. TIMES, June 6, 2006, at A1, *available at* http://www.nytimes.com/2006/06/06/washington/06scotus.html?ex=1307246400&en=7b7b1af6cbef8911&ei=5088&partner=rssnyt&emc=rss (suggesting that the change in Supreme Court Justices prompted the Court to grant certiorari); Lane, *supra* note 4 (quoting Professor Goodwin Liu's thoughts of the Court's granting of certiorari as "bad news for desegregation advocates . . . . It looks like the more conservative [J]ustices see they have a fifth vote to reverse these cases").

11. *See* Tom Curry, *O'Connor Had Immense Power as Swing Vote*, July 1, 2005, http://www.msnbc.msn.com/id/5304484/ (describing Justice O'Connor as "often the swing vote that decided high-profile cases"); *see also High Court at a Crossroads*, http://www.msnbc.msn.com/id/9531661/ (last visited Oct. 25, 2006) (detailing six significant Supreme Court decisions, ranging from partial birth abortion to state sovereign immunity, in which Justice O'Connor provided the fifth deciding vote).

12. Although Justice O'Connor announced her resignation on July 1, 2005, it was not effective until the confirmation of her successor, which occurred on January 31, 2006, with the swearing in of Justice Alito. *See* William Branigin et al., *Supreme Court Justice O'Connor Resigns*, WASH. POST, July 1, 2005, *available at* http://www.washingtonpost.com/wp-dyn/content/article/2005/07/01/AR2005070100653.html; *Alito Sworn In as Supreme Court Justice*, Jan. 31, 2006, http://www.msnbc.msn.com/id/11111624/ [hereinafter *Alito Sworn In*].

13. *See Alito Sworn In*, *supra* note 12 (describing Justice Alito as a "conservative lawyer for the Reagan administration"); Don Gonyea, *All Things Considered: Republicans Praise Alito's Con-*

ents Involved and *Meredith*, which will be the first time the Court has addressed the constitutionality of the voluntary use of race in elementary and secondary school student assignment plans.[14] While no one can know how any of the Justices will vote in the cases, many affirmative action opponents hope that the additions of Justice Alito and Chief Justice John G. Roberts, Jr.[15] to the Court will result in the prohibition of race-conscious assignment programs in public elementary and secondary schools.[16] Supporters of affirmative action fear that such a ruling will prompt and exacerbate resegregation trends currently plaguing public education.[17] Whether the Court upholds or strikes down the assignment plans employed in the two cases, *Parents Involved* and *Meredith* will significantly contribute to the jurisprudence concerning equality in public education.

Many agree that public elementary and secondary schools are more segregated today than they were prior to the *Brown v. Board of Educa-*

---

servative Credentials (NPR radio broadcast Oct. 31, 2005), *available at* http://www.npr.org/templates/story/story.php?storyId=4983450 (describing Justice Alito as "a judicial favorite of the conservative movement"); Bill Mears, *Alito's Record Shows Conservative Judge*, CNN.COM, Oct. 31, 2005, http://www.cnn.com/2005/POLITICS/10/31/alito.record/index.html (discussing Justice Alito's "conservative judicial philosophy" and relating views that he was "the most conservative member of the [Third Circuit Court of Appeals]").

14. Throughout the article, reference to the "voluntary" use of race in student assignment plans refers to school districts' consideration of students' race absent the operation of a federally mandated desegregation decree.

15. Following the death of Chief Justice William Rehnquist, Chief Justice Roberts was confirmed on September 29, 2005. *See* Charles Babington & Peter Baker, *Roberts Confirmed as 17th Chief Justice*, WASH. POST, Sept. 30, 2005, at A01, *available at* http://www.washingtonpost.com/wp-dyn/content/article/2005/09/29/AR2005092900859.html. Although void of a lengthy judicial history, Chief Justice Roberts' commentary regarding race-based affirmative action prior to taking the bench has led some to believe that he generally opposes race-based affirmative action.

> As Acting Solicitor General, Roberts' approval of a brief opposing the Federal Communications Commission's affirmative action program for broadcast licensees and later, as a private attorney, his brief on behalf of the Associated General Contractors of America in opposition to the government's highway construction program in *Adarand Constructors v. Pena* clearly indicate that had Roberts sat in the place of Justice Sandra Day O'Connor, equal access to higher education (*Grutter v. Bollinger*) and contracting (*Adarand v. Pena*) would have been foreclosed to minorities.

Press Release, American Association for Affirmative Action Opposes Confirmation of John Roberts for Chief Justice of U.S. Supreme Court (Sept. 5, 2005) (quoting Robert Ethridge, President of the American Association for Affirmative Action), *available at* http://www.affirmativeaction.org/press.jsp.

16. *See* Lane, *supra* note 4 ("Sharon Browne, principal attorney of the Pacific Legal Foundation, which supports the parents' lawsuits [in *Parents Involved* and *Meredith*], said she 'was pleased that the Court has decided to hear these cases. Together, these cases could put an end to schools using race as a factor to decide where children can attend school.'"); Greenhouse, *supra* note 10 (quoting Sharon Browne as saying, "I think the writing's on the wall, or at least I hope it is.").

17. *See* Gina Holland, *Supreme Court to Hear Schools Race Case*, CBS NEWS, June 5, 2006, http://www.cbsnews.com/stories/2006/06/05/ap/politics/mainD8I2AB700.shtml ("A ruling against the schools 'would be pretty devastating to suburban communities, small towns that have successfully maintained desegregation for a couple of generations,' he said. 'The same communities that were forced to desegregate would be forced to re-segregate.'") (quoting Gary Orfield, Director of the Civil Rights Project at Harvard University); Bob Egelko & Heather Knight, *Justices Take Cases on Race - Based Enrollment, But Prop. 209 Means California Schools Likely to be Unaffected*, S.F. CHRON., June 6, 2006, at B1 (noting views that the consideration of race in public elementary and secondary schools is necessary to "reverse growing resegregation of the schools").

tion[18] decision.[19] Current resegregation trends threaten thirty years of progress that have been made in the desegregation of African-American students,[20] thereby impeding the fulfillment of *Brown*'s promise of educational equality. Realizing the potentially devastating effects of segregated schools,[21] several school districts have voluntarily begun to employ race-conscious student assignment plans, such as those challenged in *Parents Involved* and *Meredith*, to prevent and remedy resegregation of their schools. This article examines the constitutionality of such plans and hypothesizes that the Supreme Court will strike down both student assignment plans employed in *Parents Involved* and *Meredith* as unconstitutional.

Part I begins with an analysis of factors contributing to resegregation in elementary and secondary schools. Just as the Supreme Court has been an invaluable tool by which to desegregate public schools, some of its decisions have also enabled resegregation to flourish. Part I also discusses the negative impact that school districts' adherence to the "neighborhood school concept" has had on the provision of equal educational opportunities to minority students.

Part II examines the district court and Ninth Circuit opinions in *Meredith* and *Parents Involved.* It discusses the compelling interests asserted by the school districts to justify their narrowly tailored use of race in student assignment decisions.

Part III analyzes the constitutionality of voluntary race-conscious student assignment plans as employed in *Parents Involved* and *Meredith.* Although difficult to predict, this article hypothesizes that the Court will invalidate both student assignment plans as violative of the Equal Protection Clause. This hypothesis is predicated on the Court's previous decisions and rationale concerning the use of race in the context of public education.

This article concludes with suggestions regarding policies and programs that school districts can utilize in their attempts to combat the se-

---

18. (*Brown II*), 349 U.S. 294 (1955).

19. *See* Hon. Robert L. Carter, *The Conception of* Brown, 32 FORDHAM URB. L.J. 93, 99 (2004) (concluding that "[t]here are more segregated secondary and primary schools today than existed before *Brown*"); Marvin Krislov, *Affirmative Action in Higher Education: The Value, the Method, and the Future*, 72 U. CIN. L. REV. 899, 906 (2004) (concluding that many parts of the country are experiencing segregation at levels greater than those existing when *Brown* was decided); Charles J. Ogletree, Jr., *All Deliberate Speed?:* Brown*'s Past and* Brown*'s Future*, 107 W. VA. L. REV. 625, 631 (2005) (noting that "public schools in many areas are more segregated than they were before *Brown*"); Leland Ware, *Race and Urban Space: Hypersegregated Housing Patterns and the Failure of School Desegregation*, 9 WIDENER L. SYMP. J. 55, 65 (2002) (stating that "public schools in many urban communities are more segregated now than they were in the pre-*Brown* era").

20. *See* ERICA FRANKENBERG ET AL., HARV. C.R. PROJECT, A MULTIRACIAL SOCIETY WITH SEGREGATED SCHOOLS: ARE WE LOSING THE DREAM? 4 (2003), *available at* http://www.civilrightsproject.harvard.edu/research/reseg03/AreWeLosingtheDream.pdf (discussing a twelve year decline in the desegregation of African-American students).

21. *See infra* Part I (discussing the negative effects of resegregation on public education).

vere costs imposed by racial and economic segregation in public education.

I. SCHOOL HOUSE ROCK: RESEGREGATION OF PUBLIC EDUCATIONAL INSTITUTIONS

Throughout our history, public education has occupied a significant role in our society. Its importance has been the bedrock of legal decisions concerning the provision of educational opportunities to undocumented children,[22] children with disabilities,[23] and minority students.[24] As recognized by the Supreme Court in *Brown I*:

> Today, education is perhaps the most important function of state and local governments . . . . It is the very foundation of good citizenship. Today it is a principal instrument in awakening the child to cultural values, in preparing him for later professional training, and in helping him to adjust normally to his environment. In these days, it is doubtful that any child may reasonably be expected to succeed in life if he is denied the opportunity of an education. Such an opportunity, where the state has undertaken to provide it, is a right which must be made available to all on equal terms.[25]

Unfortunately, *Brown I*'s recognition of the inherent inequality of racially segregated schools[26] has not prevented such segregation from occurring. This section explores two factors that have contributed to the resegregation of public educational institutions: first, the shift in Supreme Court jurisprudence regarding mandatory desegregation efforts; and second, local school districts' adherence to the "neighborhood school concept" when making student assignment decisions. The Supreme Court's dilution of desegregation mandates and school districts' use of racially segregated neighborhoods as criteria for student assignments have both exacerbated the resegregation trends currently afflicting public educational institutions.

A. The Court Giveth, the Court Taketh Away

The attainment of equality in public education for racial and ethnic minority students has often been pursued via legal measures. From

---

22. *See* Plyler v. Doe, 457 U.S. 202, 230 (1982) (invalidating a Texas statute that denied public education to children not legally admitted to the country).

23. *See* Cedar Rapids Comm. Sch. Dist. v. Garret, 526 U.S. 66, 67 (1999) (holding that Congress' intent "to open the doors of public education to all qualified children" required the school district to provide nursing services to a quadriplegic student in accordance with federal disability law (citing Bd. of Educ. of Hendrick Hudson Cent. Sch. Dist., Westchester County v. Rowley, 458 U.S. 176, 192 (1982))).

24. *See* Brown v. Bd. of Educ. (*Brown I*), 347 U.S. 483, 495 (1954) (invalidating segregation of races in public schools); *Grutter*, 539 U.S. at 325 (upholding the narrowly tailored use of race in higher education admissions decisions).

25. *Brown I*, 347 U.S. at 493.

26. *Id.* at 495 ("We conclude that in the field of public education the doctrine of 'separate but equal' has no place. *Separate educational facilities are inherently unequal.*" (emphasis added)).

Brown I to *Grutter*, Supreme Court intervention has helped to open the school house doors for countless numbers of students of color.[27] Despite such access, however, African-American and Hispanic students continue to lag behind their white counterparts in terms of academic achievement.[28] This phenomenon can be explained, in part, by the environments in which many minority students are educated.[29]

Due to the resegregation trend experienced by many public schools, an astounding number of African-American and Hispanic children are educated in racially and economically segregated schools. "[A]lmost three-fourths of black and Latino students attend schools that are predominantly minority."[30] Of the 2.4 million students attending schools that are 99%-100% minority, African-American and Hispanic students account for 2.3 million.[31] Unfortunately, "[t]he schools that have the highest minority enrollment also have the highest incidence of student poverty: [i]n 87% of schools that are over 90% minority (African-American and Hispanic), over half of the students come from families living in poverty."[32] These figures are particularly disturbing when one considers the disadvantages and challenges that students attending such schools must overcome to succeed academically.[33]

---

27. *See* Brief Amici Curiae of Veterans of the Southern Civil Rights Movement and Family Members of Murdered Civil Rights Activists in Support of Respondents at 8, *Grutter*, 539 U.S. 306 (2003) (No. 02-241), 2003 WL 539178 (noting that the number of African-American college graduates has increased from less than 5% in 1960 to approximately 7.5% in 2000; in addition, the number of African-American law students has increased from 1% in 1960 to 7.4% in 1996); Danielle R. Holley, *Is* Brown *Dying? Exploring the Resegregation Trend in our Public Schools*, 49 N.Y.L. SCH. L. REV. 1085, 1086 & n.4 (2004-2005) (discussing the positive impact of the *Brown* decision on racial integration in public schools); Goodwin Liu & William L. Taylor, *School Choice to Achieve Desegregation*, 74 FORDHAM L. REV. 791, 791 n.4 (2005) ("In 1965, only 15.2% of African-Americans between the ages of twenty-five and twenty-nine had attended college; by 1995, that number had risen to 44.9%. Among African-Americans in that age bracket, 15.3% had completed four or more years of college in 1995, compared to 6.8% in 1965." (citing James T. Patterson, BROWN V. BOARD OF EDUCATION: A CIVIL RIGHTS MILESTONE AND ITS TROUBLED LEGACY 230 (2001))). "In the 1990s, college enrollment by students of color increased by nearly 50%." *Id.* (citing William B. Harvey, AM. COUNCIL ON EDUC., MINORITIES IN HIGHER EDUCATION 2001-2002: NINETEENTH ANNUAL STATUS REPORT 2 (2002)).

28. *See* Eboni S. Nelson, *What Price* Grutter*? We May Have Won the Battle, But Are We Losing the War?*, 32 J.C. & U.L. 1, 8-9, 25-26 (2005) (discussing racial disparities in educational achievement).

29. *See Gratz*, 539 U.S. 244 at 299-300 (Ginsberg, J., dissenting) (noting that "African-American and Hispanic children are all too often educated in poverty-stricken and underperforming institutions").

30. FRANKENBERG ET AL., *supra* note 20, at 28.

31. *Id.*

32. James E. Ryan, *Schools, Race, and Money*, 109 YALE L.J. 249, 273 (1999); *see also* GARY ORFIELD & CHUNGMEI LEE, HARV. C.R. PROJECT, BROWN AT 50: KING'S DREAM OR PLESSY'S NIGHTMARE? 22 (2004), *available at* http://www.civilrightsproject.harvard.edu/research/reseg04/brown50.pdf (concluding that "students in highly segregated neighborhood schools are many times more likely to be in schools of concentrated poverty").

33. *See* ORFIELD & LEE, *supra* note 32, at 21-22 (detailing poverty concentrated school disadvantages such as school deterioration, lack of resources, less experienced teachers and fewer college preparatory courses).

Although the Supreme Court has issued decisions to help facilitate the provision of equal educational opportunities to minority students,[34] the Court has also issued opinions—three, in particular, referred to as the "resegregation trilogy"[35]—that have hindered the progress of desegregation.[36] The Court's decisions in *Board of Education of Oklahoma City v. Dowell*,[37] *Freeman v. Pitts*,[38] and *Missouri v. Jenkins*[39] have relaxed school districts' responsibilities and duties to eliminate all vestiges of racial segregation, thereby permitting the premature dissolution of federally mandated desegregation decrees when racial imbalance persists.[40]

The Supreme Court's decision in *Dowell* evidences its reluctance to continue taking an active role in the desegregation of public educational institutions[41] as it had in previous cases.[42] The Court's decision appears to be guided by its pronouncement that "federal supervision of local school systems [was] intended as a *temporary* measure to remedy past discrimination."[43] To hasten the return of educational decisions to local school officials, the Court set forth a less stringent test to determine whether a school system has successfully complied with a desegregation decree so as to warrant its dissolution. Unlike the Court's demand in *Green v. County School Board of New Kent County*[44] that school boards

34. *See supra* notes 24, 27; *see also* Green v. County Sch. Bd. of New Kent County, Va., 391 U.S. 430, 437-38 (1968) (placing an "affirmative duty" on school boards operating segregated systems "to take whatever steps might be necessary to convert to a unitary system in which racial discrimination would be eliminated root and branch"); Brown v. Bd. of Educ. (*Brown II*), 349 U.S. 294, 301 (1955) (instructing district courts to enter desegregation decrees that mandate the admission of African-American students into public schools "with all deliberate speed").

35. Ware, *supra* note 19, at 63.

36. *See id.* at 65 (referring to the three cases as "a three-fold shift from an affirmative duty to eliminate all vestiges of segregation to acquiescence to resegregation"). Arguably, Supreme Court cases decided prior to the resegregation trilogy have also hindered the progress of desegregation and educational equality. *See, e.g.*, Milliken v. Bradley, 418 U.S. 717 (1974) (prohibiting the imposition of multi-district desegregation policies to remedy single-district intentional discrimination); San Antonio Indep. Sch. Dist. v. Rodriguez, 411 U.S. 1 (1973) (upholding local property taxation as a constitutionally permissible method for school financing despite resulting disparities in per-student expenditures).

37. 498 U.S. 237 (1991).

38. 503 U.S. 467 (1992).

39. 515 U.S. 70 (1995).

40. *See* Nancy Levit, *Embracing Segregation: The Jurisprudence of Choice and Diversity in Race and Sex Separatism in Schools*, 2005 U. ILL. L. REV. 455, 465-73 (discussing the impact of the three cases on district courts' decisions to dissolve desegregation orders, "even if desegregation actually had not been accomplished").

41. *See* Holley, *supra* note 27, at 1090 & n.31 (describing the Supreme Court's decisions in *Dowell*, *Freeman* and *Jenkins* as evidence of its "hostility towards federal court supervision of school desegregation").

42. *See supra* note 34; *see also* Washington v. Seattle Sch. Dist. No. 1, 458 U.S. 457, 471, 487 (1982) (striking down a state initiative intended to "interfere . . . with desegregative busing"); Swann v. Charlotte-Mecklenburg Bd. of Educ., 402 U.S. 1, 25, 30 (1971) (upholding the ordering of a racial balance requirement and bus transportation as permissible tools of school desegregation).

43. *Dowell*, 498 U.S. at 238 (emphasis added); *see also id.* (stating that desegregation decrees "are not intended to operate in perpetuity"). Some scholars suggest that such statements evidence the Court's "impatience with the duration of desegregation orders," or perhaps, "an abandonment of the original purpose" of desegregation. *See* Levit, *supra* note 40, at 472 & n.91.

44. 391 U.S. 430 (1968).

develop systems "in which racial discrimination would be eliminated root and branch,"[45] the *Dowell* Court instructed lower courts to ask only "whether the Board had complied *in good faith* with the desegregation decree . . . and whether the vestiges of past discrimination had been eliminated *to the extent practicable*."[46] This test appears to concede the point that the complete elimination of segregation is impractical; therefore, school districts that demonstrate a good faith effort to desegregate and eliminate traces of past discrimination can be released from judicial control and supervision even though circumstances remain that hinder desegregation.[47]

The Court reiterated the *Dowell* test in *Freeman* as it continued to chip away at the desegregation safeguards that it had previously helped to establish.[48] In *Freeman*, respondents argued that a district court should not relinquish its supervision and control over a school system until the school district fully complies with all components of a desegregation decree.[49] The Court rejected this argument and sanctioned the incremental withdrawal of judicial supervision once a school system is determined to be in compliance with certain categories of a desegregation order.[50] In arriving at its decision, the Court once again relied heavily on its desire to return control of school systems to state and local officials.[51]

Guided by the "ultimate objective . . . to return school districts to the control of local authorities," the Court reasoned that such restoration is "essential to restore [local authorities'] true accountability in our governmental system."[52] One must be mindful, however, that local authorities' previous control of school systems resulted in unequal and segregated dual systems—systems that necessitated the imposition of court-ordered desegregation decrees in attempts to remedy them. Over ten years passed before the local school officials in *Freeman* took affirmative steps to adhere to the Supreme Court's mandate that school districts

45. *Green*, 391 U.S. at 438; *see also Swann*, 402 U.S. at 15 (stating that the Supreme Court's objective "remains to eliminate from the public schools *all* vestiges of state-imposed segregation" (emphasis added)).

46. *Dowell*, 498 U.S. at 249-50 (emphasis added).

47. *See* Holley, *supra* note 27, at 1092 (concluding that the *Dowell* test excludes the possibility of resegregation as a factor for determining unitary status so as to warrant the dissolution of a desegregation decree); Levit, *supra* note 40, at 464-65 (discussing the *Dowell* test as an invitation to lower courts to dissolve desegregation decrees even if segregation continues to exist); Ware, *supra* note 19, at 64 (concluding that the *Dowell* test allows for a finding of unitary status despite a showing that schools remained racially segregated due to housing patterns).

48. *Freeman*, 503 U.S. at 492.

49. *See id.* at 471.

50. *See id.* at 490-91.

51. *See id.* at 489-90 ("Partial relinquishment of judicial control . . . can be an important and significant step in fulfilling the district court's duty to return the operations and control of schools to local authorities.").

52. *Id.* at 489-90.

desegregate "with all deliberate speed,"[53] and such steps were initiated only after the respondents filed their lawsuit.[54] At the time the Court decided *Freeman*, over thirty-five years had passed since its decision in *Brown II*. Nevertheless, the local DeKalb County school officials continued to operate a school system that was violative of the desegregation order.[55] Such failures and delayed action should cause district courts and the Supreme Court pause when considering the arguably premature return of school systems to local control.

The *Freeman* decision may also hinder desegregation efforts because of its discussion regarding a school district's duty (or lack thereof) to remedy racial imbalance that continues to exist in its schools. The respondents in *Freeman* presented evidence demonstrating the continuance of racial imbalance in DeKalb County schools.[56] Petitioners argued that such imbalance was not caused by prior *de jure* discrimination; rather, it was due to demographic changes within the county.[57] The Supreme Court rejected the Eleventh Circuit's contention that the school district "bore the responsibility for the racial imbalance, and in order to correct that imbalance would have to take actions that 'may be administratively awkward, inconvenient, and even bizarre in some situations.'"[58] The Court clarified that "[o]nce racial imbalance traceable to the constitutional violation has been remedied, a school district is under no duty to remedy an imbalance that is caused by demographic factors."[59] When coupled with the Court's sanctioning of the incremental withdrawal of judicial supervision once a school district has been deemed to have complied with certain provisions of a desegregation decree, this pronouncement begs the question: With whom does the duty lie to desegregate schools if it does not lie with local school districts? If, as in *Freeman*, school districts are partially or fully released from their desegregation orders even though their minority students continue to attend racially segregated schools, then the likelihood of achieving true desegregation in public education and the benefits that arise from such educational environments is doubtful.

School districts' ability to remedy resegregation of their educational institutions may be further hindered by the Supreme Court's decision in *Jenkins*. The district court in *Jenkins* ordered a variety of educational programs and initiatives in its efforts to improve the educational quality

53. *Brown II*, 349 U.S. at 301.

54. *Freeman*, 503 U.S. at 472 (describing the school system's reluctant response to desegregation mandates).

55. *See id.* at 474 (discussing the district court's findings that the school system continued to be segregated with regards to "teacher and principal assignments, resource allocation, and quality of education").

56. *Id.* at 476-77.

57. *Id.* at 478.

58. *Id.* at 485 (citing *Swann*, 402 U.S. at 28).

59. *Id.* at 469 (citations omitted).

of the Kansas City, Missouri School District and to eliminate all vestiges of segregation.[60] The two measures challenged by the State were salary increases for instructional and noninstructional staff and remedial quality education programs.[61] The State argued that the requirement of salary increases for teachers and non-teaching staff exceeded the district court's remedial authority.[62] In upholding the State's challenge, the Supreme Court agreed that a district court cannot use 'interdistrict' measures to remedy 'intradistrict' constitutional violations.[63] Concluding that measures such as salary increases were motivated by the district court's pursuit of "desegregative attractiveness,"[64] the Court rejected the Eighth Circuit's contention that "[*v*]*oluntary* interdistrict remedies may be used to make meaningful integration possible in a predominantly minority district."[65] This rejection greatly restricts district courts' ability to fashion effective measures that can be used to remedy the devastating effects of segregation.

Perhaps one of the most harmful lingering effects of segregation is minority students' lack of academic achievement.[66] Greater numbers of African-American students fail to complete high school as compared to white students.[67] African-American students, many of whom attend racially imbalanced schools, routinely score lower than their white counterparts on standardized tests.[68] Fewer African-American adults, as

---

60. *See* Missouri v. Jenkins, 515 U.S. 70, 75-78 (1995) (describing the district court's ordering of class reductions, magnet school programs, capital improvements, and salary increases as measures to improve academic achievement and remedy effects of segregation).

61. *Jenkins*, 515 U.S. at 80.

62. *Id.* at 84.

63. *Id.* at 89-98 (citations omitted).

64. *Id.* at 80 (citation omitted). According to the Court, "desegregative attractiveness" refers to the implementation of programs and initiatives that will improve the attractiveness of schools within a school district such that nonminority students who are not presently attending schools within the district will decide to enroll, thereby helping to desegregate the schools. *See id.* at 91-92.

65. *Id.* at 91 (citing Jenkins v. Missouri, 855 F.2d 1295, 1302 (8th Cir. 1988)).

66. *See Brown I*, 347 U.S. at 494 (noting that segregation has a negative impact on the educational development of African-American children); Lisa J. Holmes, Comment, *After* Grutter*: Ensuring Diversity in K-12 Schools*, 52 UCLA L. REV. 563, 586-87 (2004) (discussing research suggesting that segregated educational environments may have detrimental effects on the academic development of minority children); Roslyn Arlin Mickelson, *Achieving Equality of Educational Opportunity in the Wake of Judicial Retreat from Race Sensitive Remedies: Lessons from North Carolina*, 52 AM. U. L. REV. 1477, 1485 & n.33 (2003) (citing research showing segregation's adverse effects on minority students' academic achievements).

67. In 2004, the drop out rate for African-Americans age 16-24 was 11.8% as compared to 6.8% for Whites. NAT'L CTR. FOR EDUC. STATISTICS, THE CONDITION OF EDUCATION 2006, STATUS DROPOUT RATES BY RACE/ETHNICITY tbl.26-1 (2006), *available at* http://nces.ed.gov/programs/coe/2006/section3/table.asp?tableID=481.

68. *See* Michael J. Songer, Note, *Decline of Title VII Disparate Impact: The Role of the 1991 Civil Rights Act and the Ideologies of Federal Judges*, 11 MICH. J. RACE & L. 247, 267-68 (2005). Songer reported:

> The most recent study by the Civil Rights Project at Harvard University found that 70% of minority children attend American schools with majority-minority populations. More than one-third of these children attend schools that are comprised of at least 90% African American students. African American students continue to score significantly lower than White students on standardized tests used in college and graduate school admissions.

compared to white adults, obtain a college education.[69] Such achievement gaps are due, in part, to disparities existing between the quality of teachers at poor, minority-concentrated schools and their more affluent, white counterparts.

For example, novice teachers, who are obviously not as qualified as more experienced teachers, are disproportionately assigned to high poverty, majority-minority schools.[70] The percentage of high school students attending high-minority, high-poverty schools that are taught English, science, and mathematics by "teachers who have neither a major nor certification in the subject they teach" is twice the percentage of students encountering the same experience at schools with low minority and poverty populations.[71] Obviously, such disparities have a detrimental impact on minority students' academic achievement. If such disparities could be rectified, then the positive impact on student achievement could be tremendous, and the hope of eliminating all vestiges of segregation could become a reality.

This is what the district court in *Jenkins* attempted to accomplish by ordering the State to fund salary increases in its desegregation efforts.[72] Remedial measures such as salary increases can positively affect teacher quality disparities and, consequently, student achievement disparities, by attracting more highly-qualified teachers and personnel to minority-concentrated schools. Exposing minority students to more experienced, more educated and more effective teachers will improve their educational opportunities and lessen the detrimental effects of segregation and past discrimination.[73] Unfortunately, by finding that the district court exceeded its remedial authority by ordering salary increases for school personnel, the Supreme Court deprived district courts of a valuable tool

---

Id. (footnotes omitted). Keith R. Walsh, Book Note, *Color-Blind Racism in* Grutter *and* Gratz, 24 B.C. THIRD WORLD L.J. 443, 450-51 (2004) (discussing the gap between African-American and white students' scores on standardized admissions tests).

69. *See* Walsh, *supra* note 68, at 450 & n.49 (noting that 28% of white adults are college educated as compared to 16% of African-American adults and that "[a]s of 2000, only 17.8% of African Americans over the age of twenty-five had completed four or more years of college, while 34% of their white counterparts could say the same" (citation omitted)).

70. Novice teachers are assigned to minority concentrated schools at twice the rate as those assigned to schools with low minority populations. HEATHER G. PESKE & KATI HAYCOCK, THE EDUC. TRUST, TEACHING INEQUALITY: HOW POOR AND MINORITY STUDENTS ARE SHORTCHANGED ON TEACHER QUALITY 2 (2006), *available at* http://www2.edtrust.org/NR/rdonlyres/010DBD9F-CED8-4D2B-9E0D-91B446746ED3/0/TQReportJune2006.pdf.

71. NAT'L CTR. FOR EDUC. STATISTICS, THE CONDITION OF EDUCATION, OUT-OF-FIELD TEACHING BY POVERTY CONCENTRATION AND MINORITY ENROLLMENT ¶ 1 (2004), *available at* http://nces.ed.gov/programs/coe/2004/section4/indicator24.asp; *see also* PESKE & HAYCOCK, *supra* note 70, at 3 (noting that "[i]n secondary schools serving the most minority students, almost one in three classes are assigned to an out-of-field teacher compared to about one in five in low-minority schools").

72. *See Jenkins*, 515 U.S. at 80.

73. *See* PESKE & HAYCOCK, *supra* note 70, at 11 (noting research findings indicating the positive impact interaction with highly effective teachers can have on low-performing students).

in their efforts to eradicate vestiges of segregation and create educational equality.

The establishment of equality in public education for racial and ethnic minority students is also threatened by the Court's admonition regarding the use of student achievement levels as a measure to determine partial unitary status.[74] After finding that the school district "had not reached anywhere close to its 'maximum potential because the District is still at or below national norms at many grade levels,'"[75] the district court ordered the State to continue funding quality education programs designed to improve the educational achievement of all students, especially African-Americans.[76] The State challenged the order on the grounds that improvement on test scores is not a requirement to achieve partial unitary status.[77] In upholding the State's challenge, the Court directed the district court to "sharply limit, if not dispense with, its reliance on" student performance on achievement tests in its determination of partial unitary status.[78] Although the district court maintained that a school district must achieve its "maximum potential" regarding its desegregation efforts before it can be deemed to have partially complied with a desegregation decree, the Supreme Court rejected this test and re-imposed the lower standard of "practicability" articulated in *Dowell*.[79] By rejecting the district court's more stringent test, the Court invited premature findings of partial unitary status despite the fact that minority students continue to suffer from reductions in academic achievement.

The Court also extends this invitation by its directive to the district court to be mindful of its end goal to return control of a school system to state and local officials.[80] In its efforts to expedite the return of educational decisions to local control, the Court appears to have abandoned its previous stance "that the court has not merely the power but the duty to render a decree which will so far as possible eliminate the discriminatory effects of the past as well as bar like discrimination in the future."[81] As shown by the number of schools experiencing resegregation following

74. A previously intentionally segregated school district achieves partial unitary status when a district court determines that the school district has successfully complied with certain, although not all, components of a desegregation decree so as to be deemed to no longer discriminate on the basis of race with regards to that particular component. For further discussion of the term "unitary," see Holley, *supra* note 27, at 1091 n.33.

75. *Jenkins*, 515 U.S. at 101 (citation omitted).

76. *Id.* at 73.

77. *Id.* at 101.

78. *Id.*

79. *See id.* (stating the partial unitary test as "whether the reduction in achievement by minority students attributable to prior *de jure* segregation has been remedied *to the extent practicable*" (emphasis added)).

80. *See id.* at 102 (quoting Freeman v. Pitts, 503 U.S. 467, 489 (1992)).

81. *Green*, 391 U.S. at 438 n.4 (quoting Louisiana v. United States, 380 U.S. 145, 154 (1965)).

the dissolution of desegregation decrees,[82] the withdrawal of judicial oversight, based on the relaxed standards of *Dowell*, *Freeman*, and *Jenkins*, has had a negative impact both on the elimination of discriminatory effects and on the prevention of such harmful effects in the future.[83] Such impact is due, in part, to local officials' reliance on the "neighborhood school concept" when making student assignment decisions. As demonstrated in the following section, employing student assignment methods that are based on racially segregated neighborhoods produces resegregation in public education and the detrimental effects that accompany such environments.

B. The Neighborhood School Dilemma

Historically, the neighborhood school concept, which calls for the assignment of students to schools that are in close proximity to their homes, has been a preferred method for making school assignment decisions.[84] Many argue that adherence to neighborhood schools provide educational benefits ranging from increased parental and community involvement that results in improved student achievement,[85] to reductions in transportation costs, which provide additional funding for teacher

---

82. *See, e.g.*, John Charles Boger, *Education's "Perfect Storm"? Racial Resegregation, High-Stakes Testing, and School Resource Inequities: The Case of North Carolina*, 81 N.C. L. REV. 1375, 1394-95 (2003) (noting the resegregation consequences flowing from the dissolution of a thirty-year-old desegregation decree) Boger stated:

> The consequences flowing from the *Capacchione* [*v. Charlotte-Mecklenburg Schools*] ruling were swift and dramatic: in the 2002-2003 school year, the number of Charlotte-Mecklenburg schools with minority enrollment of 91% to 100% more than doubled from the previous year—from seven elementary schools in 2001-2002 to sixteen in 2002-2003, and from two middle schools to four. There was no change in the number of elementary and middle schools with minority enrollment of 20% or less.

Id. (footnote omitted); *see also* Holley, *supra* note 27, at 1095-96 & n.60 (noting that thirty-four of the thirty-eight school districts that have achieved unitary status since the *Dowell* decision have experienced resegregation as measured by "a decrease in the exposure of black students to white students, and the exposure of Latino students to white students").

83. *See* ORFIELD & LEE, *supra* note 32, at 18 (attributing the resegregation trend "to the impact of three Supreme Court decisions between 1991 and 1995 limiting school desegregation and authorizing a return to segregated neighborhood schools").

84. *See Swann*, 402 U.S. at 28 (noting the Supreme Court's recognition that "[a]ll things being equal . . . it might well be desirable to assign pupils to schools nearest their homes"); *see also* Levit, *supra* note 40, at 456 & n.6 (referring to state initiatives to pass and implement "Neighborhood Schools Acts"). Student assignments based on proximity to one's home are especially favored when compared to the alternative of busing. *See* Davison M. Douglas, *The Quest for Freedom in the Post-*Brown *South: Desegregation and White Self-Interest*, 70 CHI.-KENT L. REV. 689, 746-47 (2004) (quoting former North Carolina Governor Robert Scott as stating, "The neighborhood-school concept has been the strength of our public education system in North Carolina and our state has been committed to that policy for some time. It is sound educational policy and must be preserved." (citation omitted)); *Id.* at 747 (quoting former President Richard Nixon as describing neighborhood schools as "the most appropriate . . . system" (citation omitted)); Levit, *supra* note 40, at 456 & n.5 (referring to Congressional anti-busing legislation setting forth the government's official policy that "students attend neighborhood schools").

85. Patrick James McQuillan & Kerry Suzanne Englert, *The Return to Neighborhood Schools, Concentrated Poverty, and Educational Opportunity: An Agenda for Reform*, 28 HASTINGS CONST. L.Q. 739, 743 (2001).

salaries and educational programs.[86] These benefits, however, are greatly outweighed by the detrimental effects that accompany many neighborhood school decisions: namely, the resegregation of elementary and secondary schools and the overwhelming challenges that are present in such environments.

Following the termination of desegregation decrees and the return of educational decisions to local control, many school districts returned to the neighborhood school concept when making their student assignment decisions.[87] Considering the rate of residential segregation in communities throughout the country, it is not surprising that such decisions have resulted in the resegregation of public schools.[88] "One-third of all African Americans in the United States live under conditions of intense racial segregation."[89] In 2000, over 230 American urban communities could be described as "hypersegregated" or "partially segregated."[90] Therefore, in accordance with student assignment policies that assign students to schools based on neighborhood proximity, schools populated by students living in these areas will also experience high levels of racial segregation,[91] which often brings about adverse educational consequences.

---

86. Kenneth O'Neil Salyer, *Beyond* Zelman*: Reinventing Neighborhood Schools*, 33 J.L. & EDUC. 283, 287-88 (2004) (discussing the advantages of neighborhood schools). Other proffered advantages of neighborhood schools include the following: reduction in student-teacher ratios, reduction in travel safety hazards, creation of sense of community, and simplification of student assignment policies. *Id.*; *see also* Kevin Brown, *Has the Supreme Court Allowed the Cure for* De Jure *Segregation to Replicate the Disease*, 78 CORNELL L. REV. 1, 35 (1992).

87. Molly S. McUsic, *The Future of* Brown v. Board of Education*: Economic Integration of the Public Schools*, 117 HARV. L. REV. 1334, 1342 (2004) (noting school districts' reversion to "neighborhood schools placed in segregated neighborhoods"); Myron Orfield, *Choice, Equal Protection, and Metropolitan Integration: The Hope of the Minneapolis Desegregation Settlement*, 24 LAW & INEQ. 269, 294, 322-23 (2006) (discussing the impact of termination of desegregation decrees and reinstitution of neighborhood schools on the resegregation of Minneapolis schools); *Speech by Theodore M. Shaw: From* Brown *to* Grutter*: The Legal Struggle for Racial Equality*, 16 WASH. U. J.L. & POL'Y 43, 52 (2004) (recounting the incident in which the Norfolk Virginia School District expressed "its intention to return to neighborhood schools by abandoning its desegregation plan after a declaration of unitary status").

88. Michael Selmi, *Race in the City: The Triumph of Diversity and the Loss of Integration*, 22 J.L. & POL. 49, 69 (2006) (noting that school segregation "follows housing segregation"); Ware, *supra* note 19, at 56 (attributing the failure of desegregation efforts in many urban schools to pervasive "segregated housing patterns").

89. Ware, *supra* note 19, at 65 (quoting DOUGLAS S. MASSEY & NANCY A. DENTON, AMERICAN APARTHEID: SEGREGATION AND THE MAKING OF THE UNDERCLASS 77 (1993)).

90. Boger, *supra* note 82, at 1402 & n.97 (detailing residential segregation levels in metropolitan areas).

91. *Id.* at 1400 (predicting that "residential segregation will prove especially likely to lead to school resegregation if districts choose student assignment strategies based on neighborhood schools"); *see also id.* at 1407-08 (discussing racial segregation increases in North Carolina schools following the implementation of neighborhood schools assignment plans); GARY ORFIELD & CHUNGMEI LEE, HARV. C.R. PROJECT, RACIAL TRANSFORMATION AND THE CHANGING NATURE OF SEGREGATION 9 (2006), *available at* http://www.civilrightsproject.harvard.edu/research/deseg/Racial_Transformation.pdf (noting that "[s]ince the Supreme Court authorized a return to segregated neighborhood schools . . . , the percentage of black students attending majority nonwhite schools increased in all regions from 66 percent in 1991 to 73 percent in 2003-4") (footnote omitted).

Research shows that students attending racially segregated schools, which are often economically segregated as well,[92] encounter tremendous challenges that greatly hinder their educational achievement. Students attending schools with majority minority student populations are often educated in "substandard and deteriorating facilities."[93] Their learning environments often suffer from "shortages of library books, computers, or laboratory equipment."[94] The teachers who educate them are often less qualified than those teaching at racially and economically diverse schools.[95] This lack of resources leads to disparities in minority students' academic achievement as measured by standardized tests scores,[96] high school drop-out and graduation rates,[97] college matriculation rates,[98] and post-graduate degrees.[99]

Not only are students attending segregated schools forced to overcome educational resources deficiencies, but they are also deprived of the educational benefits related to interacting with students who possess higher educational aspirations. Unfortunately, many minority students who live in lower-income, racially segregated neighborhoods and attend lower performing schools within those neighborhoods have low expectations regarding academic achievement. In fact, some minority communities suffer from a culture that devalues academic success,[100] which significantly undermines minority students' academic expectations and aspirations. As noted by scholar John Charles Boger:

> [A] pupil's achievement is strongly related to the educational backgrounds and aspirations of the other students in the

92. *See supra* note 32 and accompanying text.

93. Levit, *supra* note 40, at 497 (quoting Leland Ware, *Redlining Learners: Delaware's Neighborhood Schools Act*, 20 DEL. LAW. 14, 16 (2002)).

94. Boger, *supra* note 82, at 1382.

95. *Id.* at 1447 (citations omitted); *see also* Levit, *supra* note 40, at 498; *supra* notes 69-70 and accompanying text.

96. *See* Curt A. Levey, *Racial Preferences in Admissions: Myths, Harms, and Alternatives*, 66 ALB. L. REV. 489, 502 (2003) (discussing racial disparities in minority and nonminority students' standardized test scores); Walsh, *supra* note 68 (discussing disparities between Blacks and Whites regarding their performance on standardized tests).

97. In 2004, the high school drop out rate for African-American students age 16-24 was 11.8% compared to 6.8% for their white counterparts. NAT'L CTR. FOR EDUC. STATISTICS, THE CONDITION OF EDUCATION 2006, STATUS DROPOUT RATES BY RACE/ETHNICITY tbl.26-1 (2006), *available at* http://nces.ed.gov/programs/coe/2006/section3/table.asp?tableID=481. In 2005, fewer Blacks than Whites age 25-29 years old had completed high school (87% vs. 93%). NAT'L CTR. FOR EDUC. STATISTICS, THE CONDITION OF EDUCATION 2006, EDUCATIONAL ATTAINMENT tbl.31-1 (2006), *available at* http://nces.ed.gov/programs/coe/2006/section3/table.asp?tableID=492.

98. In 2005, only 49.0% of African-Americans between the ages of 25-29 had completed at least some college as compared to 64.3% of Whites. NAT'L CTR. FOR EDUC. STATISTICS, THE CONDITION OF EDUCATION 2006, EDUCATIONAL ATTAINMENT tbl.31-2 (2006), *available at* http://nces.ed.gov/programs/coe/2006/section3/table.asp?tableID=493; *see also* A. Mechele Dickerson, *Race Matters in Bankruptcy*, 61 WASH & LEE L. REV. 1725, 1769-70 (2004).

99. In 2005, 34.1% of Whites between the ages of 25-29 had obtained a bachelor's degree or higher while only 17.5% of Blacks had achieved the same educational success. NAT'L CTR. FOR EDUC. STATISTICS, THE CONDITION OF EDUCATION 2006, EDUCATIONAL ATTAINMENT tbl.31-3 (2006), *available at* http://nces.ed.gov/programs/coe/2006/section3/table.asp?tableID=494.

100. Nelson, *supra* note 28, at 26 & n.127.

> school . . . . Thus . . . if a minority pupil from a home without much educational strength is put with schoolmates with strong educational backgrounds, his achievement is likely to increase.[101]

If, in fact, "the social characteristics of a school's student body [are] the single most important school-related factor in predicting minority student achievement,"[102] student assignment plans that rely on poor, racially segregated neighborhoods will only exacerbate the current disparities existing between minority and non-minority student achievement.

Considering the detrimental impact the creation of neighborhood schools has on desegregation efforts and, consequently, the quality of education received by many minority students, one wonders why school boards continue to create and advocate for them. While school boards' decisions to adhere to neighborhood schools may be attributed, in part, to their purported benefits,[103] parents' vocal opposition to busing and school boundary proposals has also greatly influenced school boards' actions. Due to the fact that the overwhelming majority of school boards are elected positions,[104] their members must confront political pressures that are brought to bear upon them by their constituents. Having particular influence on school board members are those voting parents who organize in efforts to oppose school boundary and student assignment proposals that attempt to diversify schools, both in terms of race and socioeconomic status.[105]

---

101. Boger, *supra* note 82, at 1415 (quoting JAMES S. COLEMAN ET AL., EQUALITY OF EDUCATIONAL OPPORTUNITY 22 (1966)).

102. *Id.* (citation omitted).

103. *See supra* notes 85-86 and accompanying text.

104. *See* FREDERICK M. HESS, SCH. OF EDUC. AND DEP'T OF GOV'T, UNIV. OF VA., SCHOOL BOARDS AT THE DAWN OF THE 21ST CENTURY 5 (2002), *available at* http://www.nsba.org/site/docs/1200/1143.pdf (reporting that in a survey of 2000 school boards, 93% were entirely elected).

105. While parent groups comprise 52.1% of constituent groups that are "active" in school board elections, ethnic or racial groups only comprise 18.1%. *Id.* at 37 tbl.42. *See also* Boger, *supra* note 82, at 1399-1400 (discussing parents' resistance efforts to the proposal of assigning poor, low-performing students to schools where their children attended and to the reassignment of white, middle-class students to lower income, lower performing schools); Dana Banker, *Plantation Parents Join Busing Debate, School Boundaries Face Challenge at Meeting*, S. FLA. SUN-SENTINEL, Mar. 27, 1995, at 1B (stating the goal of parents who oppose school boundary proposals that would require their children to be bused to a predominantly Black school is to "[m]ake board members realize that this Plantation [parent] contingent is a sizable group with which to be reckoned"); John Hill, *Good Schools for All Hillsborough*, ST. PETERSBURG TIMES, May 13, 2006, at 12A (stating that "[p]arents of upscale Westchase scolded, taunted and threatened the elected board with political retaliation" because of their discontentment regarding the school board's student reassignment proposal); Ginger Jenkins, *Boundary Committee Endures Wrath of Fall Creek Residents*, HOUSTON CMTY. NEWSPAPERS ONLINE, Apr. 11, 2004, http://www.hcnonline.com/site/index.cfm?newsid=11289350&BRD=1574&PAG=461&dept_id=532207&rfi=8&xb=lutex (discussing parents' vocal opposition to school boundary proposals that would zone their children to Title 1 schools, which have high economically-disadvantaged student populations); Scott Travis, *Parents Protest Plan to Alter School Boundary*, S. FLA. SUN-SENTINEL, Sept. 12, 2000, at 1B (discussing parents' opposition to a school boundary proposal that would add 163 predominantly poor, African-American students to their children's elementary school).

The school board members elected to govern the Humble Independent School District in Humble, Texas faced similar opposition in 2003 after announcing their proposals to redraw school boundaries.[106] Because some of the boundary proposals called for certain middle-upper class, predominantly white neighborhoods to be zoned to schools that would have predominantly minority, lower-income student populations, parents and residents residing in the predominantly white neighborhoods voiced their dissent and lobbied school board members to vote to keep their children at the "good" schools.[107] Although the decision was not unanimous, the school board acquiesced and voted to accept boundary proposals that would allow the parents' children to attend the more desirable schools.[108] Unfortunately, the same boundary decision also created racially and economically segregated schools due to the extraction of white, middle-class students.[109]

If school board members continue to employ student assignment policies that rely on racially segregated housing patterns and to yield to political parental pressures that oppose diversification and, thereby, desegregation efforts, then the goal of attaining educational equality for minority students will be unrealized. In attempts to avoid the harmful costs associated with resegregation, some school districts have voluntarily implemented plans that consider students' race when making student assignment decisions. The next section examines two such plans and their attempts to further compelling interests via constitutional means.

II. TAKING MATTERS INTO THEIR OWN HANDS: PUBLIC SCHOOLS' VOLUNTARY USE OF RACE-CONSCIOUS STUDENT ASSIGNMENT PLANS

Due to the resegregation trend that is currently plaguing American public educational institutions, school districts have begun to experiment with various measures intended to diversify elementary and secondary

106. Cindy Horswell, *School Boundaries Draw Parents' Wrath/Humble ISD Stirs Campus Controversy*, HOUSTON CHRON., Nov. 15, 2003, at A31.

107. Linda Gilchriest, *Tough Choices/Humble ISD Must Decide Controversial Lines Issues*, HOUSTON CHRON., July 10, 2003, at 1 (discussing Fall Creek's ("an upscale subdivision with million-dollar homes") opposition to being zoned to Humble High School "because it would have a greater number of minority and economically disadvantaged students"); *see also* A. Mechele Dickerson, *Caught in the Trap: Pricing Racial Housing Preferences*, 103 MICH. L. REV. 1273, 1280 (2005) (suggesting that parents' desire to have their children attend "good" schools "may actually be a code for a preference to . . . have their children attend nonminority schools"); Jenkins, *supra* note 105.

108. Kristen Wright, *Humble ISD Adopts New Boundaries/New Kingwood Park to Be Scaled Down*, HOUSTON CHRON., Sept. 2, 2004, at 1.

109. Following adoption of the new boundaries, Humble High School was projected to be 45%-65% minority and 45%-55% economically disadvantaged. Meanwhile, more affluent Kingwood High School was projected to have a student body that was only 10%-20% minority and 5%-15% economically disadvantaged. The newly created high school (to which parents lobbied school board members to have their children attend) was projected to be 30%-50% minority and only 15%-25% economically disadvantaged. DEMOGRAPHIC PROJECTIONS CORE COMMITTEE RECOMMENDATION (on file with author).

schools. School districts have implemented school choice programs whereby parents can decide which schools they would like for their child to attend.[110] To encourage parents to choose schools that may have high populations of minority, economically-disadvantaged students, many school districts have introduced programs that provide pre-college courses of study such as the International Baccalaureate (IB) Diploma Program at such schools.[111] School districts have also created magnet schools and programs, which have a particular theme or curricular focus, such as science, technology, mathematics or performing arts, in their efforts to achieve diverse student bodies.[112]

Schools have also taken a more direct approach to achieve their diversity goals by considering students' race and ethnicity when making student assignment decisions. Such consideration has subjected school districts to intense and, in some cases, fatal judicial scrutiny.[113] The school districts in the following two cases, however, successfully overcame the constitutional challenges launched against their race-conscious student assignment plans at the circuit court level. It remains to be seen whether the same will be true following the Supreme Court's consideration of the plans.

A. McFarland v. Jefferson County Board of Education

In its attempts to maintain an integrated school system following the lifting of a desegregation decree, the Jefferson County Public Schools (the "Board") implemented a student assignment plan that includes racial

---

110. Robert A. Frahm, *Court Takes On Race Case, School Desegregation Could Be Affected*, HARTFORD COURANT, June 6, 2006, at A1 (discussing Connecticut's use of school choice to promote racial diversity in elementary and secondary education).

111. In 2006, the Humble Independent School District in Humble, TX announced its plans to institute the IB Diploma Program at Humble High School, which is the most racially diverse and economically-disadvantaged high school in the district. *See* What in the World is IB?, http://www.humble.k12.tx.us/ibpage.htm (last visited Oct. 21, 2006). During the 2004-2005 school year, 25.4% of Humble High School's student population was economically disadvantaged, compared to only 3.3% at Kingwood High School. The minority enrollment at Humble High School is also significantly greater than that at Kingwood High School (50.6% vs. 14%). *See* 2004-2005 ACADEMIC EXCELLENCE INDICATOR SYSTEM CAMPUS REPORTS, *available at* http://www.tea.state.tx.us/perfreport/aeis/2005/campus.srch.html.

112. *See* Hill, *supra* note 105 (discussing a school district's implementation of school choice as a means to "maintain integrated schools by making them more attractive to residents outside their neighborhoods"); Harold A. McDougall, Brown *at Sixty: The Case for Black Reparations*, 47 HOW. L.J. 863, 892 (2004) (discussing the goal of magnet schools "to accomplish or maintain desegregation").

113. *See, e.g.*, Eisenberg *ex rel.* Eisenberg v. Montgomery County Pub. Sch., 197 F.3d 123, 124-25 (4th Cir. 1999) (invalidating a race-conscious student transfer plan that denied students' transfer requests if they would have an adverse impact on the assigned or requested school's diversity levels); Tuttle *ex rel.* Tuttle v. Arlington County Sch. Bd., 195 F.3d 698, 701-02 (4th Cir. 1999) (invalidating an assignment plan that based admission into an alternative kindergarten in part on students' race and ethnicity); Wessman *ex rel.* Wessman v. Gittens, 160 F.3d 790, 791-92 (1st Cir. 1998) (invalidating Boston Latin School's race-conscious admissions policy).

guidelines.[114] While students have the ability to choose which school to attend, their ultimate assignment can be affected by the operation of the racial guidelines, which require African-American student enrollment to be "at least 15% and no more than 50%" of the student body.[115] Although many other non-racial factors affect student assignment, the racial guidelines prohibit some students' admission into particular schools or academic programs based on their race.[116] Because of such effect, students and parents challenged the constitutionality of the Board's race-conscious student assignment plan.

In reviewing the constitutionality of the plan, the district court applied strict scrutiny, which required the Board to demonstrate that its use of race furthers a compelling governmental interest and does so using narrowly tailored means.[117] In formulating what appears to be a novel justification for the use of race in education, the court held that the maintenance of racially integrated elementary and secondary schools constitutes a compelling interest.[118] In assessing the Board's asserted interests,[119] the court found that the educational and societal benefits that are derived from racial diversity in higher education are also produced in the context of elementary and secondary education.[120] The court accepted the Board's argument that "school integration benefits the system as a whole by creating a system of roughly equal components, . . . not one rich and another poor, not one Black and another White."[121] Finally, in holding that the Board's interests are compelling, the court held that the Board was not engaged in unconstitutional "racial balancing" because of its demonstrated commitment to integration and educational equality and the "academic, social and institutional benefits [they] achieve[ ]."[122]

---

114. *See* McFarland *ex rel.* McFarland v. Jefferson County Pub. Sch., 330 F. Supp. 2d 834, 836 (W.D. Ky. 2004), *aff'd* 416 F.3d 513 (6th Cir. 2005), *cert. granted sub nom.* Meredith *ex rel.* McDonald v. Jefferson County Bd. of Educ., 126 S. Ct. 2351 (U.S. June 5, 2006) (No. 05-915).

115. *McFarland*, 330 F. Supp. 2d at 842.

116. For instance, if an African-American student attempts to enroll into a school that already has a black student composition of 50%, then the racial guideline will operate to deny him admission into that particular school since his enrollment would cause the school to exceed its cap on African-American enrollment. Although the Board's open choice policy permits the student to choose the school he would like to attend, "a student's race . . . could determine whether that student receives his or her first, second, third or fourth choice of school." *See id.*

117. *See id.* at 837, 848-49.

118. *See id.* at 855.

119. The Court's statement of the Board's asserted interests is as follows:

> To give all students the benefits of an education in a racially integrated school and to maintain community commitment to the entire school system precisely express the Board's own vision of *Brown* 's promise. The benefits the [Board] hopes to achieve go to the heart of its educational mission: (1) a better academic education for all students; (2) better appreciation of our political and cultural heritage for all students; (3) more competitive and attractive public schools; and (4) broader community support for all [the Board] schools.

Id. at 849 n.29.

120. *See id.* at 853.

121. *Id.* at 854.

122. *Id.* at 855.

Not only did the district court find that the Board used race in its pursuit of compelling interests, but it also concluded that, in most respects, it utilized narrowly tailored means to pursue such interests.[123] The court applied the following four criteria in determining the constitutionality of the race-conscious student assignment plan:

> (1) whether the 2001 Plan amounts to a quota that seeks a fixed number of desirable minority students and insulates one group of applicants from another, (2) whether the applicant is afforded individualized review, (3) whether the 2001 Plan "unduly harm[s] members of any racial group," and (4) whether [the Board] has given "serious, good faith consideration of workable race-neutral alternatives" to achieve its goals.[124]

In finding that the racial guidelines did not operate as a quota, the district court reasoned that they represented a "quite flexible and broad target range," such as that permitted in *Grutter*, and not a "relatively precise target."[125] This reasoning, however, fails to address the fact that the "target range" is actually a Board *requirement* that African-American students comprise 15%-50% of a school's student enrollment.[126] The Board's formulation of its diversity goal as a numerical mandate may prove to be fatal in its quest to seek constitutional approval from the Supreme Court.[127]

Related to the quota criteria is the narrowly tailored requirement that race-conscious student assignment plans afford each student holistic, individualized review. Unlike other courts that have held that the requirement is inapplicable in the context of elementary and secondary education,[128] the district court considered the requirement and found that the Board's plan allows for individualized review, albeit "of a different kind in a different context than the Supreme Court found in *Grutter*."[129] The court reasoned that the Board considers many aspects of each student's application when determining student assignments. "[R]ace is simply one possible factor among many, acting only occasionally as a

---

123. *See id.* at 855-62.

124. *Id.* at 856 (alteration in original) (internal citations omitted).

125. *Id.* at 857. The Court also relied on the varying actual percentages of Black students present at individual schools (20.1%-50.4%) to support its conclusion that the guidelines did not operate as a quota. *See id.*

126. *See id.* at 842 (stating that "the 2001 Plan *requires* each school to seek a Black student enrollment of at least 15% and no more than 50%" (emphasis added)).

127. For further discussion, see *infra* Part III.

128. *See, e.g.*, Parents Involved in Cmty. Sch. v. Seattle Sch. Dist. No. 1 (*Parents II*), 426 F.3d 1162, 1183 (9th Cir. 2005) ("[I]f a noncompetitive, voluntary student assignment plan is otherwise narrowly tailored, a district need not consider each student in a individualized, holistic manner."), *cert. granted* 126 S. Ct. 2351 (U.S. June 5, 2006) (No. 05-908); Comfort v. Lynn Sch. Comm., 418 F.3d 1, 18 (1st Cir. 2005), ("Unlike the *Gratz* and *Grutter* policies, the Lynn Plan is designed to achieve racial diversity rather than viewpoint diversity. The only relevant criterion, then, is a student's race; individualized consideration beyond that is irrelevant to the compelling interest.") (footnote omitted), *cert. denied sub nom.* Comfort *ex rel.* Neumyer, 126 S. Ct. 798 (2005).

129. *McFarland*, 330 F. Supp. 2d at 859.

permissible 'tipping' factor in most of the [Board] assignment process."[130] Because the Board successfully demonstrated that its plan complied with this as well as the other narrowly tailored requirements, the court concluded that its use of race in student assignments was constitutionally permissible.[131]

B. Parents Involved in Community Schools v. Seattle School District No. 1

Employing similar rationale as that utilized by the Board in *McFarland*, the Seattle School District No. 1 (the "District") also adopted an open choice student assignment plan in its attempts to create racially diverse schools and to prevent racial imbalance that would result from adherence to the neighborhood school concept.[132] The plan allows parents to choose which of the ten high schools they want their children to attend, provided a particular school has availability.[133] To address situations in which a school is oversubscribed,[134] the District employs four tiebreakers, the second one being a student's race.[135] Although the District has never engaged in *de jure* segregation and, therefore, has never been ordered to desegregate,[136] as had the *McFarland* Board, it voluntarily uses the racial tiebreaker to ensure diversity or "balance" in the racial composition of its public high schools.[137] The operation of the racial tiebreaker is as follows: If a school's student population deviates from the goal of 40% white and 60% minority (+/-15%), then the racial tiebreaker is used to grant automatic admission to those students whose race will enable the school to move closer to the desired racial composition.[138] Conversely, the racial tiebreaker also operates to deny admission to those students whose race does not further the District's diversity goals.[139] Because the District, a state actor, utilizes student assignment policies that are based, in part, on race, such policies are subject to strict

130. *Id.*

131. The Court did conclude, however, that with regard to the traditional school assignments in which African-American and white students are placed on separate assignment tracks, the narrowly-tailored requirement was not met; therefore, the Board's use of race was constitutionally impermissible. *See id.* at 862-64. The Court of Appeals affirmed the district court's judgment without issuing a detailed written opinion. *See* McFarland *ex rel.* McFarland v. Jefferson County Pub. Sch., 416 F.3d 513 (6th Cir. 2005), *cert. granted sub nom.* Meredith *ex rel.* McDonald v. Jefferson County Bd. of Educ., 126 S. Ct. 2351 (U.S. June 5, 2006) (No. 05-915).

132. *See* Parents Involved in Cmty. Sch. v. Seattle Sch. Dist. No. 1 (*Parents I*), 377 F.3d 949, 954-55, (9th Cir. 2004), *rev'd en banc*, 426 F.3d 1162 (9th Cir. 2005), *cert. granted*, 126 S. Ct. 2351 (U.S. June 5, 2006) (No. 05-908).

133. *Id.*

134. A school is considered to be "oversubscribed" "when more students want to attend that school than there are spaces available." *See id.* at 955.

135. *See id.*

136. *See id.* at 954.

137. *See id.* at 955.

138. *See id.* at 955-56.

139. *See id.* at 955 n.7.

scrutiny, and thus, must employ "narrowly tailored measures that further compelling governmental interests."[140]

In *Parents I*, the Ninth Circuit found that the racial tiebreaker program did not pass constitutional scrutiny. While the court recognized the pursuit of educational and societal benefits that accompany racially diverse learning environments as a compelling interest,[141] it found that the racial tiebreaker was not narrowly tailored to further such interest.[142] Upon rehearing *en banc*, the Ninth Circuit in *Parents II* sanctioned the use of the racial tiebreaker and found that the measure was narrowly tailored to further the District's compelling interest in achieving racially and ethnically diverse student bodies.[143] Similar to the district court in *McFarland*, the court also recognized another compelling interest—"ameliorating racial isolation or concentration in . . . high schools by ensuring . . . [student] assignments do not simply replicate . . . segregated housing patterns."[144]

Both courts in *Parents I* and *Parents II* agreed that "one compelling reason for considering race is to achieve the educational benefits of diversity."[145] Both courts found that the District's educational goals complied with the constitutionally permissible diversity rationale as set forth by the Supreme Court in *Grutter*.[146] In so doing, the court in *Parents I* alluded to the prevention of racial isolation as a permissible goal,[147] while *Parents II* directly held that "ameliorating real, identifiable *de facto* racial segregation" is a separate compelling interest.[148]

Although the Supreme Court has never recognized the elimination of *de facto* racial segregation as a compelling interest,[149] other lower

---

140. *Id.* at 960 (quoting Gratz v. Bollinger, 539 U.S. 244, 270 (2003)). *Contra Parents II*, 426 F.3d at 1194 (Kozinski, J., concurring) (advocating a rational basis standard of review "because the Seattle plan carries none of the baggage the Supreme Court has found objectionable in cases where it has applied strict scrutiny and narrow tailoring").

141. *See Parents I*, 377 F.3d at 964.

142. *See id.* at 969.

143. *See Parents II*, 426 F.3d at 1166.

144. *Id*; *see also* James E. Ryan, *Voluntary Integration: Asking the Right Questions*, 67 OHIO ST. L.J. 327, 334 (2006) (formulating the constitutional issue related to voluntary race-conscious student assignment plans as "whether [public schools] have a compelling interest in creating or maintaining a racially integrated student body").

145. *Parents II*, 426 F.3d at 1173; *see also Parents I*, 377 F.3d at 964.

146. *See Parents I*, 377 F.3d at 962 (discussing the Supreme Court's sanctioning of the diversity rationale in *Grutter*); *id.* at 963 ("[E]ach of the School District's proffered interests in using its racial tiebreaker falls comfortably within the diversity rationale as . . . articulated to (and embraced by) the Court."); *see also Parents II*, 426 F.3d at 1173 (describing *Grutter*'s compelling interest as "the promotion of the specific educational and societal benefits that flow from diversity").

147. *See Parents I*, 377 F.3d at 963.

148. *See Parents II*, 426 F.3d at 1178-79 (emphasis added).

149. *See, e.g.*, Freeman v. Pitts, 503 U.S. 467, 494 (1992) ("Racial balance is not to be achieved for its own sake. It is to be pursued when racial imbalance has been caused by a *constitutional violation*." (emphasis added)).

courts have.[150] In advocating for a new compelling interest for using race in an education context, the Ninth Circuit employed the following reasoning:

> The benefits that flow from integration (or desegregation) exist whether or not a state actor was responsible for the earlier racial isolation. *Brown*'s statement that "in the field of public education . . . [s]eparate educational facilities are inherently unequal" retains its validity today. The District is entitled to seek the benefits of racial integration and avoid the harms of segregation even in the absence of a court order deeming it a violator of the U.S. Constitution.[151]

The court also relied on the Supreme Court's school desegregation jurisprudence to justify its sanctioning of school districts' voluntary race-conscious integration efforts.[152]

Unlike the three-judge panel in *Parents I*, the *Parents II* Court held that the race-conscious student assignment plan used by the District was narrowly tailored to achieve its compelling interests. The contrary holdings may be due, in part, to the differing narrowly-tailored tests utilized by the courts. *Parents I* identified and applied the following six narrowly-tailored requirements: (1) prohibition of racial quotas; (2) flexible, individualized consideration of each applicant; (3) prohibition of mechanical or conclusive consideration of race; (4) earnest consideration of race-neutral alternatives; (5) minimization of adverse impact on non-preferred group members; and (6) time limitation.[153] *Parents II*, however, identified the following five factors and only applied factors two through five: "(1) individualized consideration of applicants; (2) the absence of quotas; (3) serious, good-faith consideration of race-neutral alternatives to the affirmative action program; (4) that no member of any racial group was unduly harmed; and (5) that the program had a sunset provision or some other end point."[154]

In finding the individualized consideration factor inapplicable to the District's plan, the Ninth Circuit relied heavily on the different contexts of higher education admissions and secondary education assignments.[155] The court argued that the protections afforded by individualized consideration in a competitive university admission context are not relevant in a

150. *See Parents II*, 426 F.3d at 1178 (citing district and appellate court decisions holding that the creation and maintenance of desegregated schools serve compelling governmental interests).

151. *Id.* at 1179 (alteration in original) (citation omitted).

152. *See id.*

153. *See Parents I*, 377 F.3d at 968-69.

154. *Parents II*, 426 F.3d at 1180.

155. *See* Ryan, *supra* note 144, at 335-36, 339 (arguing that the narrow tailoring test must be formulated in light of the context in which race is used).

non-competitive student assignment context.[156] The Supreme Court in *Grutter* and *Gratz* employed this requirement "in order to prevent race from being used as a mechanical proxy for an applicant's qualifications."[157] As asserted by the Ninth Circuit, the requirement is unnecessary in the present case because students' qualifications are unrelated to their assignment to a particular school.[158] If students' qualifications, such as performance on standardized tests, grades, and artistic and athletic abilities, are not factors in student assignment decisions, then a holistic, individualized review or consideration of such factors is not necessary.[159]

The court also argued that the differences in compelling interests advanced by universities and elementary and secondary schools warrant the non-application of individualized review.[160] While the use of race in both contexts seeks to obtain the social and educational benefits of diversity, the university context lacks the second compelling interest that is present in the high school context, which is preventing the replication of segregated housing patterns in public education.[161] "Because race itself is the relevant consideration when attempting to ameliorate *de facto* segregation, the District's tiebreaker must necessarily focus on the race of its students."[162] In the court's opinion, to require school districts to focus on attributes other than race, such as leadership potential, grades, or life experiences, would undermine their ability to achieve and maintain racially integrated schools.

The court in *Parents I* did not appear to address the different contexts of higher and secondary education as they relate to the individualized consideration requirement. They merely recognized the requirement as a narrowly-tailored factor and applied it to the case. In so doing, the court found that instead of considering several different factors to determine student assignment (as constitutionally mandated in *Grutter* and *Gratz*), the racial tiebreaker "automatically and mechanically admits . . . [and denies] hundreds of white and non-white applicants *solely* because of their race."[163] The court concluded that such operation fails the narrowly-tailored test as set forth in *Grutter* by establishing a "*de jure* [pol-

156. *See Parents II*, 426 F.3d at 1180-81; *see also* Ryan, *supra* note 144, at 335-36, 339-44 (arguing that given the different context of employing non-merit based, non-competitive race-conscious assignment plans, public schools should not be required to give individualized consideration to each student).

157. *Parents II*, 426 F.3d at 1181.

158. *See id.*

159. *See id*; *see also* Holmes, *supra* note 66, at 595-96 (asserting similar arguments regarding the inapplicability of *Grutter*'s individualized consideration requirement to "non-merit-based race-conscious student assignment" programs).

160. *See Parents II*, 426 F.3d at 1183.

161. *See id.*

162. *Id.* (emphasis added).

163. *See* Parents Involved in Cmty. Sch. v. Seattle Sch. Dist. No. 1 (*Parents I*), 377 F.3d 949, 969 (9th Cir. 2004) (emphasis added), *rev'd en banc*, 426 F.3d 1162 (9th Cir. 2005), *cert. granted*, 126 S. Ct. 2351 (U.S. June 5, 2006) (No. 05-908).

icy] of automatic acceptance or rejection based on a[ ] single 'soft' variable."[164] As demonstrated by the conflicting holdings in *Parents I* and *Parents II*, the Supreme Court's formulation of the compelling interests (if any) and the narrowly tailored requirements to advance such interests will have a significant impact on its findings regarding the constitutionality of voluntary race-conscious student assignment plans.

III. A GLIMPSE INSIDE THE COURT'S CRYSTAL BALL: THE BLEAK FUTURE FOR RACE-CONSCIOUS STUDENT ASSIGNMENT PLANS

When one considers the importance of the issues raised in *Parents Involved* and *Meredith* and their potential impact on the provision of educational opportunities to minority students, it is clear that the decisions will significantly contribute to the jurisprudence concerning public education in this country. In determining the constitutionality of race-conscious student assignment plans, the Supreme Court will either sanction or prohibit school districts' use of race as a means to create and maintain racially diverse learning environments. Unfortunately, the Court's reasoning and holdings in previous cases involving the use of race in education present difficult and, in all likelihood, insurmountable challenges to the sanctioning of voluntary race-conscious student assignment plans as employed in the cases at bar.

In assessing the constitutionality of voluntary race-conscious student assignment plans, the Supreme Court must first determine whether the plans serve a compelling interest.[165] Although the Court has never provided a precise definition of what constitutes a "compelling interest,"[166] the term is generally assumed to refer to those interests that are "of the highest order," "overriding," or "unusually important."[167] To date, the Court has recognized two compelling interests that justify the government's constitutional use of race: (1) to remedy past discrimina-

164. *Id.* at 970 (alterations in original) (quoting Grutter v. Bollinger, 539 U.S. 306, 337 (2003)).

165. According to the Supreme Court's holding in *Adarand Constructors, Inc. v. Peña*, all government imposed racial classifications "must be analyzed by a reviewing court under strict scrutiny." *See* Adarand Constructors, Inc. v. Peña, 515 U.S. 200, 227 (1995). Therefore, the school districts' race-conscious student assignment plans "are constitutional only if they are narrowly tailored measures that further compelling governmental interests." *Id.* For a contrary view regarding the appropriate standard of review, *Parents II*, 426 F.3d at 1194 (Kozinski, J., concurring) (advocating a rational basis standard of review "[b]ecause the Seattle plan carries none of the baggage the Supreme Court has found objectionable in cases where it has applied strict scrutiny and narrow tailoring"), *cert. granted*, 126 S. Ct. 2351 (U.S. June 5, 2006) (No. 05-908).

166. *See* Thomas R. Bender, *Does the Right to Trial by Jury Place Constitutional Limits on Prejudgment Interest?*, 39 SUFFOLK U. L. REV. 935, 950-51 (2006).

167. *Id.* at 950; *see also* McFarland *ex rel.* McFarland v. Jefferson County Pub. Sch., 330 F. Supp. 2d 834, 850 (W.D. Ky. 2004) ("Whether an asserted interest is truly compelling is revealed only by assessing the objective validity of the goal, its importance to [the government actor] and the sincerity of [the government actor's] interest."), *aff'd*, 416 F.3d 513 (6th Cir. 2005), *cert. granted sub nom.* Meredith *ex rel.* McDonald v. Jefferson County Bd. of Educ., 126 S. Ct. 2351 (U.S. June 5, 2006) (No. 05-915).

tion;[168] and (2) to achieve student body diversity in higher education.[169] The school districts in *Meredith* and *Parents Involved* ask the Court to recognize a third—to achieve and maintain racially integrated elementary and secondary schools.[170] Considering the Court's prior discussions and holdings regarding government's remedial authority in the context of *de facto* segregation and its prohibition against racial balancing, it is unlikely that it will "expand[ ] the range of permissible uses of race"[171] to include the creation and maintenance of racially diverse public schools. Even if the school districts succeed in demonstrating a compelling interest, the Court will likely prohibit their continued use of race under the challenged plans due to their failure to meet narrowly-tailored requirements.

A. De Jure vs. De Facto Segregation

Directly addressing the constitutionality of the voluntary use of race to remedy *de facto* segregation in public education will be a case of first impression for the Court.[172] The Court, however, has had previous opportunities to consider the use of race to remedy *de jure* segregation in the educational context.[173] In its desegregation jurisprudence, the Court has permitted school districts to employ race-conscious measures in their attempts to eliminate unconstitutional dual educational systems.[174] The measures, however, were restricted to circumstances in which schools' student bodies and faculties were racially imbalanced as a result of the districts' intentional discrimination. Such circumstances do not exist in *Parents Involved* and *Meredith*.

As previously discussed, the District in *Parents Involved* has never experienced legal segregation and, therefore, has never been subject to a desegregation decree.[175] The District's use of race does not seek to remedy the effects of intentional discrimination but rather to prevent racial imbalance that would result from student assignments based on racially segregated housing patterns. The same is true for the Board's utilization of race in *McFarland*.

---

168. *See* Wygant v. Jackson Bd. of Educ., 476 U.S. 267, 274 (1986).

169. *See Grutter*, 539 U.S. at 325.

170. *See* Brief in Opposition at 11-13, *Meredith*, 126 S. Ct. 2351 (No. 05-915), 2006 WL 448513; Brief in Opposition at 16, Parents Involved in Community Sch., 126 S. Ct. 2351 (U.S. June 5, 2006) (No. 05-908), 2006 WL 789611.

171. *Grutter*, 539 U.S. at 357 (Thomas, J., concurring in part and dissenting in part).

172. *See Parents II*, 426 F.3d at 1173 (noting that "the Supreme Court has never decided a case involving the consideration of race in a voluntarily imposed school assignment plan intended to promote racially and ethnically diverse secondary schools").

173. *See* Swann v. Charlotte-Mecklenburg Bd. of Educ., 402 U.S. 1, 15 (1971); United States v. Montgomery County Bd. of Educ., 395 U.S. 225, 225 (1969).

174. *See Montgomery County Bd. of Educ.*, 395 U.S. at 234-36 (sanctioning the establishment of racial ratios for school faculties as a desegregation measure); *see also Swann*, 402 U.S. at 25 (permitting the use of racial mathematical ratios to ensure student body diversity).

175. *See Parents I*, 377 F.3d at 954.

Although the Board had previously been subject to a desegregation decree, the decree was dissolved in June 2000, ten months prior to the Board's adoption of the race-conscious student assignment plan.[176] To justify the dissolution of the decree, the district court found that "[t]o the greatest extent practicable, the Decree has eliminated the vestiges associated with the former policy of segregation and its pernicious effects."[177] Therefore, arguably, the Board's use of racial guidelines is not necessary to eliminate vestiges of racial discrimination since such effects have been deemed to already have been eliminated. Instead, the Board utilizes the racial guidelines to maintain the racially integrated schools created under the desegregation decree.

As noted by the district court responsible for lifting the decree in *McFarland*, student assignment racial guidelines and ratios "[a]re shielded from normal constitutional scrutiny" if employed under a federally mandated desegregation order.[178] Due to school districts' blatant disregard for the Supreme Court's mandate to desegregate, there existed an urgent need for courts to take an active role in directing desegregation efforts.[179] Within this role, courts issued various desegregation mandates, and school districts implemented various policies and programs in their efforts to comply with such mandates.[180] Even though "voluntary school integration" may be viewed "as an extension of the Supreme Court's school desegregation jurisprudence,"[181] it does not necessarily follow that policies implemented under the legal protection of a desegregation decree will survive constitutional scrutiny once the order has been lifted.[182]

As argued in *Parents Involved* and *Meredith*, the context in which state actors use race and ethnicity is extremely important when determin-

176. *See McFarland*, 330 F. Supp. 2d at 841.

177. Hampton v. Jefferson County Bd. of Educ. (*Hampton II*), 102 F. Supp. 2d 358, 360 (W.D. Ky. 2000).

178. *Hampton II*, 102 F. Supp. 2d at 377; *see also* Hampton v. Jefferson County Bd. of Educ. (*Hampton I*) 72 F. Supp. 2d 753, 777 (W.D. Ky. 1999) ("When the Board acts pursuant to the continuing Decree, it acts lawfully.").

179. *See* Freeman v. Pitts, 503 U.S. 467, 471-72 (1992) (acknowledging school districts' delay in complying with Brown v. Bd. of Educ. (*Brown I*), 347 U.S. 483 (1954), and Brown v. Bd. of Educ. (*Brown II*), 349 U.S. 294 (1955), desegregation mandates); *Brown II*, 349 U.S. at 301 (instructing district courts to enter desegregation decrees to require schools to desegregate "with all deliberate speed").

180. *See, e.g.*, Green v. County Sch. Bd. of New Kent County, 391 U.S. 430, 437-38 (1968); *Swann*, 402 U.S. at 25.

181. *McFarland*, 330 F. Supp. 2d at 851; *see also Parents II*, 426 F.3d at 1179 (concluding that the Supreme Court's reference to "the *voluntary integration* of schools as sound educational policy within the discretion of local school officials" supports the Court's finding that "[t]he District is entitled to seek the benefits of racial integration and avoid the harms of segregation even in the absence of a court order deeming it a violator of the U.S. Constitution" (emphasis in original)).

182. *See, e.g.*, *Hampton II*, 102 F. Supp. 2d at 379, 381 (holding that the Board's race-conscious magnet school student assignment plan that had previously been permissible under the desegregation decree was not narrowly tailored to achieve a compelling governmental interest).

ing the constitutionality of their usage.[183] Just as the benefits attained by using race in elementary and secondary education may differ from those attained from using race in higher education, the necessity of racial considerations in federally mandated student assignment plans may differ from the necessity of such considerations in voluntary plans. In *Jenkins*, the Supreme Court clarified that its pronouncement in *Brown I* "was tied *purely* to *de jure* segregation, not *de facto* segregation."[184] Because states had intentionally required Blacks to attend separate, inferior schools, states had an affirmative duty to implement those measures that would effectively eliminate dual educational systems.[185] The Court found that measures involving racial guidelines and ratios were necessary to remedy the harms caused by *de jure* segregation.[186] Once states had practically eliminated the harms associated with *de jure* segregation, the Court held that desegregation duties had been fulfilled since "mere *de facto* segregation (unaccompanied by discriminatory inequalities in educational resources) does not constitute a continuing harm after the end of *de jure* segregation."[187] In *Freeman*, the Court further clarified that with regard to its jurisprudence concerning the imposition of "'awkward,' 'inconvenient,' and 'even bizarre' measures to achieve racial balance in student assignment," such measures were reserved to the context of *de jure* segregation, not phases "when the imbalance is attributable . . . to independent demographic forces."[188]

The current Supreme Court may rely on this rationale to find that the elimination of racial isolation attributable to *de facto* segregation in public schools does not justify the use of racial guidelines and tiebreakers in voluntary student assignment plans. In its reluctance to expand the justifications for the voluntary use of racial classifications, the Court may confine such race-based measures to the context of *de jure* segregation, which, as previously discussed, is inapplicable in the present cases.[189]

B. Racial Balancing

Despite the various contexts in which race and ethnicity have been employed to achieve governmental interests, the Supreme Court has routinely rejected voluntary racial balancing as a permissible interest to jus-

---

183. *See Parents II*, 426 F.3d at 1173 ("[C]ontext matters when reviewing race-based governmental action under the Equal Protection Clause." (quoting *Grutter*, 539 U.S. at 326)); *McFarland*, 330 F. Supp. 2d at 849 ("The different context 'matters' because, under the Equal Protection Clause, 'not every decision influenced by race is equally objectionable.'" (quoting *Grutter*, 539 U.S. at 327)).

184. Missouri v. Jenkins, 515 U.S. 70, 120 (1995) (Thomas, J., concurring) (emphasis added).

185. *See Green*, 391 U.S. at 437-38.

186. *See supra* note 168.

187. *Jenkins*, 515 U.S. at 122 (Thomas, J., concurring).

188. *Freeman*, 503 U.S. at 493.

189. *See Parents II*, 426 F.3d at 1208 n.17 (Bea, J., dissenting) (noting that the Supreme Court's desegregation jurisprudence sanctions the use of race "to combat past *de jure* segregation," not "to achieve racial balance absent *de jure* segregation").

tify their usage.[190] In rejecting racial balancing "for its own sake," the Court in *Freeman* limited its pursuit to those circumstances in which "racial imbalance has been caused by a constitutional violation."[191] Considering the arguments advanced by the petitioners in *Parents Involved* and *Meredith*,[192] it is apparent that both school districts will have to overcome the Court's prohibition against racial balancing to sustain their utilization of race-conscious student assignment plans.

In *Grutter*, the Supreme Court attempted to distinguish between racial balancing and the pursuit of a "critical mass" of minority students. According to the Court, a school's attempt "to assure within [a] student body some specified percentage of a particular group merely because of its race or ethnic origin" amounts to unconstitutional racial balancing.[193] If, however, a school defines its diversity pursuits "by reference to the educational benefits that diversity is designed to produce," then such pursuits may be constitutionally permissible.[194] The respondents in *Parents Involved* and *Meredith* argue that their race-conscious plans satisfy this test.

The respondents in *Parents Involved* argue that the District's plan, including the integration tiebreaker, does not amount to racial balancing because it does "not seek to achieve a pre-determined racial distribution in any school," as proscribed by the Constitution.[195] Rather, the plan seeks to afford white and minority students the opportunity to attend popular schools that may not be close to their neighborhoods.[196] Similarly, the respondents in *Meredith* also argue that their use of racial

---

190. *See Grutter*, 539 U.S. at 330 ("[O]utright racial balancing . . . is patently unconstitutional."); Regents of the Univ. of Cal. v. Bakke, 438 U.S. 265, 307 (1978) (rejecting racial balancing as facially invalid).

191. *Freeman*, 503 U.S. at 494.

192. For example, one of the questions presented by the petitioner in *Parents Involved* asks the following:

> May a school district that is not racially segregated and that normally permits a student to attend any high school of her choosing deny a child admission to her chosen school solely because of her race in an effort to achieve a desired *racial balance* in particular schools, or does such *racial balancing* violate the Equal Protection Clause of the Fourteenth Amendment?

Petition for a Writ of Certiorari at i, Parents Involved in Cmty. Schs. v. Seattle Sch. Dist. No. 1, 126 S. Ct. 2351 (2006) (No. 05-908), 2006 WL 1579631 (emphasis added); *see also* Brief of Petitioner at 5, Meredith v. Jefferson County Bd. of Educ., 126 S. Ct. 2351 (2006) (No. 05-915), 2006 WL 2433475 (arguing that the Board's imposition of racial guidelines "is simply an action for the sake of reflecting racial distribution").

193. *Grutter*, 539 U.S. at 329-30.

194. *Id.* at 330. The Majority's proffered distinction drew much disagreement from other Justices. *See, e.g.*, *id.* at 355 (Thomas, J., concurring in part and dissenting in part) (questioning how the Law School's interest in educational benefits is not racial balancing considering the Law School's apparent belief "that only a racially mixed student body can lead to the educational benefits it seeks"); *id.* at 379, 383 (Rehnquist, C.J., dissenting) (arguing that "[s]tripped of its 'critical mass' veil, the Law School's program is revealed as a naked effort to achieve racial balancing" due to its precise attention to numbers when making admissions decisions).

195. Brief in Opposition at 17, *Parents Involved*, 126 S. Ct. 2351 (No. 05-908), 2006 WL 789611.

196. *Id.*

guidelines in student assignments is not motivated by constitutionally impermissible interests.[197] Rather, the guidelines are used to promote the Board's good faith interest in maintaining racial integration in its schools and the educational benefits that flow from such environments.[198] The district court agreed with this argument and relied on the fact that the Board had "precisely described the academic, social and institutional benefits it achieves from integrated schools" to demonstrate that it had not implemented the racial guidelines to achieve racial balancing "merely for its own sake."[199] This argument, however, fails to adequately address the potentially defeating counterargument that the 15%-50% racial guidelines are mechanical *mandates* intended to assure a *specified* percentage of African-American students in each school.[200] Such racial mandates, which could be termed "quotas," are absolutely proscribed by the Constitution.[201]

As defined by the Supreme Court:

> Quotas "impose a fixed number or percentage which must be attained, or which cannot be exceeded," and "insulate the individual from comparison with all other candidates for the available seats." In contrast, "a permissible goal . . . requires only a good-faith effort . . . to come within a range demarcated by the goal itself," and permits consideration of race as a "plus" factor in any given case while still ensuring that each candidate "competes with all other qualified applicants."[202]

The attainment of a student body that is composed of no fewer than 15% and no more than 50% African-American students is not a "goal" that the Board strives to achieve. Rather, it is a fixed percentage with which schools are *required* to seek compliance.[203] The respondents, in fact, state that "[t]he Plan provides that each school (except preschools, kindergartens, alternative and special schools, and the four exempted magnet schools) *shall* have not less than 15% and not more than 50% black students."[204] Including such directive does not appear to comport with the Supreme Court's sanctioning of the use of race-conscious measures in public education.

Considering that both student assignment plans seek to create and maintain racially balanced schools, both are vulnerable to the Court's proscription of unconstitutional racial balancing. Now that Justice O'Connor, the drafter of the *Grutter* majority, is no longer on the bench,

197. *See* Brief in Opposition at 14, *Meredith*, 126 S. Ct. 2351 (No. 05-915), 2006 WL 448513.
198. *See id.*
199. *McFarland*, 330 F. Supp. 2d at 855.
200. *See supra* notes 125-26 and accompanying text.
201. *See Grutter*, 539 U.S. at 334; *Bakke*, 438 U.S. at 315.
202. *Grutter*, 539 U.S. at 335 (citations omitted).
203. *See supra* note 126.
204. Brief in Opposition at 3-4, *Meredith*, 126 S. Ct. 2351 (No. 05-915) (emphasis added).

it is not apparent that the current members of the Court will accept the racial balancing test as articulated by the majority in *Grutter*. Rather, the Court may employ a more exacting standard to ensure that the interests motivating the utilization of voluntary race-conscious plans are constitutionally permissible.

C. Individualized Consideration

A final impediment to the constitutionality of the race-conscious plans is their failure to meet narrowly-tailored requirements. As required by the standard of review set forth in *Grutter*, all admissions plans that use racial classifications must be narrowly tailored to further compelling interests.[205] Constitutional race-conscious admissions plans are "flexible enough to consider all pertinent elements of diversity" and "ensure that each applicant is evaluated as an individual and not in a way that makes race or ethnicity the defining feature of the application." [206] Unfortunately, the race-conscious student assignment plans utilized in *Parents Involved* and *Meredith* fail both criteria.

Although the Ninth Circuit held that the non-competitive context of elementary and secondary education does not require individualized review,[207] it is doubtful that the Supreme Court will adopt a similar view. While it is true that "context matters when reviewing" race-based measures;[208] the context of elementary and secondary education does not warrant the inapplicability of individualized consideration. Rather, it is, perhaps, the most pertinent context that necessitates individualized review.

All racial classifications are subject to strict scrutiny to guard against the infringement of personal rights guaranteed by the Equal Protection Clause of the Constitution.[209] Strict scrutiny is necessary to protect individuals from the potential stigmatic harms imposed by group-based racial classifications.[210] More so than in other contexts, such protections must be afforded to children in elementary and secondary education. There is, perhaps, no other more necessary context for such protections than elementary and secondary education. The potential harms that can result from telling a child that he or she cannot attend a particular school because he or she is of the wrong race are immeasurable. "Harms such as promotion of racial inferiority, strengthening of racial stereotypes, [and] heightening of racial hostility"[211] are precisely those harms

205. *See Grutter*, 539 U.S. at 308.
206. *Id.* at 309 (citation omitted).
207. *See Parents II*, 426 F.3d at 1183.
208. *Grutter*, 539 U.S. at 308.
209. *See* Adarand Constructors, Inc. v. Peña, 515 U.S. 200, 227 (1995).
210. *See* Richmond v. J.A. Croson Co., 488 U.S. 469, 493 (1989).
211. Nelson, *supra* note 28, at 38.

that the Court's desegregation cases attempted to remedy.[212] It is, therefore, highly improbable that the current Supreme Court would permit the use of racial classifications in elementary and secondary education without requiring that they meet *every* element of strict scrutiny.

Contrary to the narrowly-tailored criteria set forth in *Grutter*, the student assignment plans in question do not afford meaningful consideration to diversity elements other than race and ethnicity. The district court in *McFarland* found that the Board's plan is constitutional because it considers other diversity factors "such as place of residence and student choice of school or program."[213] Such argument cannot sustain the constitutionality of the plan because the operation of the plan is such that these factors are effectively irrelevant if a student attempts to enroll in a school where the racial composition will fall outside the racial guidelines if he or she is admitted. Despite the student's other "diversity factors," he or she will most likely be denied admission.[214] The racial tiebreaker employed in *Parents Involved* operates in a similar manner. Depending on the racial makeup of a particular school to which a student is applying for admission, his or her race can be the determinative factor in deciding whether he or she is admitted or denied.[215] In both plans, race operates as the defining and decisive feature of a student's application, not as a constitutionally permissible "plus" factor.[216] Therefore, the plans are not narrowly tailored and, thus, cannot pass constitutional scrutiny.

CONCLUSION: FULFILLING *BROWN*'S MANDATE

In assessing the constitutionality of voluntary race-conscious student assignment plans in the context of *de facto* racial isolation in elementary and secondary schools, the Supreme Court will be guided by its previous holdings and rationales. As it attempts to balance the proffered interests in creating and maintaining racial integration against the constitutional protections provided by the Equal Protection Clause of the Constitution, the Court will be guided by the principle that "[t]he Constitution does not prevent individuals from choosing to live together, to work

---

212. *See, e.g.*, *Brown I*, 347 U.S. at 494 (noting that legally sanctioned racial segregation produces feelings of inferiority, which detrimentally "affects the motivation of a child to learn").

213. *McFarland*, 330 F. Supp. 2d at 859.

214. As noted by the district court:

> [W]here the racial composition of an entire school lies near either end of the racial guidelines, the application of any student for open enrollment, transfer or even to a magnet program could be affected. In a specific case, a student's race, whether Black or White, could determine whether that student receives his or her first, second, third or fourth choice of school.

Id. at 842.

215. *See Parents I*, 377 F.3d at 955-56 (explaining that the racial tiebreaker operates to grant automatic admission to students who are of the preferred race needed to help schools attain the desired racial ratio of white and minority students).

216. *See* Gratz v. Bollinger, 539 U.S. 244, 271-72 (2003) (invalidating a race-conscious admissions policy because of its use of race as the decisive factor in an admissions decision rather than as a "plus" factor along with many different diversity criteria).

together, or to send their children to school together, *so long as the State does not interfere with their choices on the basis of race.*"[217] Once the inquiry has been completed, the challenged plans will most likely be invalidated. In light of this probable outcome, local, state, and federal officials should immediately engage in the development of race-neutral programs and policies that can effectively address the harmful effects of resegregation of public schools.

School officials should not retard the progress that has been made in the provision of educational opportunities to minority students by paving "a one-way street" to racially and economically segregated neighborhood schools.[218] As previously discussed, students attending such schools face challenges, which are often insurmountable, that range from less qualified teachers[219] to a culture of lower academic expectations.[220] To combat these challenges, schools should employ race-neutral student assignment plans[221] and implement educational policies that effectively address deficiencies in the provision of equal educational opportunities to minority students.

Some schools have already begun to experiment with race-neutral measures in their efforts to achieve racially diverse student bodies.[222] Such measures include the consideration of "diversity in student achievement" and "diversity in socioeconomic status."[223] Limiting concentrations of low-performing students in schools will impact student body diversity since minority students often perform lower than their white counterparts on academic measures.[224] Similarly, assigning students to schools based on their socioeconomic status can also achieve racial diversity because of the existing racial gaps in socioeconomic status.[225] Such "class-based" assignment plans are also beneficial be-

---

217. *Jenkins*, 515 U.S. at 121 (emphasis added).

218. *Hampton II*, 102 F. Supp. 2d at 379.

219. *See supra* notes 70-71, 93-99 and accompanying text.

220. *See supra* note 100 and accompanying text.

221. In the context of student assignments, "race-neutral" refers to those plans that do not classify students based on their race or ethnicity. Such plans are not "race-blind" in that they ignore the effects of race on educational opportunities. They simply do not consider a student's race when assigning him or her to a particular school. *See* Nelson, *supra* note 28, at 7-11 (discussing the meaning of "race-neutral" alternatives in the context of higher education admissions decisions).

222. *See, e.g.*, Boger, *supra* note 82, at 1397-1400 (discussing the implementation of race-neutral student assignment plans in Wake County, North Carolina).

223. *Id.* at 1397.

224. For example, in 2004, black and Hispanic children age 9, 13 and 17 had lower average reading scale scores than white students. *See* NAT'L CTR. FOR EDUC. STATISTICS, DIGEST OF EDUCATION STATISTICS: 2005 tbl.108 (2006) *available at* http://nces.ed.gov/programs/digest/d05/tables/dt05_108.asp. The same was true for their performance in mathematics. *See id.* at tbl.118, *available at* http://nces.ed.gov/programs/digest/d05/tables/dt05_118.asp. In 2001, the average geography and U.S. history scores for white students were higher than those achieved by black and Hispanic students. *See id.* at tbl.116, *available at* http://nces.ed.gov/programs/digest/d05/tables/dt05_116.asp.

225. *See* NAT'L CTR. FOR EDUC. STATISTICS, THE CONDITION OF EDUCATION 2006 tbl.6-1 (2006), *available at* http://nces.ed.gov/programs/coe/2006/section1/table.asp?tableID=440 (indicating that 70% of black Fourth graders and 73% of Hispanic Fourth graders are eligible for free or

cause they provide the added benefit of socioeconomic diversity, which may, in fact, be more educationally beneficial than racial diversity.

Some scholars have concluded that "[n]o other single social measure is consistently more strongly related than poverty to school achievement."[226] Consequently, "overall socioeconomic composition of schools seem[ ] more predictive of academic achievement than [does] a student's individual socioeconomic status."[227] If this is true, school officials should direct their attention to achieving and maintaining socioeconomic diversity rather than racial diversity. Presumably, such efforts would not be subject to the heightened and, potentially, fatal standard of strict scrutiny because they neither employ racial classifications nor seek to achieve racial diversity benefits.[228] Rather, they seek to achieve the educational benefits of socioeconomic integration.

In their attempts to provide equal educational opportunities to all students, school officials should implement policies to remedy the disparities that currently exist between minority, economically disadvantaged schools and their non-minority economically advantaged counterparts.[229] As often noted by many scholars, "to those who need the best our education system has to offer, we give the least. The least well-trained teachers. The lowest-level curriculum. The oldest books. The least instructional time. Our lowest expectations. Less, indeed, of everything that we believe makes a difference."[230] As previously discussed, one glaring disparity is the level of teacher quality.[231] Students attending high minority, low socioeconomic schools are disproportionately subjected to being taught by less qualified teachers.[232] Such inequitable

---

reduced lunch, as compared to only 24% of white Fourth graders); *see also* Dickerson, *supra* note 98, at 1756-68 (noting significant racial disparities in wealth as shown by levels of home ownership, personal assets and business ownership).

226. Boger, *supra* note 82, at 1416.

227. *Id.* at 1416-17; *see also supra* notes 101-02 and accompanying text; Orfield, *supra* note 87, at 280 (concluding that peer socioeconomic status accounts for more than 75% of the difference between minority and white students' academic achievement).

228. *See* Boger, *supra* note 82, at 1398-99 (concluding that race-neutral student assignment plans should not be subject to strict scrutiny as long as they have not "been adopted as a mere pretext for continuing racial assignments"); *see also* Levit, *supra* note 40, at 511 (encouraging schools to "first try experiments that are more likely to be successful and less likely to be unconstitutional" in their efforts to achieve educational goals).

229. *See* Dickerson, *supra* note 107, at 1291 n.82 suggesting the following:

> School disparities also could be eliminated by increasing the attractiveness of "bad" schools (for example, by giving the school a disproportionate share of new technology, equipment or supplies, addressing its facility maintenance needs before the needs of other schools, allowing smaller classes and student/teacher ratios, by giving the teachers in the school greater flexibility in the classroom, etc.).

230. *See* Susan P. Leviton & Matthew H. Joseph, *An Adequate Education for All Maryland's Children: Morally Right, Economically Necessary, and Constitutionally Required*, 52 MD. L. REV. 1137, 1142 (1993).

231. *See supra* notes 69-70 and accompanying text.

232. *See id*; *see also* Linda Darling-Hammond, *Teacher Quality and Student Achievement: A Review of State Policy Evidence*, 8 EDUC. POL'Y ANALYSIS ARCHIVES 1, ¶ 85 (2000), http://epaa.asu.edu/epaa/v8n1/ (reporting findings that poor minority students are taught by less qualified teachers than their non-minority socially advantaged peers).

learning environments negatively affect not only the quality of education that students receive,[233] but also their psychological well-being by sending and reinforcing messages "that society doesn't care enough about whether they learn."[234]

To combat such debilitating effects, school officials should invest in the quality of their teachers, especially those teaching in lower-performing schools, by implementing initiatives that are designed to improve teacher qualifications and effectiveness, such as pre-service teacher education, mentoring programs, and continual professional development.[235] School officials should also provide incentives to encourage more qualified teachers to teach at lower-performing schools. Such incentives could be immediate, such as salary increases or bonuses,[236] or they could be long-term, such as early retirement opportunities. More qualified teachers may be enticed to teach at high minority, low socio-economic schools if doing so afforded them the opportunity to be eligible for retirement five or ten years earlier than their counterparts teaching at more affluent schools. Coupled with intensive recruitment efforts at the high school and college levels, schools implementing such beneficial policies could see a significant improvement in the quality of their teachers and, consequently, the academic quality of their students.[237]

Implementing race-neutral assignment policies and teacher quality initiatives is merely the beginning in addressing the significant costs imposed by segregated learning environments. To fulfill *Brown*'s mandate of educational equality, economically disadvantaged minority students must have the opportunity to interact with peers from diverse backgrounds to broaden and heighten their educational goals and possibilities.[238] Whether or not the Supreme Court allows schools to facilitate this interaction through the use of race-conscious student assignment plans, our schools and our country have the moral responsibility to ensure that such interaction takes place and that it occurs within educational institutions that provide all students access to equal resources necessary to create and fulfill their academic dreams.

---

233. *See id.* (concluding that student outcomes and student achievement are negatively affected by poor teacher quality).

234. *See* Jeannie Oakes, *Education Inadequacy, Inequality, and Failed State Policy: A Synthesis of Expert Reports Prepared for* Williams v. State of California 1, 9-10, *available at* http://www.decentschools.org/expert_reports/oakes_report.pdf.

235. *See* Darling-Hammond, *supra* note 232, ¶ 57.

236. *See* Dickerson, *supra* note 107, at 1291 n.82 (proposing the awarding of bonuses to highly qualified teachers as incentives to teach at low-income, minority schools).

237. *See* Darling-Hammond, *supra* note 232, ¶¶ 56-57 (describing significant student achievement gains made in North Carolina and Connecticut following the states' enactment of substantial reforms targeting teacher quality).

238. *See supra* notes 100-01 and accompanying text.

THE RISE, DEVELOPMENT AND FUTURE DIRECTIONS OF CRITICAL RACE THEORY AND RELATED SCHOLARSHIP

ATHENA D. MUTUA†

TABLE OF CONTENTS

---

† Athena D. Mutua is an Associate Professor of Law at the University at Buffalo Law School. I would like to thank Rebecca French, Janet Lindgren, Isabel Marcus, Martha McCluskey, Makau Mutua, Stephanie Phillips, John Schlegel and Francisco Valdes for their astute comments on earlier versions of this paper. I especially would like to thank Stephanie Phillips who, when she discovered I was interested in learning more about critical race theory, took me by the hand, accompanied me to one of the last CRT workshops and initially co-taught with me several CRT seminars at the University at Buffalo. Throughout this process she regaled me with the stories about CRT, its various personalities, many of whom I had met in other contexts, and the twists, turn, and intrigues of its development. It was a lot of fun and I am very grateful for her insights.

This article was originally drafted for the Sage Handbook on Race and Ethnic Studies (Patricia Hill Collins and John Solomos eds., forthcoming), in which a version of it appears.

INTRODUCTION

This essay tells the story of the rise, development and future directions of critical race theory and related scholarship. [1] In telling the story, I suggest that critical race theory (CRT) rises, in part, as a challenge to the emergence of colorblind ideology in law, a major theme of the scholarship.[2] I contend that conflict, as a process of intellectual and institutional growth, marks the development of critical race theory and provides concrete and experiential examples of some of its key insights and themes. These conflicts are waged in various institutional settings over the structural and discursive meanings of race and the role that race plays in society, an argument made in part, by Kimberlé Crenshaw,[3] and a story drawn in parts from her, Stephanie Phillips,[4] and Cheryl Harris.[5]

---

1. *See generally* CRITICAL RACE THEORY: THE KEY WRITINGS THAT FORMED THE MOVEMENT (Kimberlé Crenshaw, Neil Gotanda, Gary Peller & Kendall Thomas eds., 1995) [hereinafter CRT: KEY WRITINGS]; CRITICAL RACE THEORY: THE CUTTING EDGE (Richard Delgado & Jean Stefancic eds., 2d ed. 2000) [hereinafter CRT: CUTTING EDGE]; RACE AND RACES: CASES AND RESOURCES FOR A DIVERSE AMERICA (Juan Perea, Richard Delgado, Angela Harris, & Stephanie Wildman eds., 2000); CRITICAL RACE THEORY: CASES, MATERIALS AND PROBLEMS (Dorothy A. Brown ed., 2003) [hereinafter CRT: CASES]; RICHARD DELGADO & JEAN STEFANCIC, CRITICAL RACE THEORY: AN INTRODUCTION (2001) [hereinafter DELGADO & STEFANCIC, CRT: AN INTRODUCTION]; CROSSROADS, DIRECTIONS, AND A NEW CRITICAL RACE THEORY (Francisco Valdes, Jerome McCristal Culp, & Angela P. Harris eds., 2002); THE LATINO/A CONDITION: A CRITICAL READER (Richard Delgado & Jean Stefancic eds., 1998) [hereinafter THE LATINO/A CONDITION]; FRANK H. WU, YELLOW: RACE IN AMERICA BEYOND BLACK AND WHITE (2002); Keith Aoki, *The Scholarship of Reconstruction and the Politics of Backlash*, 81 IOWA L. REV. 1467 (1996); Robert S. Chang, *Toward an Asian American Legal Scholarship: Critical Race Theory, Post-Structuralism, and Narrative Space*, 81 CAL. L. REV. 1241 (1993), *reprinted in* CRT: CUTTING EDGE, *supra*, at 354-68 and *reprinted in part in* POWER, PRIVILEGE AND LAW: A CIVIL RIGHTS READER 211-15 (Leslie Bender & Daan Braveman eds., 1995) [hereinafter POWER, PRIVILEGE AND LAW]; Cheryl I. Harris, *Critical Race Studies: An Introduction*, 49 UCLA L. REV. 1215 (2002); Natsu Taylor Saito, *Alien and Non-Alien Alike: Citizenship, "Foreignness," and Racial Hierarchy in American Law*, 76 OR. L. REV. 261 (1997); Symposium, *Critical Race Theory*, 82 CAL. L. REV. 741 (1994); Symposium, *Critical Race Lawyering: Reopening the Emmett Till Case: Lessons and Challenges for Critical Race Practice*, 73 FORDHAM L. REV. 2101 (2005).

2. *See, e.g.*, Neil Gotanda, *A Critique of "Our Constitution Is Colorblind,"* 44 STAN. L. REV. 1 (1991), *reprinted in* CRT: KEY WRITINGS, *supra* note 1, at 257-75; *see also* DERRICK BELL, RACE, RACISM, AND AMERICAN LAW, 115-35 (5th ed. 2004); DELGADO & STEFANCIC, CRT: AN INTRODUCTION, *supra* note 1, at 21-22; Harris, *supra* note 1, at 1229-30; Tanya Katerí Hernandez, *"Multiracial" Discourse: Racial Classifications in an Era of Color-Blind Jurisprudence*, 57 MD. L. REV. 97 (1998); discussion *infra* notes 25-36, 78-88, 168-203 and accompanying text.

3. CRT: KEY WRITINGS, *supra* note 1, at xiii-xxxii; Kimberlé Williams Crenshaw, *The First Decade: Critical Reflections, or "A Foot in the Closing Door,"* 49 UCLA L. REV. 1343 (2002), *reprinted in* CROSSROADS, DIRECTIONS, AND A NEW CRITICAL RACE THEORY, *supra* note 1, at 9-31.

4. *See generally* Stephanie L. Phillips, *The Convergence of the Critical Race Theory Workshop with LatCrit Theory: A History*, 53 U. MIAMI L. REV. 1247 (1999).

5. *See generally* Harris, *supra* note 1. In telling the story of CRT's development, Crenshaw specifically notes at the outset that CRT did not develop in the "abstract but in the context shaped by

Conflict within CRT in turn, to some extent spurs the development of CRT-related scholarship, such as Asian American Legal Scholarship, Critical Race Feminism and specifically Latino and Latina Critical Schools (LatCrit).[6] Though this related scholarship could be seen as fragmenting the CRT movement,[7] I suggest, focusing primarily on LatCrit, that it has actually deepened it. However, a significant area that CRT has not adequately addressed is the issue of class and its relationship to race and other subordinating structures. I examine reasons why this is the case even though CRT scholars have repeatedly called for analyses of the relationship between race and class and propose critical class analyses or classcrits as a necessary future direction of CRT and related scholarship.[8]

A few caveats are in order. This story could be told in as many different ways as there are CRT theorists.[9] Further, although this rendition presents CRT as a "fully unified school of thought," this is not the case, as CRT remains a work in progress.[10] In addition, while I present what

---

specific institutional struggles over concrete issues that were set in motion by certain individuals." Crenshaw, *supra* note 3, at 9.

6. *See infra* notes, 46-52, 104-16 and accompanying text.

7. *See* Otis T. Bryant, *Factualism Within the Critical Race Theory Movement: Fact or Fiction? A Comparative Analysis of the Emerging LatCrit and AsianCrit Self-Identification Theories and a Continued Critique (and Defense) of the Black/White Binary* (discussing the fragmentation of the movement and suggesting that the differences are small) (manuscript on file with author).

8. Martha McCluskey and I, with support of the Baldy Center for Law and Social Policy at University at Buffalo recently planned a workshop scheduled for January 2007 to explore the possibility of launching a new network of scholars interested in developing a progressive legal economic theory and politics. We explained that the term "ClassCrits'" (taken from an earlier draft of this paper) reflected "our interest in focusing on economics through the lens of critical legal scholarship movements, such as critical legal studies, critical feminist theory, critical race theory, LatCrit, and queer theory. That is, we start[ed] with the assumption that economics in law is inextricably political and fundamentally tied to questions of systemic status-based subordination." We explained our interest and posed a number of questions in an invitation roughly as follows:

> Economic inequality has become a growing problem locally, nationally and internationally. In light of this central social and political reality, it is time to foreground economics in progressive jurisprudence and to reconsider longstanding assumptions and approaches in legal scholarship and practice. We aim to provide an alternative to the predominant discussions of "law and economics" grounded in neoclassical economic theory and its denial of "class." Many legal scholars are now interested in challenging or broadening some of the assumptions of neoclassical economics. Nonetheless, the question of class and the role of institutionalized inequality still lurks beneath the surface of most discussions of economics in legal academia.
>
> Here are a few of the questions we want to explore through a critical analysis of economic inequality. How might a critical class analysis of law and economic inequality build on, differ from or respond to other approaches to analyzing law and economic inequality, such as Law and Economics, poverty law, labor law, socioeconomic and legal realism? How might a focus on class contribute to the debates within critical theory and practice, such as how to address legal and societal subordination? What insights from earlier work in critical legal studies and legal realism might be revived, updated, and improved to better address contemporary concerns and debates? What is the relationship between subordinated identity based on economic class and institutionalized structures of economic subordination? How might the perspectives and insights of critical feminism and critical race theory help develop previous work on economic class in law? How might ClassCrits build upon the idea of anti- subordination praxis and intersectionality? What can we do to build a ClassCrits network?

9. *See* Crenshaw, *supra* note 3, at 9; *see also* Harris, *supra* note 1, at 1218 (making a similar point).

10. Crenshaw, *supra* note 3, at 20.

are commonly agreed to be the "tenets" of CRT,[11] not every precept presented here is held valid by every self-described critical race scholar. Perhaps that is as it should be. And finally, although I provide an overview of Asian American Legal Scholarship, Critical Race Feminism, and LatCrit as CRT-related scholarship, I primarily focus on the thematic and institutional development of "LatCrit" because it in part grows out of internal conflict within CRT and because of its deep institutionalization, especially given the cessation of the CRT workshop.[12]

Parts One through Four of this article present a narrative of the origins and development of Critical Race Theory. Part One provides an overview of CRT and related scholarship and some of its basic themes. Part Two discusses CRT's intellectual antecedents. Part Three discusses a series of four sets of conflicts that I suggest have contributed to its intellectual content and institutional development. These involve conflicts between future critical race scholars with the Harvard Law School administration in the early eighties, later conflicts with critical legal studies, and a debate at the University of California Los Angeles Law School (UCLA) around the issue of affirmative action in the context of the California Proposition 209. They also involve conflicts internal to CRT, one challenging CRT views on sexual oppression and another involving the experiences and perspectives of non-black people of color, the latter in part leading to the establishment of LatCrit. Part Four then lays out CRT's basic tenets and methodological fingerprints.

Part Five builds upon the context developed in the first four sections of the paper and applies critical theory insights and methods to an historical analysis of law and race. In doing so, I seek to provide a counter-narrative, as Derrick Bell suggests,[13] to the dominant story about race and law. The dominant story suggests that the struggle for racial justice, though long and incremental, is nevertheless forward-moving, progressive, and eventually triumphant, given the American creed and precepts.[14] The counter-narrative challenges this, suggesting that racial justice has not triumphed and that the white supremacy of American law first based explicitly on a theory of white superiority and black inferiority continues now under the guise, through the operation, and on a theory of colorblindness. Part Five, thus, seeks to *do* Critical Race Theory. Part Six, summarizes a number of the key insights of related CRT scholarship with a particular focus on LatCrit. While it has been suggested that the development of this related scholarship has fragmented the CRT pro-

11. *See infra* note 131 and accompanying text.

12. *See infra* notes 131-34 and accompanying text.

13. BELL, *supra* note 2, at 21-22; *see also* DELGADO & STEFANCIC, CRT: AN INTRODUCTION, *supra* note 1, at 40-41.

14. BELL, *supra* note 2, at 21-22.

ject,[15] I suggest that these movements and insights necessarily inform the CRT project.

Part Seven briefly takes up an important refrain and critique of CRT: that it more systematically explore the relationship between class and race.[16] This I think requires serious engagement with the notion and structures of class. Drawing on similar work and critiques by Richard Delgado,[17] I explore possible reasons why CRT, which has made significant contributions to feminist legal scholarship,[18] and gay, lesbian, and queer legal scholarship,[19] in addition to race scholarship, has not adequately engaged the issues of class. I suggest class analyses or the founding of classcrits as a future direction of CRT scholarship and advocacy.

I. OVERVIEW

One of the most significant developments in law on issues of race and ethnicity in the last twenty years is the development of Critical Race Theory (CRT) and related scholarship.[20] The name, "Critical Race Theory" was coined in the late 1980's by Kimberlé Crenshaw who explained that the theory represented a racial analysis, intervention and critique of traditional civil rights theory on the one hand, and of Critical Legal Studies insights on the other.[21] Its basic premises are that race and racism are endemic to the American normative order and a pillar of American institutional and community life. Further, it suggests that law does not merely reflect and mediate pre-existing racialized social conflicts and relations.[22] Instead law, as part of the social fabric and the larger hegemonic order, constitutes, constructs and produces races and race relations

15. *See infra* note 117.

16. *See infra* note 247.

17. *See generally* Richard Delgado, *Crossroads and Blind Alleys: A Critical Examination of Recent Writings About Race, Crossroads, Directions, and a New Critical Race Theory*, 82 TEX. L. REV. 121 (2003) [hereinafter Delgado, *Blind Alleys*]; Richard Delgado, *Two Ways to Think About Race: Reflections on the Id, the Ego, and Other Reformist Theories of Equal Protection*, 89 GEO. L.J. 2279 (2001) [hereinafter Delgado, *Two Ways to Think About Race*]; Delgado, *The Current Landscape of Race: Old Targets, New Opportunities*, 104 MICH. L. REV. 1269 (2006) [hereinafter Delgado, *The Current Landscape of Race*]; *infra* notes 247-98.

18. I am thinking, for instance, of intersectional theory as a significant contribution to feminist legal theory. *See* Kimberlé Crenshaw, *Mapping the Margins: Intersectionality, Identity Politics, and Violence Against Women of Color*, 43 STAN. L. REV. 1241 (1991), *reprinted in part in* CRT: KEY WRITINGS, *supra* note 1, at 357-83 [hereinafter Crenshaw, *Mapping the Margins*]; KIMBERLÉ CRENSHAW, A BLACK FEMINIST CRITIQUE OF ANTIDISCRIMINATION LAW AND POLITICS, *in* THE POLITICS OF LAW: A PROGRESSIVE CRITIQUE 356-80 (David Kairys ed., 3d ed. 1998) [hereinafter THE POLITICS OF LAW], *see also infra* notes 221-27 and accompanying text.

19. Here I am thinking about multidimensionality theory as a contribution to gay, lesbian, and queer theory. *See, e.g.*, Darren Lenard Hutchinson, *Ignoring the Sexualization of Race: Heteronormativity, Critical Race Theory and Anti-Racist Politics*, 47 BUFF. L. REV. 1 (1999); Francisco Valdes, *Afterword: Beyond Sexual Orientation in Queer Legal Theory: Majoritarianism, Multidimensionality, and Responsibility in Social Justice Scholarship*, 75 DENV. U. L. REV. 1409 (1998); *see also infra* notes 228-33 and accompanying text.

20. *See generally*, *supra* note 1.

21. Crenshaw, *supra* note 3, at 9-31; CRT: KEY WRITINGS, *supra* note 1, at xiii-xxxii.

22. Harris, *supra* note 1, at 1216-17.

in a way that supports white supremacy.[23] Critical Race Theory, as Cheryl Harris explains, "coheres in the drive to excavate the relationship between the law, legal doctrine, ideology, and [white] racial power and the motivation 'not merely to understand the vexed bond between law and racial power but to change it.'"[24]

Critical Race Theory arose during the ascendance of, and as a challenge to the ideology of colorblindness in law,[25] which asserts that race, like eye color, is and should be irrelevant to the determination of individuals' opportunities. A noble sentiment perhaps, but Critical Race Theory, while maintaining that race is not like eye color,[26] argues that legal colorblindness operates as if a colorblind society already exists and has always existed in the United States.[27] In doing so, it ignores and cements the racial caste system constructed in part by law.[28] In other

23. *Id.* For recent discussions of the construction of racial identity at the personal level, see, e.g., Michelle Adams, *Radical Integration*, 94 CAL. L. REV. 261 (2006); Angela Onyuachi-Willig, *Undercover Other*, 94 CAL. L. REV. 873 (2006).

24. Harris, *supra* note 1, at 1218 (*quoting* CRT: KEY WRITINGS, *supra* note 1, at xiii).

25. *See* CRT: KEY WRITINGS, *supra* note 1, at xvi-xvii (noting that at the emergence of CRT, many of its principle figures experienced the "boundaries of 'acceptable' racial discourse as becoming suddenly narrowed" and explaining that while there were differences between liberal and conservative positions they "defined and constructed 'racism' the same way, [in contrast to future CRT scholars] as the opposite of color-blindness"); *see also* Crenshaw, *supra* note 3, at 22 (noting that CRT came into "existence at the twilight of what had been a transformative social period" referencing the civil rights movement of the sixties and seventies). In retrospect, it becomes clear that CRT's initial conflicts are related to the colorblind perspectives of their adversaries. *See infra* notes 78-92. Further, CRT emerges at the same time that the Supreme Court's interpretations of the Equal Protection Clause in the racial context increasingly shift toward and are based on notions of colorblind individualism. *See infra* notes 30-36, 171-206 and accompanying text. For discussions of colorblindness, see BELL, *supra* note 2, at 115-35; DELGADO & STEFANCIC, CRT: AN INTRODUCTION, *supra* note 1, at 21-22; Gotanda, *supra* note 2, at 257-75. *See generally* MICHAEL K. BROWN ET. AL., WHITEWASHING RACE: THE MYTH OF A COLOR-BLIND SOCIETY 193 (2003) [hereinafter BROWN ET AL., WHITEWASHING RACE] (reviewed, from a CRT perspective by Delgado, *The Current Landscape of Race*, *supra* note 17); Cheryl Harris, *Review Essay: Whitewashing Race: Scapegoating Culture*, 94 CAL. L. REV. 907 (2006); Girardeau A. Spann, *Affirmative Action and Discrimination*, 39 HOW. L.J. 1 (1995); Eric K. Yamamoto, Carly Minyer, & Karen Winter, *Contextual Strict Scrutiny*, 49 HOW. L.J. 241 (2006).

26. Harris, *supra* note 1, at 1229. Critical race theorists would argue that race is socially constructed in a process, in which social and materially relevant meanings are assigned to certain biological traits, whereas eye color is a biological trait to which no overriding social meaning has been assigned. For good treatments of the social construction of race, see BELL, *supra* note 2, at 5-8; Robert S. Chang, *Critiquing "Race" and Its Uses: Critical Race Theory's Uncompleted Argument*, *in* CROSSROADS, DIRECTIONS, AND A NEW CRITICAL RACE THEORY, *supra* note 1, at 87-96; IAN F. HANEY LÒPEZ, WHITE BY LAW: THE LEGAL CONSTRUCTION OF RACE (1996) *reprinted in* CRT: CUTTING EDGE, *supra* note 1, at 626-34; MICHAEL OMI & HOWARD WINANT, RACIAL FORMATION IN THE UNITED STATES FROM THE 1960S TO THE 1990S 9-13 (2d ed., 1994); Gotanda, *supra* note 2, at 257-75; Ian F. Haney López, *The Social Construction of Race: Some Observations on Illusion, Fabrication, and Choice*, 29 HARV. C.R.-C.L. L. REV. 1, 3-10 (1994), *reprinted in part in* CRT: CUTTING EDGE, *supra* note 1, at 191-203. W.E.B. Du Bois anticipated this position. *See The Conservation of Race*, in W.E.B. DU BOIS SPEAKS: SPEECHES AND ADDRESSES, 1890-1919, 50-54 (Philip Foner ed., 1998); *see generally* FRANZ BOAS, RACE, LANGUAGE, AND CULTURE 3-195 (1940) (these pages include multiple essays discussing race).

27. This argument is not meant to suggest that the ideal society in the future would be a colorblind society. An ideal society might be one that embraces and respects human diversity.

28. *See* BROWN ET AL., WHITEWASHING RACE, *supra* note 25, at 193; Gotanda, *supra* note 2, at 274; *see also* BELL, *supra* note 2, at 126-27; DELGADO & STEFANCIC, AN INTRODUCTION *supra* note 1, at 21-22; Harris, *supra* note 1, at 1229-30; Harris, *supra* note 25 (arguing in reference to

words, it maintains the oppressive conditions and lack of opportunities for subordinated groups that continue to be structured by the historical and modern use of race in law and throughout the society. It amounts to what has been called "colorblind racism."[29]

Colorblindness, or more specifically, "colorblind individualism,"[30] as an ideology applied in law, fully emerges after the civil rights movement but has deep historical roots even in American law itself. One of its earliest and clearest articulations is found in the dissent to the infamous U.S. Supreme Court case, *Plessy v. Ferguson*.[31] In *Plessy*, the Court held that the Equal Protection Clause of the fourteenth amendment to the U.S. Constitution, enacted as part of the Civil War amendments that abolished slavery, allowed the separate but equal treatment of racialized people. *Plessy* thus made legal the practices of racial, "Jim Crow" segregation in the United States[32] that the civil rights movement, almost one hundred years later, would challenge. Justice John Marshall Harlan, the lone dissenter in the case, in rejecting the holding asserts the colorblind claim: "[I]n the eye of the law, there is in this country no superior, dominant, ruling class of citizens . . . Our Constitution is color-blind, and neither knows nor tolerates classes among citizens."[33] The separate but equal legal doctrine of *Plessy* was overturned in the *Brown v. Board of*

---

conservative interpretations that race was not a factor in the aftereffects of Hurricane Katrina, that colorblindness is a perspective that narrows the definition of race to that arising from the aberrant individual's intentional action); Hernandez, *supra* note 2. From this colorblind perspective, the response to the question of whether race played a role in the government's slow response to hurricane victims, the argument is no one or group of individuals could be said with certainty to have intentionally delayed governmental action. Of course this does not address more unconscious influences in individuals and it does not explain why, out of the 1.3 million people who make up the New Orleans metropolitan area—all hit by the hurricane—the 120,000 or so people stranded in New Orleans after the hurricane were overwhelmingly black? *See* Virginia R. Dominguez, *Seeing and Not Seeing: Complicity in Surprise*, *in* SOCIAL SCIENCE RESEARCH COUNCIL, UNDERSTANDING KATRINA: PERSPECTIVES FROM THE SOCIAL SCIENCES (2005), http://understandingkatrina.ssrc.org/Dominguez/pf/ (providing these statistics and making a similar point); *see also infra* notes 297-316 (discussing class, race and Hurricane Katrina).

29. *See generally* EDUARDO BONILLA-SILVA, WHITE SUPREMACY AND RACISM IN THE POST-CIVIL RIGHTS ERA 137-66 (2001); LESLIE G. CARR, "COLOR-BLIND" RACISM 107-170 (1997).

30. This term is used to capture both the ideas of race neutrality and the emphasis in American law on the individual. For instance, Kevin Brown uses this term to capture these two different dynamics in discussing the demise of school desegregation efforts in the U. S. *See* KEVIN DION BROWN, RACE, LAW AND EDUCATION IN THE POST-DESEGREGATION ERA: FOUR PERSPECTIVES ON DESEGREGATION AND RESEGREGATION 103-28 (2005); *see also* John O. Calmore, *New Demographics and the Voting Rights Act: Race-Conscious Voting Rights and the New Demography in a Multiracing America*, 79 N.C. L. REV. 1253, 1271-72 (2001) (discussing the way the Supreme Court is transporting "individualized colorblindness" across different context and explaining why this is ridiculous in the voting dilution context, a context that only makes sense in the terms of identifiable groups and yet blind to white group bloc voting).

31. 163 U.S. 537 (1896). *See generally* CHARLES A. LOFGREN, THE PLESSY CASE: A LEGAL-HISTORICAL INTERPRETATION (1987).

32. *See* C. VANN WOODWARD, THE STRANGE CAREER OF JIM CROW (3d rev. ed., 1974). "Jim Crow" laws were laws that imposed racial separation in schools, common carriers, public accommodations, and in public facilities generally, both governmental and private.

33. *Plessy*, 163 U.S. at 559 (Harlan, J., dissenting).

Education case,[34] decided in 1954, which made segregation in schools unconstitutional.

The rediscovery in the seventies of colorblind notions in law, was given a boost by the appropriation by conservative forces of colorblind discourse within the civil rights movement,[35] as captured in the 1963 speech of Martin Luther King, Jr., a civil rights hero. King aspired to a time when every person would be judged "by the content of his character rather than the color of his skin."[36] However, while Dr. King believed, worked, and died fighting for peace and racial and economic justice, the application of legal colorblindness has worked to undermine that dream. A central theme of Critical Race Theory, therefore, is to explore the ways in which legal colorblindness, in supplanting overt legal racial ordering, has not only allowed law to ignore the social and institutional structures of oppression created historically and recreated presently in law and practice but also has blunted efforts to dismantle the racial caste system, working instead to maintain it. Critical Race Theory's main goal is the liberation of minorities and other socially subordinated people; its stance is one of "antisubordination."[37]

Critical Race Theory supports its claim by analyzing cases, laws, and legal patterns that unearth the many ways in which law constitutes and/or supports the status quo of white racial power and black and non-white subordination. For example, CRT scholars have examined the

---

34. 347 U.S. 483 (1954). *See generally* RICHARD KLUGER, SIMPLE JUSTICE: THE HISTORY OF BROWN V. BOARD OF EDUCATION AND BLACK AMERICA'S STRUGGLE FOR EQUALITY (2d ed., 2004). For consideration of the consequences of *Brown*, see generally CHARLES J. OGLETREE, JR., ALL DELIBERATE SPEED: REFLECTIONS ON THE FIRST HALF CENTURY OF BROWN V. BOARD OF EDUCATION (2004); JAMES T. PATTERSON, BROWN V. BOARD OF EDUCATION: A CIVIL RIGHTS MILESTONE AND ITS TROUBLING LEGACY (2001), *reviewed by* Carlo A. Pedriali, *Under a Critical Race Theory Lens*, 7 AFR.-AM. L. & POL'Y REP. 93 (2005).

35. *See* BELL, *supra* note 2, at 115 (referring to modern-day colorblind ideology as a rediscovered constitutional rationale). For early conservative misappropriations of the civil rights aspiration as captured by King, see generally PAUL SEABURY, REVERSE DISCRIMINATION (Barry Gross ed., 1977); NATHAN GLAZER, AFFIRMATIVE DISCRIMINATION: ETHNIC INEQUALITY AND PUBLIC POLICY (1975).

36. Martin Luther King, Jr., *I Have a Dream*, *in* A TESTAMENT OF HOPE: THE ESSENTIAL WRITINGS OF MARTIN LUTHER KING, JR., 217, 219 (James Melvin Washington ed., 1986), *quoted in* CRT: KEY WRITINGS, *supra* note 1, at xv. Dr. King was obviously speaking of judgments of personal moral worth, not about a theory of constitutional interpretation. Constitutional colorblindness is part of the Supreme Court's rationale for the application of strict scrutiny to all racial classifications, even those intended to remedy the effects of discrimination. *See* Adarand Constructors, Inc. v. Pena, 515 U.S. 200, 227 (1995); BELL, *supra* note 2, at 115-35.

37. *See* CATHERINE A. MACKINNON, FEMINISM UNMODIFIED: DISCOURSES ON LIFE AND LAW (1987), *reprinted in part in* POWER, PRIVILEGE, AND LAW, *supra* note 1, at 479-81; Ruth Colker, *Anti-Subordination Above All: Sex, Race, and Equal Protection*, 61 N.Y.U. L. REV. 1003, 1005 (1986); Francisco Valdes, *Outsider Scholars, Critical Race Theory, and "OutCrit" Perspectivity: Postsubordination Vision as Jurisprudential Method*, *in* CROSSROADS, DIRECTIONS, AND A NEW CRITICAL RACE THEORY, *supra* note 1, at 399-409; Elizabeth M. Iglesias, *Structures of Subordination? Women of Color at the Intersection of Title VII and the LLRA. Not!*, 28 HARV. C.R.-C.L. L. REV. 395 (1993), *reprinted in part in* CRITICAL RACE FEMINISM: A READER 317-32 (Adrien Katherine Wing 1997) [hereinafter CRITICAL RACE FEMINISM]; Elizabeth M. Iglesias & Francisco Valdes, *Afterword to LatCrit V Symposium: LatCrit at Five: Institutionalizing a Post-Subordination Future*, 78 DENV. U. L. REV. 1249, 1265-66 (2001).

ways that past naturalization laws,[38] together with cases such as *Dred Scott* both constructed whiteness and defined it as a condition of citizenship.[39] They have examined how current laws legitimate racial profiling of black and other non-white peoples, thereby reinforcing elements of the caste system developed throughout the nation's history.[40] And they have analyzed the ways in which race-neutral housing laws facilitated white flight and suburban sprawl after the *Brown* decision, perpetuating in new form the old pattern of racial residential segregation.[41] From this perspective, CRT rejects the conventional claims of lawyers, judges, and others that law, through the professional processes of reasoned analysis of abstract rules such as equality, is neutral, objective, and distinct from and outside the realm of politics and political choices.[42] Having emerged from a critique of civil rights, Critical Race theorists initially focused on constitutional and civil rights issues. However, they now explore the relationship between white racial power and law in a range of topics from business law[43] to international law.[44]

---

38. Uniform Naturalization Act of 1790, 1 Stat. 103 (repealed 1795).

39. Ozawa v. United States, 260 U.S. 178, 195 (1922); Scott v. Sandford, 60 U.S. 393, 419-20 (1857). *See generally* DON E. FEHRENBACHER, THE DRED SCOTT CASE: ITS SIGNIFICANCE IN AMERICAN LAW AND POLITICS (1978).

40. *See* Frank Rudy Cooper, *The Un-Balanced Fourth Amendment: A Cultural Study of the Drug War, Racial Profiling and Arvizu*, 47 VILL. L. REV. 851 (2002); *see, e.g.*, Whren v. United States, 517 U.S. 806 (1996); *see also* BELL, *supra* note 2, at xx; William H. Buckman & John Lamberth, *U.S. Drug Laws: The New Jim Crow?: Challenging Racial Profiles: Attacking Jim Crow on the Interstate*, 10 TEMP. POL. & CIV. RTS. L. REV. 387 (2001); Angela J. Davis, *Race, Cops, and Traffic Stops*, 51 U. MIAMI L. REV. 425 (1997); David A. Harris, *"Driving While Black" and All Other Traffic Offences: The Supreme Court and Pretextual Traffic Stops*, 87 J. CRIM. L. & CRIMINOLOGY 544 (1997); Tracey Maclin, *Race and the Fourth Amendment*, 51 VAND. L. REV. 338 (1991), *reprinted in part in* CRT: CASES, *supra* note 1, at 206-28; Katheryn K. F. Russell, *"Driving While Black": Corollary Phenomena and Collateral Consequences*, 40 B.C. L. REV. 717 (1999); United States v. Harvey, 16 F.3d 109, 114 (6th Cir. 1994) (Keith, J., dissenting).

41. John O. Calmore, *Spatial Equality and the Kerner Commission Report: A Back-to-the-Future Essay*, 71 N.C. L. REV. 1487, 1492-94 (1993), *reprinted in part in* SOCIAL JUSTICE: PROFESSIONALS, COMMUNITIES, AND LAW 884-92 (Martha R. Mahoney, John O. Calmore, & Stephanie M. Wildman eds., 2003) [hereinafter SOCIAL JUSTICE].

42. *See* Angela P. Harris, *Foreword: The Jurisprudence of Reconstruction*, 82 CAL. L. REV. 741, 746 (1994). *See generally* Phillips, *supra* note 4.

43. *See, e.g.*, CRT: CASES, *supra* note 1, at 138-177; Keith Aoki, *The Stakes of Intellectual Property Law*, *in* THE POLITICS OF LAW, *supra* note 18, at 259-78; Ian Ayres, *Fair Driving: Gender and Race Discrimination in Retail Car Negotiations*, 104 HARV. L. REV. 817 (1991), *reprinted in part in* A WOMAN'S PLACE IS IN THE MARKETPLACE: GENDER AND ECONOMICS 295-318 (Emma Coleman Jordan & Angela P. Harris eds., 2006) [hereinafter A WOMAN'S PLACE]; Ian Ayres, *Further Evidence of Discrimination in New Car Negotiations and Estimates of Its Cause*, 94 MICH. L. REV. 109 (1995); Steven W. Bender, *Consumer Protection for Latinos: Overcoming Language Fraud and English-Only in the Marketplace*, 45 AM. U. L. REV. 1027 (1996); Anthony R. Chase, *Race, Culture, and Contract Law: From the Cotton Field to the Courtroom*, 28 CONN. L. REV. 5, 6-7 (1995); Mechele Dickerson, *Race Matters in Bankruptcy*, 61 WASH. & LEE L. REV. 1725 (2004); Anne-Marie G. Harris, *Shopping While Black: Applying 42 U.S.C. § 1981 to Cases of Consumer Racial Profiling*, 23 B.C. THIRD WORLD L.J. 1 (2003); Cheryl L. Harris, *Whiteness as Property*, 106 HARV. L. REV. 1707 (1993), *reprinted in part in* CRT: CASES, *supra* note 1, at 279-87, *and in* CRT: KEY WRITINGS, *supra* note 1, at 276-91, *and in* A WOMAN'S PLACE, *supra* at 357-71; Emily M.S. Houh, *Critical Interventions: Towards an Expansive Equality Approach to the Doctrine of Good Faith in Contract Law*, 88 CORNELL L. REV. 1025 (2003); Blake Morant, *The Relevance of Race and Disparity in Discussions of Contract Law*, 31 NEW ENG. L. REV. 889 (1997); Muriel Morisey, *Teaching* Williams v. Walker-Thomas Furniture Co., 3 TEMP. POL. & CIV. RTS. L. REV. 89 (1993),

In addition, CRT helped spawn the development of the Latina and Latino Critical Theory (LatCrit)[45] and Asian American critical legal analyses and movements.[46] By shifting the Critical Race Theory lens to other racialized groups, these analyses brought in important discussions of both historical and contemporary issues of citizenship and immigration law as sources of racial and ethnic subordination, as well as, for example, language suppression and stereotypes (Latina/os)[47] or the model minority myth (Asian Americans).[48] These issues were less visible in the original context of CRT's employment of the white over black paradigm and the particularities of the African American experience as analytical frameworks.

The LatCrit movement is particularly interesting because it explicitly incorporates and builds upon feminist legal insights and queer theory as foundational philosophies, explores international issues, and explicitly and consciously articulates the social justice position of antisubordination—a stance against all forms of oppression. Further, LatCrit scholars

---

reprinted in part in CRT: CASES, *supra* note 1, at 160-64; David A. Skeel, Jr., *Racial Dimensions of Credit and Bankruptcy*, 61 WASH. & LEE L. REV. 1695 (2004); Mylinh Uy, *Tax and Race: The Impact on Asian Americans*, 11 ASIAN L.J. 117 (2004); Neil G. Williams, *Offer, Acceptance, and Improper Considerations: A Common-Law Model for the Prohibition of Racial Discrimination in the Contracting Process*, 62 GEO. WASH. L. REV. 183 (1994), *reprinted in part in* CRT: CASES, *supra* note 1, at 138-43, 165-68, 176-77.

44. *See generally* CROSSROADS, DIRECTIONS, AND A NEW CRITICAL RACE THEORY, *supra* note 1, at 303-75; Keith Aoki, *Space Invaders: Critical Geography, the "Third World" in International Law and Critical Race Theory*, 45 VILL. L. REV. 913 (2000), *reprinted in part in* SOCIAL JUSTICE, *supra* note 41, at 880-82; Elizabeth M. Iglesias, *Human Rights in International Economic Law: Locating Latinas/os in the Linkage Debates*, 28 U. MIAMI INTER-AM. L. REV. 361 (1996); CRITICAL RACE FEMINISM, *supra* note 37, at 339-85; GLOBAL CRITICAL RACE FEMINISM: AN INTERNATIONAL READER (Adrien Katherine Wing ed., 2000) [hereinafter GLOBAL CRITICAL RACE FEMINISM]; Chantal Thomas, *Critical Race Theory and Postcolonial Development Theory: Observations on Methodology*, 45 VILL. L. REV. 1195 (2000).

45. *See* Francisco Valdes, *Afterword: Theorizing "OutCrit" Theories: Coalitional Method and Comparative Jurisprudential Experience – RaceCrits, QueerCrits, and LatCrits*, 53 U. MIAMI L. REV. 1266, 1266 (1999).

46. *See generally* WU, *supra* note 1; Chang, *supra* note 1, at 355; Margaret Chon & Donna E. Arzt, *Judgments Judged and Wrongs Remembered: Examining the Japanese American Civil Liberties Cases on Their Sixtieth Anniversary: Walking While Muslim*, 68 LAW & CONTEMP. PROBS. 215 (2005); Frank H. Wu, *The Arrival of Asian Americans: An Agenda for Legal Scholarship*, 10 ASIAN L.J. 1 (2003). Recently these scholars have called their work Asian American Jurisprudence. *See, e.g.*, John Hayakawa Torok, *Asian American Jurisprudence: On Curriculum*, 2005 MICH. ST. L. REV. 635.

47. Richard Delgado & Jean Stefancic, *Images of the Outsider in American Law and Literature: Can Free Expression Remedy Systemic Ills?*, 77 CORNELL L. REV. 1258, 1273 (1992), *reprinted in part in* POWER, PRIVILEGE AND LAW, *supra* note 1, at 82-89.

48. Frank H. Wu, *Changing America: Three Arguments About Asian Americans and the Law*, 45 AM. U. L. REV. 811, 813-14 (1996); Chang, *supra* note 1; Uy, *supra* note 43, at 131-32. *See generally* NAZLI KHRIA, BECOMING ASIAN AMERICAN: SECOND-GENERATION CHINESE AND KOREAN IDENTITIES (2002); DIANA TING LIU WU, ASIAN PACIFIC AMERICANS IN THE WORKPLACE (1997); STACEY J. LEE, UNRAVELING THE "MODEL MINORITY": LISTENING TO ASIAN AMERICAN YOUTH (1996).

have established an institutional framework for the future development of LatCrit and other Critical Race Theory scholarship.[49]

While feminist theory, particularly feminist legal theory and black feminist theory, were inherent in the original Critical Race Theory scholarship through the work of people like Crenshaw and Angela Harris,[50] Critical Race Feminism, a term coined originally by Richard Delgado, and a field adopted and promoted specifically by Adrien Wing, has taken off as a separate, newly-developing theory and body of scholarship.[51] Critical Race Feminism builds on Critical Race Theory as well as insights specifically from black feminist theory and is also a leftist legal critique that primarily focuses on the intersections of race, ethnicity, and/or colonialism on the one hand and gender on the other. In addition, it explores the international manifestations of racialized gender oppression. The exploration of the sex/gender system, generally, its relationship to the racial order, and the living reality of sexual minorities of color, coupled with Critical Race Theory's embrace of the larger social justice project of working toward the liberation of all, has resulted in many critical race theorists also examining the ways in which law subordinates sexual minorities.[52]

Together these issues have led to insights about the ways in which identity is multidimensional. For instance, critical race theorists, among others, point out that people are not simply raced (black, white, yellow, or "Hispanic"); they are also gendered, (masculine, feminine or transgendered) and possess sexual identities (heterosexual, homosexual, or bisexual), etc. From this perspective, every person's identity is multidimensional.[53] This insight of multidimensionality goes further. The so-

49. LatCrit literature is readily available at the LatCrit website, Latina and Latino Critical Theory (2004), http://personal.law.miami.edu/~fvaldes/latcrit/latcrit/index.html (listing, *inter alia*, 16 colloquia and symposia on LatCrit). *See generally*, THE LATINO/A CONDITION, *supra* note 1; Colloquium: *Representing Latina/o Communities: Critical Race Theory and Practice*, 9 LA RAZA L.J. 1 (1996); Panel: *Latina/o Identity and Pan-Ethnicity: Toward LatCrit Subjectivities*, 2 HARV. LATINO L. REV. 175 (1997); Joint Symposium: *Latinas/os and the Law*, 85 CAL. L. REV. 1087 (1997).

50. *See generally* CRT: KEY WRITINGS, *supra* note 1; Angela P. Harris, *Race and Essentialism in Feminist Legal Theory*, 42 STAN. L. REV. 581 (1990), *reprinted in part in* POWER, PRIVILEGE AND LAW, *supra* note 1, at 484-85, 539-43, *and in* CRT: CUTTING EDGE, *supra* note 1, at 253-66, *and in* CRITICAL RACE FEMINISM, *supra* note 35, at 11-18.

51. *See* CRITICAL RACE FEMINISM, *supra* note 37, at 1; GLOBAL CRITICAL RACE FEMINISM, *supra* note 44; CRT: CUTTING EDGE, *supra* note 1, at 261-74; Adrien K. Wing & Christine A. Willis, *From Theory to Praxis: Black Women, Gays, and Critical Race Feminism*, 4 AFR.-AM. L. & POL'Y REP. 1 (1999); *see also Women of Color in Legal Academia: A Biographic and Bibliographic Guide*, 16 HARV. WOMEN'S L.J. 1 (1993).

52. *See generally* Hutchinson, *supra* note 19; Francisco Valdes, *Queer Margins, Queer Ethics: A Call to Account for Race and Ethnicity in the Law, Theory and Politics of "Sexual Orientation,"* 48 HASTINGS L.J. 1293 (1997).

53. Berta Esperanza Hernández-Truyol, *Women's Rights as Human Rights – Rules and Realities and the Role of Culture: A Formula for Reform*, 21 BROOK. J. INT'L L. 605, 667-68 (1996). *See generally* Darren Lenard Hutchinson, *Identity Crisis: "Intersectionality," "Multidimensionality," and the Development of an Adequate Theory of Subordination*, 6 MICH. J. RACE & L. 285 (2001), *reprinted in part in* SOCIAL JUSTICE, *supra* note 41.

cial structures of race, gender, sexuality, and class as systems of power are interrelated and mutually reinforcing. They create multiple and intersecting positions of subordination for members of the groups they disadvantage. So for example, racism in the United States is patriarchal and patriarchy in the United State is racist. As Dorothy Roberts points out, some of the first laws passed in America involved changing the legal/social practice of children inheriting the status of their fathers, to children inheriting the status of their mother if their mothers were (black) slaves, while patrolling and later prohibiting sexual relations between white women and black men. "Black women [were forced to] produce[] children who were legally Black to replenish the master's supply of slaves . . . [while] White women [were compelled to] produce[] white children to continue the master's legacy."[54] Black men passed on little.[55] The racism that both black men and women experienced was thus, also gendered. To these analyses, scrutiny of the specific relations of class could be added as well as an analysis of compulsory heterosexuality. As such, the structure of black oppression was and has remained multidimensional.

CRT theorists have often explored the material harms caused by race, gender, and sexuality as mutually reinforcing social systems embedded in and constructed, albeit not exclusively, by law; and they have suggested that these systems reinforce the reproduction of class within the United States. However, they have not in any sustained manner theorized the ways in which class functions as a site for identity formation, or the various ways in which class as a specific function of the creation and distribution of resources, operates both independently and mutually with other subordinating structures to limit the material well-being of people including racial and other minorities. These kinds of analyses await further development and constitute a necessary future direction of CRT.

II. INTELLECTUAL ANTECEDENTS

Many scholars have described the origins of Critical Race Theory. They suggest that it owes its intellectual genesis to three intellectual movements. The first is the civil rights movement and the critical assessments of its effects in changing the actual conditions of black life. Second, CRT builds upon the themes and critical understandings of law exposed by the Critical Legal Studies movement.[56] And third, it incor-

54. Dorothy E. Roberts, *Racism and the Patriarchy in the Meaning of Motherhood*, 1 AM. U. J. GENDER SOC. POL'Y & L. 1, 8 (1993), *reprinted in* FEMINIST LEGAL THEORY, VOLUME II (Frances E. Olsen ed., 1995). *See generally* JACQUELINE JONES, LABOR OF LOVE, LABOR OF SORROW: BLACK WOMEN, WORK AND THE FAMILY FROM SLAVERY TO THE PRESENT (1985).

55. Presumably where black men were free and had children with a free woman, he passed on his legacy.

56. *See generally* CRT: KEY WRITINGS, *supra* note 1; CRT: CUTTING EDGE, *supra* note 1. For a discussion on the Critical Legal Studies (CLS) movement, see generally CRITICAL LEGAL STUDIES (Jones Boyle ed., 1992); CRITICAL LEGAL STUDIES (Allan C. Hutchinson ed., 1989);

porates many of the insights and theorizing of feminist legal and other feminist scholars.[57]

Richard Delgado traces CRT genesis to the early seventies when several legal scholars began to express doubt about the effectiveness of the civil rights movement's legal strategy to racial justice. For instance, Derrick Bell, considered a forefather of CRT, in an essay in 1976, suggested that civil rights attorneys' approach to litigating school cases for purposes of desegregating entire school districts (and balancing them racially) might be at odds with their clients'—African American families—very real hopes and concrete goals of immediately improving their children's educations.[58] In another article, Bell argued that the result in the 1954 *Brown v. Board of Education* decision, although heralded as a triumph of the civil rights legal strategy, might be better explained by what he called "interest convergence,"[59] a theme that has become a mainstay of critical race analysis.[60] *Brown*, he suggested, came about not because of some belated realization by whites of the harms black children suffered under segregation,[61] but rather, because of its value to

---

CRITICAL LEGAL STUDIES (Peter Fitzpatrick & Alan Hunt eds., 1987); CRITICAL LEGAL STUDIES: ARTICLES, NOTES, AND BOOK REVIEWS SELECTED FROM THE PAGES OF THE HARVARD REVIEW (1986); MARK KELMAN, A GUIDE TO CRITICAL LEGAL STUDIES (1987); ROBERTO MANGABEIRA UNGER, THE CRITICAL LEGAL STUDIES MOVEMENT (1986); Symposium: *Critical Legal Studies*, 36 STAN. L. REV. 1-674 (1984); Mark Tushnet, *Critical Legal Studies: A Political History*, 100 YALE L.J. 1515 (1991).

57. *See generally* CRT: CUTTING EDGE, *supra* note 1, at 499-550.

58. Derrick Bell, *Serving Two Masters: Integration Ideals and Client Interests in School Desegregation Litigation*, 85 YALE L.J. 470 (1976), *reprinted in* CRT: KEY WRITINGS, *supra* note 1, at 5-19, *and in part in* CRT: CUTTING EDGE, *supra* note 1, at 228-38, *and in* SOCIAL JUSTICE, *supra* note 41, at 328-33; *see* DELGADO & STEFANCIC, CRT: AN INTRODUCTION, *supra* note 1, at 30-31.

59. Derrick Bell, Brown v. Board of Education *and the Interest-Convergence Dilemma*, 93 HARV. L. REV. 518, 524 (1980), *reprinted in* CRT: KEY WRITINGS, *supra* note 1, at 20-29 [hereinafter Bell, *Interest Convergence Dilemma*]; *see also* DERRICK BELL, AND WE ARE NOT SAVED: THE ELUSIVE QUEST FOR RACIAL JUSTICE 51-74 (1987). Bell's theory had a precursor in the "Slave Power" theory constructed by antislavery elements, such as future Chief Justice Samuel P. Chase, prior to the Civil War. They argued that the South was under the political control of slave owners who, through the Democratic Party, dominated all branches of the national government, which was its tool for the denial, in the interests of slavery, of the freedoms of Northern whites. This became part of the ideology of the Republican Party whose original commitment was to prevent the extension of slavery. *See generally* ERIC FONER, FREE SOIL, FREE LABOR, FREE MEN: THE IDEOLOGY OF THE REPUBLICAN PARTY BEFORE THE CIVIL WAR 73-102 (1970); MICHAEL F. HOLT, THE POLITICAL CRISIS OF THE 1850S (1978).

60. Delgado, *Blind Alleys*, *supra* note 17, at 124; *see, e.g.*, Steven A. Ramirez, *Games CEOs Plan and Interest Convergence Theory: Why Diversity Lags in America's Boardrooms and What to Do About It*, 61 WASH. & LEE L. REV. 1583 (2004); Michael Z. Green, *Addressing Race Discrimination Under Title VII after Forty Years: The Promise of ADR as Interest-Convergence*, 48 HOW. L.J. 937 (2005); John Hayakawa Torok, *"Interest Convergence" and the Liberalization of Discriminatory Immigration and Naturalization Laws Affecting Asian Americans, 1943-1965*, 1995 CHINESE AMERICA: HISTORY & PERSPECTIVES 1-15; Marlia Banning, *Critical Race Theory and Interest Convergence in the Desegregation of Higher Education*, *in* RACE IS – RACE ISN'T: CRITICAL RACE THEORY AND QUALITATIVE STUDIES IN EDUCATION (Laurence Parker, Donna Deyhle, & Sofia Villenas eds., 1999).

61. Bell, *Interest-Convergence Dilemma*, *supra* note 59, at 523-24. The Brown decision also ignored, and apparently could not contemplate, the harms segregation visited upon white students. These harms as Kevin Brown suggests include a false sense of superiority. *See* BROWN, *supra* note 30, at 164-67. He notes that prominent social scientists filed an amicus brief in the Brown cases

whites—a value and interest that converged with black aspirations for freedom and well-being.[62] The decision, he suggests, was intended to and "helped to provide immediate credibility to America's struggle with Communist countries to win the hearts and minds of emerging third world peoples."[63] Bell's argument that the goal of the U.S. government in advocating for the decision in *Brown* had little to do with improving the education or life chances of black children, rang particularly accurate in 2004 when, fifty years after the decision, education in the United States was found to be as segregated (and disadvantaging) in terms of race as it had been at the time of *Brown*.[64] Bell later argued that racism was pervasive in the American social order, that law was imbued with it, and that racism was a permanent feature of American society including its legal system.[65]

---

suggesting that segregation hurt all children both black and white, not simply black children. The scientists noted that children taught prejudice learn:

> to gain personal status in an unrealistic and non-adaptive way. When comparing themselves to members of the minority group, they are not required to evaluate themselves in terms of the more basic standards of actual personal ability and achievement. The culture [allows] . . . them to direct their feelings of hostility and aggression against whole groups of people . . . perceived as weaker than themselves. They often develop patterns of guilt feeling, rationalizations and other mechanisms which they must use in an attempt to protect themselves from recognizing the essential injustice . . . .

Id. at 165. Brown concludes that "if segregation created a false sense of inferiority within blacks, then it must have also generated the psychological harm of a false sense of superiority in whites." *Id.* at 165-66. He explains this as the dual psychological harm of segregation, the two different sides of the same delusion suffered by the entire society. *Id.* at 166.

The difficulties experienced by minority teachers (as well as women) in the classroom in many ways capture this harm in that white students who often know little about the subject matter of the class act as if there is little a minority teacher can teach them. Consider Derrick Bell's experience at Stanford Law School in 1986 when first year law students presumed to know more about constitutional law than Professor Bell did and therefore critiqued the class as not covering the appropriate subject matter. The Stanford administration, though later apologizing, initially capitulated to the student's estimation. *See* Derrick Bell, *The Price and Pain of Racial Perspective*, STAN. L. SCH. J., May 9, 1986, at 5. In fact these teachers have written quite a bit about their classroom difficulties as minority teachers and the other challenges they face in the context of the predominately white male American law school. *See, e.g.*, Richard Delgado & Derrick Bell, *Minority Law Professors' Lives: The Bell-Delgado Survey*, 24 HARV. C.R.-C.L. L. REV. 349, 360 (1989); Okianer Christian Dark, *Just My 'Magination*, 10 HARV. BLACKLETTER L.J. 21, 23 (1993); Trina Grillo, *Tenure, and Minority Women Law Professors: Separating the Strands*, 31 U.S.F. L. REV. 747, 753-54 (1997); Reginald Leamon Robinson, *Teaching from the Margins: Race as a Pedagogical Sub-text: A Critical Essay*, 19 W. NEW ENG. L. REV. 151, 152 (1997). For the perspective of critical students, see Kathryn Pourmand Nordick, *Essay: A Critical Look at Student Resistance to Non-Traditional Law School Professors*, 27 W. NEW ENG. L. REV. 173, 174-75 (2005).

62. Bell, *Interest-Convergence Dilemma*, *supra* note 59, at 524.

63. Bell, *Interest-Convergence Dilemma*, *supra* note 59, at 524; *see also* Mary L. Dudziak, *Desegregation as a Cold War Imperative*, 41 STAN. L. REV. 61 (1998), *reprinted in* CRT: CUTTING EDGE, *supra* note 1, at 106-17 (noting that the government stated this in their amicus brief submitted in the *Brown* case and further substantiating Bell's intuition).

64. This point was made during celebrations marking the fiftieth anniversary of *Brown*. *See, e.g.*, Greg Toppo, *Integrated Schools Still a Dream 50 Years Later*, USA TODAY, Apr. 28, 2004, at A1.

65. BELL, *supra* note 2, at 1; *see* Bell, *supra* note 59, at 518; DERRICK BELL, FACES AT THE BOTTOM OF THE WELL: THE PERMANENCE OF RACISM 3 (1993).

> Ignored in the rush to proclaim color blindness as the judicial panacea to claims of racial injustice is the fact that virtually all policies adopted as protections against racial injus-

At about the same time, the Critical Legal Studies movement based in the legal academy was growing. This movement questioned the entire edifice of law as an objective arbiter of social conflict distinct from the messiness of politics and political choices. This movement, which included scholars such as Alan Freeman, Peter Gabel, Duncan Kennedy, and Mark Tushnet, successfully demonstrated the ways in which legal rules were, in and of themselves, not determinative of a particular result.[66] They showed that, for any given rule, there were multiple, contrasting and conflicting rules whose resolution required actors to make choices. These choices were political ones that generally reflected, supported, and legitimized the social power of dominant classes.[67]

However, they rejected what they termed "vulgar instrumentalist" or "structuralist" accounts of law that understand it as merely a tool and reflection of bourgeoisie/elite interests and ideas, or as a merely superstructural phenomenon determined by the underlying economic base. Rather, drawing in part on the Italian Marxist philosopher Antonio Gramsci's idea of hegemony,[68] Critical Legal Studies scholars understand law as a complex system with many functions, one of which is to exercise and simultaneously *legitimate* the use of institutional violence within the *prevailing* social arrangements in a way that gains the consent

---

tices suffered by blacks and other people of color in this country have actually proven to be of more value to whites.

BELL, *supra* note 2, at xx; *see also* Derrick Bell, *Racial Realism*, 24 CONN. L. REV. 363, 373 (1992); Derrick Bell, *Racial Realism – After We're Gone: Prudent Speculations on America in a Post-Racial Epoch*, 34 ST. LOUIS U. L.J. 393 (1990), *reprinted in* CRT: CUTTING EDGE, *supra* note 1, at 2-8 [hereinafter Bell, *Racial Realism*].

66. *See* THE POLITICS OF LAW, *supra* note 18, at 1-7 (providing a good overview of Critical Legal Studies themes). This book has been revised several times. The first edition in 1982 was updated and followed by a revised edition in 1990, and then again in a third edition in 1998, *supra* note 1. I like each of these editions but find the 1982 publication the best for introducing many of the basic CLS concepts and initial ideas. In this edition the ideas tend to be more fully explained, whereas in later publications some key insights are merely summarized. The later editions, however, introduce new thinking developed in a covered area and introduce additional essays that explore a wider breadth of legal fields.

67. *See* Victor Rabinowitz, *The Radical Tradition in the Law, in* THE POLITICS OF LAW, *supra* note 18, at 686; *see, e.g.*, ROBERT L. HAYMAN, JR., NANCY LEWIS & RICHARD DELGADO, JURISPRUDENCE CLASSICAL AND CONTEMPORARY: FROM NATURAL LAW TO POSTMODERNISM 402-13 (2002).

68. Hegemony is an illusive concept, but is to be counterposed to "direct domination" as a paradoxical "'spontaneous' consent given by the great masses of the population to the general direction imposed on social life by the dominant fundamental group; this consent is 'historically' caused by the prestige (and consequent confidence) which the dominant group enjoys because of its position and function in the world of production." SELECTIONS FROM THE PRISON NOTEBOOKS OF ANTONIO GRAMSCI 12, 161, 170, 416-17 (Quintin Hoare & Geoffrey Nowell Smith eds., 1971); *see also* Robert W. Gordon, *Some Critical Theories of Law and Their Critics, in* THE POLITICS OF LAW, *supra* note 18, at 647-48; Douglass Litowitz, *Gramsci, Hegemony and the Law*, 2000 BYU L. REV. 515, 515-16 (2000); ALAN HUNT, DICHOTOMY AND CONTRADICTION IN THE SOCIOLOGY OF LAW, *in* MARXISM AND LAW 86-87 (Piers Beirne & Richard Quinney eds., 1982); Duncan Kennedy, *Antonio Gramsci and the Legal System*, 6 ALSA F. 32, 32 (1982). For an application of Gramsci's concept of hegemony to antebellum slavery, see EUGENE V. GENOVESE, ROLL, JORDAN, ROLL: THE WORLD THE SLAVES MADE 25-49 (1976), *reprinted in part in* MARXISM AND LAW, *supra* at 279-94.

and acquiescence of the subordinated to their conditions.[69] Law does this political work by deploying a distinct and elaborate discourse and body of knowledge (popularly perceived as objective and apolitical) to justify its decisions. These decisions, while generally supporting the power of elite groups, sometimes actually restrain the exercise of power and occasionally provide justice to ordinary people. In doing so, however, they lend legitimacy to law and to many of the existing social arrangements and institutions of which law is a part. Thus, while there may indeed be "a difference between arbitrary power and [the] rule of law," law may also be "in some part sham."[70]

So for example, in a path-breaking article, Alan Freeman argued that antidiscrimination law offered a credible measure of tangible progress without in any way disturbing the basic class structure of the American society.[71] This was accomplished by using concepts such as intent, fault, colorblindness and formal equality, which over time ultimately located the problem of racism in the intentional actions of bad actors instead of the established caste system that included the conscious and unconscious habitual human and institutional practices of racial ordering. The remedy to the problem, defined in this manner, was to compel the bad actors to act differently instead of changing or dismantling the caste system of embedded racial arrangements.[72] Thus, although judges declared that law would treat everyone the same, they did so without regard to and so without changing the conditions that stratified people(s) socially and materially in the first place. As such, antidiscrimination law outlawed the obvious and explicit manifestations of racism (the "white only" signs of the Jim Crow era) and thereby provided credible evidence that the law and the basic structure of society were fair,

---

69. My notion here of legal discourse as backed by violence and justifying its use comes from Robert Cover. *See generally*, Robert Cover, *Violence and the Word*, 95 YALE L.J. 1601 (1986). I thank my colleague John Henry Schlegal for pointing out that I should clarify this.

70. Rabinowitz, *supra* note 67, at 688 (citing E.P. Thompson, the celebrated English radical historian who acknowledged the element of false consciousness induced by the legal order, but affirmed the rule of law as a potential basis for progressive change; *see* E.P. THOMPSON, WHIGS AND HUNTERS: THE ORIGIN OF THE BLACK ACT 265-66 (1975), *reprinted in part in* MARXISM AND LAW, *supra* note 68, at 130-37). Another of Gramsci's concepts is the "organic intellectual," who is, not a social technician (like most lawyers) but rather someone who instructs popular consciousness "precisely in order to construct an intellectual-moral bloc which can make politically possible the intellectual progress of the race and not only of small intellectual groups." SELECTIONS FROM THE PRISON NOTEBOOKS OF ANTONIO GRAMSCI, *supra* note 68, at 332-33. For a CRT application of these concepts which regards Martin Luther King as one such organic intellectual, see ANTHONY E. COOK, BEYOND CRITICAL LEGAL STUDIES: THE RECONSTRUCTIVE THEOLOGY OF DR. MARTIN LUTHER KING, JR., *in* CRT: KEY WRITINGS, *supra* note 1, at 90-91.

71. Alan Freeman, *Legitimizing Racial Discrimination Through Antidiscrimination Law: A Critical Review of Supreme Court Doctrine*, 62 MINN. L. REV. 1049 (1978), *reprinted in* MARXISM AND LAW, *supra* note 68, at 210-35 [hereinafter Freeman, *Legitimizing Racial Discrimination*]; *see also* Alan Freeman, *Antidiscrimination Law: from 1954 to 1989: Uncertainity, Contradiction, Rationalization, Denial*, *in* THE POLITICS OF LAW, *supra* note 18, at 285-311.

72. Freeman also suggests that the Court's momentary flirtation with attacking the actual structural conditions of subordination through ordering school desegregation lent support to its rhetoric of change.

without disturbing the structural and systemic manifestations, including the maldistribution of resources, of that same deeply embedded racism.

And finally, though perhaps not initially obvious, CRT owes a debt to yet another school of thought: feminism. As Delgado notes, CRT builds upon feminist "insights into the relationship between power and the construction of social roles, as well as the unseen, largely invisible collection of patterns and habits that make up patriarchy and other types of domination."[73] In addition, the idea of antisubordination, the central stance of "race crits," can be traced not only to race scholars but also to feminist scholarship.[74]

III. CONFLICT AS AN ENGINE OF CRT INTELLECTUAL AND INSTITUTIONAL GROWTH

These ideas, according to Kimberlé Crenshaw, met in the actual persons of students attending Harvard Law School in 1981 where a struggle raged over the meaning of race. This struggle was a part of a larger continuum of student movements at universities in the 1970s and 1980s advocating for ethnic study departments, against South African apartheid, and for diversity in student admissions and faculty staffing.[75] The struggle implicated notions of race consciousness, affirmative action, and the presumptive existence of meritocracy and proved to be a catalyst for the theory's institutionalization. It also began a pattern of conflict that would engender crucial CRT insights and fuel its growth and entrenchment.

Specifically, the conflict at Harvard led to the Alternative Course, described by Crenshaw as CRT's first institutional expression. Second, conflict with Critical Legal Studies helped to inspire the formal establishment of the CRT workshop. Third, internal conflict and critique over the commitments and focus of the workshop led to the exploration of the experiences of additional racialized and oppressed groups. The later founding of LatCrit, resulting in part from these conflicts, brought about concerns about fragmentation of the movement.[76] Nevertheless, it embodied and represented a continuing site for the germination of "Race Crit" insights. And last, the conflict over Proposition 209[77] and its visions of colorblind justice in the context of the University of California Los Angeles Law School led to CRT's entrenchment.

---

73. DELGADO & STEFANCIC, CRT: AN INTRODUCTION, *supra* note 1, at 5.

74. *See supra* note 37.

75. Sumi Cho & Robert Westley, *Historicizing Critical Race Theory's Cutting Edge: Key Movements That Performed the Theory*, 33 U.C. DAVIS L. REV. 1377 (2000), *reprinted in* CROSSROADS, DIRECTIONS, AND A NEW CRITICAL RACE THEORY, *supra* note 1, at 32-70.

76. Crenshaw, *supra* note 3, at 1356-57.

77. California Civil Rights Initiative, California Ballot Proposition 209 (adopted Nov. 6, 1986) (codified at CAL. CONST. art. I, § 31) [hereinafter California Proposition 209].

A. Alternative Course: Confronting Colorblindness

After the departure of Derrick Bell from Harvard Law School in 1981, a group of students asked the administration to continue to offer his semester-long course on Constitutional Law and Minority Issues.[78] This course, according to Crenshaw, fore-grounded race in the context of legal issues rather than analyzing laws from the typical liberal individual rights perspective of "how do we get blacks some rights" while reconciling other *vested* interests.[79] Further it assessed laws on the basis of their effectiveness in changing the actual conditions of black subordination while also analyzing the ways in which many of these same laws actually contributed to racial subordination.[80]

These students, seeing the hole in the curriculum created by Bell's departure as an opportunity to "desegregate the faculty," asked the administration to hire a person of color to teach the course.[81] The Dean's response, in particular, was instructive to students who later institutionalized Critical Race Theory. He questioned any value a course on "Constitutional Law and Minorities Issues" might add to a curriculum that already offered courses such as constitutional law and employment discrimination law, both of which dealt with "those" issues. [82] The Dean asked "why the students would not prefer an excellent white professor over a mediocre black one."[83] And he suggested that there were no people of color in the country "qualified" to be hired at Harvard Law School.[84] His comment reflected a particular perspective that provoked what later came to be key CRT themes. First, his comments suggested that qualifications or merit were something other than socially defined worth that themselves might embody racial meanings and structure.[85] Second, although this merit supposedly had nothing to do with color because merit was colorblind;[86] it was nonetheless captured by the then current predominantly white professors and not in any black professors in the country at the time.[87] And, third, that Harvard, as a race-neutral

78. CRT: KEY WRITINGS, *supra* note 1; Crenshaw, *supra* note 3, at 1348. Bell later returned to Harvard Law School. During his second stay, he staged a protest in support of Harvard hiring black women. The protest is documented in part in his book, CONFRONTING AUTHORITY. *See generally* DERRICK BELL, CONFRONTING AUTHORITY: REFLECTIONS OF AN ARDENT PROTESTOR (1994).

79. Crenshaw, *supra* note 3, at 1347.

80. CRT: KEY WRITINGS, *supra* note 1, at xxi; Crenshaw, *supra* note 3, at 1347.

81. Crenshaw, *supra* note 3, at 1345.

82. CRT: KEY WRITINGS, *supra* note 1, at xxi.

83. *Id.* at xx.

84. *Id.*; Harris, *supra* note 1, at 1221 n.10.

85. *See, e.g.*, Daria Roithmayr, *Deconstructing the Distinction Between Bias and Merit*, 85 CAL. L. REV. 1449, 1452-1453 (1997).

86. *See, e.g.*, Harris, *supra* note 25, at 911-12 (on colorblindness).

87. *See* Crenshaw, *supra* note 3, at 1348-49 (discussing the pool problem); *see also* CHO & WESTLEY, *supra* note 73, at 46 (explaining that the "pool argument" was not credible and noting that the percentage of faculty of color teaching in law schools grew by 85% in two years (1989-1991) after the first nationwide strike by students advocating for faculty diversity).

institution, supposedly did not hire people on the basis of color, and yet the faculty was overwhelmingly white. Crenshaw notes: "This framing of the issue gave many of us involved in that struggle a clear sense about how conceptions such as colorblindness and merit functioned as rhetoric of racial power in presumptively race-neutral institutions."[88]

Harvard responded to the students' request by hiring two distinguished civil rights lawyers, neither of whom was a person of color, to teach a three-week mini course. The students rejected this offer, setting off a national debate on affirmative action.[89] Instead they organized an alternative course in which they invited a number of professors to teach various parts of the course.[90] Several of the participants later became prominent figures in CRT, including such scholars as Denise Carty-Benia, Richard Delgado, Linda Green, Neil Gotanda, and Charles Lawrence, who were already law professors, Kimberlé Crenshaw, a law student, and Mari Matsuda, a graduate student at the time. Crenshaw marks the "Alternative Course" as CRT's first institutional expression. Its second institutional expression was in 1989 when Crenshaw, together with Stephanie Phillips, Neil Gotanda and others and with the support of Richard Delgado and the backing of David Trubek, then the director of the Institute of Legal Studies at Wisconsin, organized the first CRT workshop in Wisconsin.[91] Budding scholars such as Angela Harris were to attend this workshop.[92]

B. Conflict with CLS: The African American Experience as an Analytical and Methodological Framework

In the intervening time, many of these and other future "race crits" had been meeting informally, and in separate sessions, at the Critical Legal Studies meetings, retreats, and summer camps. This engagement had both an intellectual and institutional contribution to CRT. First, the CRT workshop was patterned after the Critical Legal Studies (CLS) summer workshops involving small groups of people to "explore a range of topics."[93] Further, future race crits had clashed with CLS scholars in the 1985 and 1987 conferences, when they sought to critique CLS both at the level of practice and theory. At the level of practice, they questioned the whiteness of CLS (as well as the elite maleness of it) and the way these social positionings affected, and potentially limited, CLS analyses. At the level of theory, race crits questioned one of the major theoretical

88. Crenshaw, *supra* note 3, at 1345.
89. CRT: KEY WRITINGS, *supra* note 1, at xxi.
90. *Id.*
91. Crenshaw, *supra* note 3, at 1361.
92. *Id.* at 1361 n.19. Other participants at the Wisconsin workshop, according to Stephanie Phillips, included Paulette Caldwell, John Calmore, Harlon Dalton, Kendall Thomas, and Patricia Williams. *Id.*
93. Crenshaw, *supra* note 3, at 1359.

critiques of CLS scholarship, the trashing of rights.[94] Some CLS scholars argued that the idea of legal rights and "rights talk"—"I have this or that right"—engaged in by the legal and popular public alike, actually obscured and narrowed the actual concrete conflicts that underlie the talk. Further, they suggested that rights talk limited people's ability to creatively imagine alternative frameworks and solutions to problems that might include measures or activities outside the realm of law. Some CLS scholars thus advocated abandoning and trashing rights.[95] Future race crits partly rejected this line of argument because it neglected the reality that blacks and other people of color had in fact used the language of legal rights through the civil rights movement to effect some change in their social treatment, even if using the rights discourse had failed to alter the fundamental conditions of their oppression.[96]

In using the lens of the African American experience to critically examine and challenge the CLS critique of rights, African American insights and experience became a central part of Critical Race Theory's analytical framework. Further, testing CLS insights against the actual experiences and perspectives of African Americans, as embodied in the civil rights movement for instance, informed CRT's methodology. And, use of the African American experience also arguably fueled CRT's optimism and commitment to the progressive use of law and to the modernist ideals of justice, equality, and dignity.

Nonetheless, CLS scholars' hostile resistance to this critique, and the maelstrom it created, cast a shadow on the race crits' continued engagement with CLS. In doing so, it sowed the seeds that a leftist, critical law engagement with race required its own space. The critical theory race workshop became that space.

---

94. *See generally Minority Critique of the Critical Legal Studies Movement*, 22 HARV. C.R.-C.L. L. REV. 297 (1987) (including articles by Richard Delgado, Mari Matsuda, Patricia Williams, and Harlon Dalton); Kimberlé Williams Crenshaw, *Race,Reform, and Retrenchment: Transformation and Legitimation in Antidiscrimination Law*, 101 HARV. L. REV. 1331, 1356-86 (1988), *reprinted in part in* CRT: CASES, *supra* note 1, at 27-38; Richard Delgado, *The Ethereal Scholar: Does Critical Legal Studies Have What Minorities Want?*, 22 HARV. C.R.-C.L. L. REV. 301, 302 (1987).

95. *See, e.g.*, Freeman, *Legitimizing Racial Discrimination*, *supra* note 71; Peter Gabel, *The Phenomenology of Rights-Consciousness and the Pact of the Withdrawn Selves*, 62 TEX. L. REV. 1563, 1590 (1984); Frances Olsen, *Statutory Rape: A Feminist Critique of Rights Analysis*, 63 TEX. L. REV. 387, 427 (1984); Mark Tushnet, *The Critique of Rights*, 47 SMU L. REV. 23 (1993), *reprinted in part in* POWER, PRIVILEGE AND LAW, *supra* note 1, at 549-55; Mark Tushnet, *An Essay on Rights*, 62 TEX. L. REV. 1386, 1382-83 (1984). The CLS critique of rights, just as some other CLS ideas, was drawn from the Legal Realists of the 1920s and 1930s. *See* Elizabeth Mensch, *The History of Mainstream Legal Thought*, *in* THE POLITICS OF LAW, *supra* note 18, at 34-35; *see, e.g.*, Felix Cohen, *Transcendental Nonsense and the Functional Approach*, 35 COLUM. L. REV. 809 (1935).

96. CRT: KEY WRITINGS, *supra* note 1, at xxiii-xxiv; DELGADO & STEFANCIC, CRT: AN INTRODUCTION *supra* note 1, at 23-25; PATRICIA J. WILLIAMS, THE ALCHEMY OF RACE AND RIGHTS: DIARY OF A LAW PROFESSOR (1991), *reprinted in part in* POWER, PRIVILEGE AND LAW, *supra* note 1, at 35-36, 558-64; Crenshaw, *supra* note 94, at 1334; Phillips, *supra* note 4, at 1249; GLOBAL CRITICAL RACE FEMINISM: AN INTERNATIONAL READER, *supra* note 44, at 3. Some in CLS by now acknowledge the partial validity of the CRT counter-critique of their rights critique. *See, e.g.*, Gordon, *supra* note 68, at 657-58.

C. *Internal Conflict within the CRT Workshop: Expanding the Analytical Framework and Building the Commitment to Antisubordination*

The critical race workshop met annually from 1989 until 1997.[97] Attendance at the workshop was by application and "invitation only"—a policy that contributed to the critique of CRT as elitist.[98] However, according to Phillips, the policy was meant to facilitate sustained engagement over a five-day period of a small group of people committed to "radical transformative politics."[99] As such, the workshop did not issue "a general invitation to all legal scholars of color, no matter how conservative or parochial, to simply come hang out."[100] Rather, she suggests, this was the workshop's attempt to institute what Frank Valdes later called "a move from color to consciousness," the idea that "alliances are best built on shared substantive commitments, perhaps stemming from similar experiences . . . with subordination, rather than traditional fault lines like race or ethnicity."[101] But, this principle was not extended to include white scholars with similar commitments. White scholars were excluded from participating in the workshop, a decision which generated debate as to whether this was a pragmatic attempt to construct safe space and inhibit the reproduction of white racial hierarchy, or simply an unprincipled decision.[102] In either case, the question became moot with the cessation of annual CRT workshops and the almost simultaneous founding of the annual LatCrit conferences with its commitment to anti-essentialist practice and its open-door policy that welcomed whites.[103]

LatCrit developed in part in response to the conflicts in the workshop over two primary issues.[104] According to Phillips, the first issue, erupting during the 1990 workshop, was whether CRT's commitment to

---

97. The reason for the cessation of the CRT workshops, as far as I can tell, was that the workshop lacked a firm institutional framework to perpetuate its continuation. CRT scholars attending a workshop would be asked or volunteer to host the next workshop at their school the following year. A committee would then be established to guide the program. This differed from the process eventually established by LatCrit conference where host were identified and secured two years in advance and worked with standing officers and officials.

At the 1997 CRT workshop, Robert Westley, a coordinator of the workshop together with Sumi Cho, approached Stephanie Phillips about hosting the CRT workshop for the following year. I, who was attending the CRT workshop for the first time and was at that time an adjunct faculty member at the University at Buffalo Law School, encouraged her. Stephanie, however was always reluctant to host the conference again, having hosted a workshop in the early nineties. I did not push her on this and later went on to host another project. Thus, we unwittingly became part of the story of the workshop's demise.

98. Phillips, *supra* note 4, at 1249 n.4.

99. *Id.*

100. *Id.*

101. *Id.*

102. *Id.*

103. Crenshaw, *supra* note 3, at 1362-63; Phillips, *supra* note 4, at 1249 n.4.

104. Phillips, *supra* note 4, at 1251-52; *see generally* Francisco Valdes, *Foreword: Latina/o Ethnicities, Critical Race Theory, and Post-Identity Politics in Postmodern Legal Culture: From Practices to Possibilities*, 9 LA RAZA L.J. 1 (1996) [hereinafter Valdes, *Ethnicities*]; *see also* Francisco Valdes, *Under Construction - LatCrit Consciousness, Community, and Theory*, 85 CAL. L. REV. 1087, 1090 (1997) [hereinafter Valdes, *Under Construction*].

racial justice included a commitment to justice and liberation from other forms of oppression, particularly the oppression of gay, lesbian, and transgendered people.[105] It took, she suggests, almost eight years for the workshop to fully embrace the position that antiracial struggle does or should include the fight against oppression of sexual minorities.[106] The second issue involved many of the nonblack people of color participants at the 1992 workshop challenging the workshop's almost exclusive focus on the history and conditions of African Americans to the exclusion of the conditions of nonblack/nonwhite people. Phillips suggests that this was one of the earliest critiques of what is typically known as the black-white paradigm, [107] but which came to be called the white over black paradigm[108] to emphasize the hierarchal ordering of race and alluding to its inclusion of other groups in between whites and blacks.[109] Unlike the issue of sexuality, she notes that the collective response of the workshop to this critique was confessed ignorance,[110] apologies and the embarrassment of some who perceived their actions as CRT having done to nonblack peoples of color what CLS did to them. She argues that the workshop thereafter began to explore the experiences of other racialized groups and came to agree that:

> [R]acism is not only historical slavery, Jim Crow laws and gerrymandered voting districts in the South; it is also immigration laws and internment camps; it is stolen land grants and silenced languages;

---

105. Phillips, *supra* note 4, at 1250.

106. *Id.* at 1250-51.

107. *Id.* at 1252. Many scholars have critiqued and examined the limits of the white/black or white over black paradigm. *See, e.g.*, Delgado, *The Current Landscape of Race*, *supra* note 17, at 1272; Rachel F. Moran, *Neither Black Nor White*, 2 HARV. LATINO L. REV. 61, 81-82 (1997); Juan F. Perea, *The Black/White Binary Paradigm of Race: The "Normal Science" of American Racial Thought*, 85 CAL. L. REV. 1213, 1220, 1254 (1997) (arguing that the black/white paradigm promotes the invisibility and marginalized Latina/o experiences). Devon W. Carbado has surveyed and examined a variety of these critiques. *See* Devon W. Carbado, *Critical Race Studies: Race to the Bottom*, 49 UCLA L. REV. 1283, 1305-12. I agree that there are limitations to the paradigm, it cannot possibly capture the many valences of race in America. *See* Athena D. Mutua, *Mapping Intellectual/Political Foundations and Future Self Critical Directions: Shifting Bottoms and Rotating Centers: Reflections on LatCrit III and the Black/White Paradigm*, 53 U. MIAMI L. REV. 1177, 1179-80 (1999) [hereinafter Mutua, *Shifting Bottoms*]. However, I see the black experience and blackness as central to and paradigmatic of a colorized racial system in which blackness or anything but blackness is the chief obsession of white power. *Id.* at 1181-82. *See also* Harris, *supra* note 25, at 916 ("[T]he Black/White paradigm may not accurately reflect racial demographics, because, in part, it does not seek to do so. Instead, it describes racial power."). Further, I see no problem in focusing on the African American experience, or the Latina/o experience, or any other group histories or experiences for that matter. It seems to me there is a difference between a practical focus - concentrating your attention or efforts on one set of experiences—and assuming that this one set of experiences is the whole or represent all there is. *See* Athena D. Mutua, *Theorizing Progressive Black Masculinities*, *in* PROGRESSIVE BLACK MASCULINITIES 34 (Athena D. Mutua ed., 2006) [hereinafter PROGRESSIVE BLACK MASCULINITIES] (discussing the multiple struggles inherent in a project to transform the structures of domination that order American society, and noting the difference between a practical focus on one struggle, and assuming that it is the only one as opposed to part of a larger struggle).

108. Phillips, *supra* note 4, at 1252.

109. Mutua, *Shifting Bottoms*, *supra* note 107, at 1189.

110. Phillips, *supra* note 4, at 1253.

> it is standardized tests based on standardized culture; it is invisibility and lost identity.[111]

The critique of the white over black paradigm opened the window of recognition into the ways that the racialization of other groups involved the interplay of laws and practices not readily exposed by looking at the domestic experiences and conditions of blacks in the United States. It thus, opened doors for CRT to examine the relationship between race and ethnicity, racism, and nativisim, and racism, nationalism, and colonialism.[112] It also brought into focus laws related to immigration, citizenship, foreignness, language, and assimilation, among other issues as they related to U.S. foreign policy, and transnational and international law.

These insights together with the way in which some Latina/o scholars, such as Francisco Valdes, experienced the workshop conflicts led them to institutionalize a separate space for the exploration of issues germane to Latina/o communities and Latina/o identity.[113] The insights also led to a call for (initially made by Robert Chang) and the subsequent development of a line of Asian American legal scholarship.[114] Further, these insights intermittently drew in scholars who understood the Native American experience both in terms of national oppression and racial oppression.[115] And finally it led to and encompassed work that explored whiteness as a practice of exclusion and genocide, as a hidden norm and as a site of unearned privilege by whites through the work of "white crits."[116]

The establishment of the LatCrit annual conference together with the development of other related scholarship raised concerns over the fragmentation of Critical Race Theory and the potential explosion of multiple identity categories and projects.[117] However, not only did these

---

111. *Id.* at 1254.

112. For those CRT theorists who were familiar with black nationalist discourses, for example, which understood black America as an internal colony and pushed Pan-Africanism, these theorists may have made the connections between racism, nativism, and colonialism, even in the context of the African American experience.

113. Valdes, *Ethnicities*, *supra* note 104, at 8-9.

114. *See* Chang, *supra* note 1, at 1247-49; Aoki, *supra* note 1, at 1476-79.

115. *See, e.g.*, ROBERT A. WILLIAMS, THE AMERICAN INDIAN IN WESTERN LEGAL THOUGHT: THE DISCOURSES OF CONQUEST (1990), *reprinted in part in* POWER, PRIVILEGE AND LAW, *supra* note 1, at 169-72; Robert A. Williams, *Documents of Barbarism: The Contemporary Legacy of European Racism and Colonialism in the Narrative Traditions of Federal Indian Law*, 31 ARIZ. L. REV. 237 (1987), *reprinted in* CRT: CUTTING EDGE, *supra* note 1, at 98-109.

116. *See, e.g.*, RICHARD DELGADO & JEAN STEFANCIC, CRITICAL WHITE STUDIES: LOOKING BEHIND THE MIRROR 316 (1997); RUTH FRANKENBERG, DISPLACING WHITENESS: ESSAYS IN SOCIAL AND CULTURAL CRITICISM (1997); PEGGY MCINTOSH, WHITE PRIVILEGE AND MALE PRIVILEGE: A PERSONAL ACCOUNT OF COMING TO SEE CORRESPONDENCES THROUGH WORK IN WOMEN'S STUDIES, *reprinted in* POWER, PRIVILEGE AND LAW, *supra* note 1, at 23; DAVID R. ROEDIGER, THE WAGES OF WHITENESS: RACE AND THE MAKING OF THE AMERICAN WORKING CLASS (1991); Barbara Flagg, *"Was Blind, But Now I See": White Race Consciousness and the Requirement of Discriminatory Intent*, 91 MICH. L. REV. 953, 956-57 (1993).

117. *See, e.g.*, DELGADO & STEFANCIC, CRT: AN INTRODUCTION, *supra* note 1, at 6; Crenshaw, *supra* note 3, at 1364. *See generally*, Bryant, *supra* note 7.

developments spur and contribute many important and necessary intellectual insights, they expanded CRT's analytical framework to include other racialized or otherwise oppressed groups' experiences.[118] It also expanded CRT's commitment to the liberation of black and other oppressed people of color to include a commitment to the liberation of all oppressed and subordinated peoples.

Though critical race scholarship continued to grow through individual scholarship and group-initiated symposia, the founding of LatCrit filled an institutional gap left by the cessation of the annual CRT workshop. Even though various iterations of CRT workshops have since been held sporadically, another more permanent institutional framework for promulgating Critical Race Theory has not developed.[119]

D. UCLA, Proposition 209 and the Entrenchment of CRT: Confronting the Whiteness of Colorblindness Again

CRT continues to become more entrenched in the legal academy. Many law schools now offer courses in CRT.[120] In 2002, the University of California Los Angeles (UCLA) established the first concentration in CRT offered by an elite law school.[121] This event, like the birth of CRT, owes its establishment to the racialized conflict over affirmative action generated by California's Proposition 209 and the ways in which this played out in the institutional context of UCLA. Proposition 209 made illegal the consideration of race in California schools' admissions policy.[122] The result was a significant decline in minorities attending UCLA.[123] The Proposition appealed to and enforced an ideology of co-

---

118. One tension that arose with this expansion was whether critical race theory should maintain the black experience as a central focus or whether the CRT label should represent and act as an umbrella group for all of the different racial and other antisubordination projects. This idea was broached at the third annual conference of LatCrit where a panel explored whether a separate "black crit" enterprise should be established and devoted to the African American experience, and the experiences of others perceived as racially black such as black Latinos. Phillips, *supra* note 4, at 1251 n.10. Phillips suggested CRT be an umbrella group but seemed simultaneously opposed to establishing yet another separate critical race/black enterprise. *See* Phillips, *supra* note 4, at 1254-55. As a practical matter, CRT seems to serve as an umbrella label but one that is often qualified with the term "related scholarship," an approach I have tentatively adopted in this piece. *See, e.g.*, Harris, *supra*, note 1, at 1215 (also using these terms). At the same time, CRT has maintained a central focus on the African American experience, with analysis of other racialized experiences captured by the labels of LatCrit, Asian American Legal Scholarship, etc. This centrality, in part is due to how CRT, though always multicultural in membership, developed but also because the African American experience is often viewed as paradigmatic of race in the U.S. *See supra* note 107.

119. For instance, scholars have held subsequent iterations of the workshop at the American University Washington College of Law. These were distinct from the Critical Race Theory Conference held at Yale in 1997.

120. In a survey conducted by LatCrit in 2002, some 23 law schools (out of approximately 180 ABA approved law schools) had courses called or related to Critical Race Theory. *See* Robert S. Chang, *"Forget the Alamo": Race Courses as a Struggle over History and Collective Memory*, 13 BERKELEY LA RAZA L.J. 113 (2002) (LatCrit 2002 symposium).

121. Harris, *supra* note 1, at 1215-16.

122. *Id.* at 1221-25.

123. *Id.* at 1223-25.

lorblindness.[124] But as Cheryl Harris notes, "the ideology . . . could not hide from view what was before our very eyes; racial diversity was eroding."[125] In fact, the imposition of colorblind rules in the community context of UCLA revealed, what much of CRT implies—that, given the normative and social structure of the United States, colorblindness is a proxy for whiteness. Harris describes UCLA after Proposition 209:

> Given the fact that the school is physically located in southern California—an area teeming with racial complexity—the virtual absence of the full range of diversity within its walls is a constant and stark reminder of the entrenched nature of racial difference in terms of geography, educational opportunity, and access. Admission into the law school community is defined and constituted by rules that capture and reinforce certain background difference and inequalities, particularly those regarding race and class.[126] . . . However, colorblindness does not in fact ignore race; it rests upon and reflects an investment in a particular conception of race in which race is divested of its historical, societal, or experiential meaning.[127]

In reaction, the UCLA faculty, as Cheryl Harris explains, after serious debate, decided to establish a CRT concentration called Critical Race Studies.[128] They did so in part to signal UCLA's continuing commitment to racial equality despite Proposition 209.[129] Thus, while proponents of 209 pushed color-blindness as an appropriate approach to race, they made the color of colorblindness clear—whiteness; and further entrenched critical race consciousness of difference and Critical Race Theory in the legal academy.

IV. CRT TENETS AND METHODOLOGY

The basic tenets of Critical Race Theory remain true to the original ideas discussed in the 1990 CRT workshop. With little modification, Critical Race Theory:

> 1. holds that racism is pervasive and endemic to, rather than a deviation from, American norms;[130]
>
> 2. [rejects] dominant claims of meritocracy, neutrality, objectivity and color-blindness;

---

124. *Id.* at 1229-30.
125. *Id.* at 1230.
126. *Id.* at 1229.
127. *Id.*
128. *Id.* at 1232-34.
129. *Id.* at 1230.
130. As W.E.B. Du Bois wrote, "the problem of the twentieth century is the problem of the color-line." W.E.B. DU BOIS, THE SOULS OF BLACK FOLK 7 (Henry Louis Gates Jr. & Terri Hume Oliver eds., 1999) (1903).

> 3. [rejects] ahistoricism, and insists on contextual, historical analysis of law;
>
> 4. challenges the presumptive legitimacy of social institutions;
>
> 5. insists on recognition of both the experiential knowledge and critical consciousness of people of color in understanding law and society;
>
> 6. is interdisciplinary and eclectic (drawing upon, *inter alia*, liberalism, poststructuralist, feminism, Marxism, critical legal theory, postmodernism, and pragmatism) with the claim that the intersection of race and the law overruns disciplinary boundaries; and
>
> 7. works toward the liberation of people of color as it embraces the larger project of liberating all oppressed people.[131]

The purpose of CRT, its *raison d'etre*, is twofold. First its purpose is to demonstrate the many ways in which white supremacy is endemic to American society by "exposing the facets of law and legal discourse that create racial categories and legitimate racial subordination."[132] Second, its purpose is to destabilize and change this relationship, in part by challenging or proposing alternative laws, among other things, in order to contribute to the liberation of oppressed people. As Jerome Culp notes, Critical Race Theory may mean many different things to different people, but "there is a common belief in an opposition to oppression."[133] And, CRT scholars such as Matsuda and Hutchinson, as well as LatCrit and others have issued a clarion call that *Antisubordination,* a stance against all forms of oppression and subordination, be both the commitment of race scholars and the *principle* upon which racial justice, particularly *equality*, be understood and practiced.[134]

These tenets and the overall commitment to antisubordination that CRT scholars evidence also provide crucial insight into CRT methodological tendencies. CRT is said to have no single, unifying methodology.[135] Rather it is eclectic, drawing from various schools, disciplines and approaches.[136] Harris, in providing some examples of the different methodologies employed by Race Crits, notes that they include structur-

---

131. Phillips, *supra* note 4, at 1249-50.

132. CRT: KEY WRITINGS, *supra* note 1, at xiii.

133. Jerome McCristal Culp., Jr., *To the Bone: Race and White Privilege*, 83 MINN. L. REV. 1637, 1638 (1999).

134. *See* Mari J. Matsuda, *Voices of America: Accent, Antidiscrimination Law, and a Jurisprudence for the Last Reconstruction*, 100 YALE L.J. 1329, 1329-1404 (1991), *reprinted in part in* RACE AND RACES: CASES AND RESOURCES FOR A DIVERSE AMERICA, *supra* note 1, at 557-61; *see also* Darren Lenard Hutchinson, *"Unexplainable on Grounds Other Than Race": The Inversion of Privilege and Subordination in Equal Protection Jurisprudence*, 2003 U. ILL. L. REV. 615, 622.

135. *See e.g.*, Harris, *supra* note 1, at 1218; *infra* note 146 (Dorothy Brown making a similar point).

136. Harris, *supra* note 1, at 1217-18.

alism and historical, doctrinal (legal), empirical and economic analyses.[137]

CRT approaches can, however, be said to possess some unifying themes or methodological tendencies. These include a particular focus on context and history.[138] CRT suggests that a rule or principle may mean different things in different contexts and/or historical periods. So, for instance, they have argued that the idea of colorblindness, first expressed in Justice Harlan's 1896 dissent in *Plessy v. Ferguson*, can be understood at that time as a progressive idea in the context of a society in which law sanctioned the explicit and systematic oppression of blacks after slavery. However, a colorblind approach to race in the current era, when the subordination of blacks is no longer explicit but remains systematic, is no longer a progressive approach.[139] Thus, as an abstract principle its meaning and progressive potential is neither universal nor trans–historical. CRT, therefore, pays particular attention to the specificity of context in order to understand the meanings of a particular concept or practice, to evaluate a particular position and to render additional information and ideas.

Further, CRT argues that as rules and principles mean different things in different contexts that they *should* mean different things in different contexts. So for instance, equality might mean symmetrical or "same treatment" in a society without vast racial, gender, and class inequalities but might mean and require affirmative practices to bring about equality for historically disadvantaged groups, treating them differently than the privileged, in a society with these alarming disparities.[140]

In addition, CRT scholars listen to and scrutinize the voices, understandings and experiences of marginalized and oppressed peoples to situate, test, and inspire the examination of particular and/or novel approaches to law.[141] The idea of distinctive minority voices recognizes, for example, that not every Native American critiques the American holiday "Columbus Day."[142] But it does understand that, given Native

137. *Id.* at 1218 n.6.

138. *Id.* at 1229; *see* Houh, *supra* note 43, at 1061-62.

139. *But see* Gotanda, *supra* note 2, at 257.

140. *See, e.g.*, Hutchinson, *supra* note 134, at 646 (discussing equality as not symmetrical).

141. *See e.g.*, Mari Matsuda, *Looking to the Bottom: Critical Legal Studies and Reparations*, 22 HARV. C.R.-C.L. L. REV. 323, 324 (1987) (urging CLS to draw on the experiences and writings of marginalized groups to inform their theory).

142. In 1992, the five hundredth anniversary of Columbus' arrival in the Western Hemisphere was marked by Indian protests across the country. MATTHEW DENNIS, RED, WHITE, AND BLUE LETTER DAYS: AN AMERICAN CALENDAR 154-55 (2002). Columbus Day had not been officially celebrated until the tercentenary in 1892, one hundred years earlier, when "some – especially African-Americans – began to contest the Columbus celebrations, not so much the accepted view of Columbus the Man but rather the image of Columbia the land of freedom, opportunity and progress." *Id.* at 148. *See generally* IDA B. WELLS, FREDERICK DOUGLASS, IRVINE GARLAND PENN, & FERDINAND L. BARNETT, THE REASON WHY THE COLORED AMERICAN IS NOT IN THE WORLD'S COLUMBIAN EXPOSITION (Robert W. Rydell ed., 1999); Sylvia Wynter, *1492: A New World View, in* RACE, DISCOURSE AND THE ORIGIN OF THE AMERICAS: A NEW WORLD VIEW 5-57 (Vera Law-

American history, the conditions of oppression, and the cultural nature of their resistance, Native Americans might find the idea of Columbus *discovering* America problematic, and not exactly a cause for celebration.[143] This understanding has led CRT scholars to excavate forgotten or overlooked histories, rules, and cases, as well as the cultural practices, stories, and perspectives of marginalized groups as sources for grounding their analysis.

In this vein, race crits have often successfully employed storytelling or narrative to explore alternative meanings, insights, and perspectives on an issue. Some of the leading legal storytellers include Derrick Bell (*And We Are Not Saved* (1989), *Faces at the Bottom of the Well* (1993), and *Gospel Choirs* (1996)), Richard Delgado in his *Rodrigo* series (1996) and *When Equality Ends* (1999), and Patricia Williams in *The Alchemy of Race and Rights* (1991).[144] Legal storytelling has garnered significant critique, including criticisms (1) that such stories are not subject to empirical or other typical methods of evaluation; (2) that challenge the idea of a particular minority voice or perspective; and (3) that charge that such stories tend to distort the "truth,"[145] a truth understood by many CRT theorists, as simply the common sense understandings that arise under the current hegemonic ideologies and practices. It has further led to the suggestion as Dorothy Brown points out, that CRT stands against empiricism as a form of argumentation and verification because it refutes narrative.[146] This idea is buttressed by CRT's embrace of CLS'

---

rence Hyatt & Rex Nettleford eds., 1995) (for a critique of Columbus and what he stands for from a modern black perspective).

143. *See* James Barron, *He's the Explorer/Exploiter You Just Have to Love/Hate*, N.Y. TIMES, Oct. 12, 1992, at B1.

144. *See generally* Bell, *supra* note 59, at 20-29; Bell, *supra* note 65; DERRICK BELL, GOSPEL CHOIRS: PSALMS OF SURVIVAL FOR AN ALIEN LAND CALLED HOME (1996); RICHARD DELGADO, THE RODRIGO CHRONICLES: CONVERSATIONS ABOUT AMERICA AND RACE (1996), *reprinted in part in* CRT: CUTTING EDGE, *supra* note 1, at 388-396; RICHARD DELGADO, WHEN EQUALITY ENDS: STORIES ABOUT RACE AND RESISTANCE (1999); WILLIAMS, *supra* note 96; DELGADO & STEFANCIC, CRT: AN INTRODUCTION, *supra* note 1, at 37-49; CRT: CUTTING EDGE, *supra* note 1, at 41-91; Margaret E. Montoya, *Celebrating Racialized Legal Narratives*, *in* CROSSROADS, DIRECTIONS, AND A NEW CRITICAL RACE THEORY, *supra* note 1, at 243-301; Michael A. Olivas, *The Chronicles, My Grandfather's Stories, and Immigration Law: The Slave Traders' Chronicle as Racial History*, 34 ST. LOUIS U. L.J. 425 (1990), *reprinted in* CRT: CUTTING EDGE, *supra* note 1, at 9-20; Thomas Ross, *The Richmond Narratives*, 68 TEX. L. REV. 381 (1989), *reprinted in* CRT: CUTTING EDGE, *supra* note 1, at 42-51; William N. Eskridge, Jr., *Gaylegal Narratives*, 46 STAN. L. REV. 607 (1994); Lucie E. White, *Subordination, Rhetorical Survival Skills, and Sunday Shoes: Notes on the Hearing of Mrs. G*, 38 BUFF. L. REV. 1 (1990), *reprinted in part in* POWER, PRIVILEGE AND LAW, *supra* note 1, at 565-81; Rachel F. Moran, *Full Circle*, *in* CRITICAL RACE FEMINISM, *supra* note 37, at 113-17; CARL GUTIÉRREZ-JONES, CRITICAL RACE NARRATIVES: A STUDY OF RACE, RHETORIC, AND INJURY (2001).

145. Daniel A. Farber & Suzanna Sherry, *Telling Stories Out of School: An Essay on Legal Narratives*, 45 STAN. L. REV. 807 (1993). *But see* Jane B. Baron, *Resistance to Stories*, 67 S. CAL. L. REV. 255 (1994); Richard Delgado, *On Telling Stories in School: A Reply to Farber and Sherry*, 46 VAND. L. REV. 665 (1993).

146. Dorothy A. Brown, in the Symposium, *Critical Race Theory: The Next Frontier: Fighting Racism in the Twenty-First Century*, 61 WASH. & LEE L. REV. 1485 (2004), suggested that CRT rejects numbers as neutral, that the privileging of numbers refutes narrative, and thus empirical research may be incompatible with CRT. *Id.* at 1486-87. But she argued that empirical evidence is

critique of the Law and Economics movement,[147] a scholarly tendency that often employs empirical data. While most law is viewed from the liberal perspective of individual rights perceived as neutral and objective, the Law and Economics School, like CLS, is critical of that approach. Law and Economics, however, is a conservative approach, which according to crits, simply replaces law's claims of impartially and neutrality with similar claims for the field of economics.[148] Economics, however, is neither neutral nor objective. Rather, it too involves political choices both at the level of practice and study. And, arguably, both are replete with the values, assumptions, presumptions and dictates about human behavior and the operation of society as determined and understood by the current economic order of capitalism.[149]

Harris, however, notes that CRT scholars have employed both empirical data and economic analysis.[150] Nevertheless, an argument based on empiricism or economic analysis, according to CRT, is just that, a form of argumentation in which political choices are made as to what should be included or excluded and what is important or not, as well as how the facts or statistics should be interpreted. Similar empirical data could presumably be used, like various rules, to support contrary and alternative arguments and interpretations.[151]

---

necessary to reach out to white America. *Id.* at 1489. Darren Hutchinson argues that CRT theorists usually rely on law and legal reasoning, but could buttress their arguments by also relying on political science data that, for instance, have used polling to demonstrate that the Supreme Court largely responds to majoritarian concerns in its decision-making and facilitates majoritarian interests. Darren Lenard Hutchinson, *Critical Race Theory: History, Evolution, and New Frontiers: Critical Race Histories: In and Out*, 53 AM. U. L. REV. 1187, 1213-14 (2004). That was also a common theme in the public law subfield of political science. *See generally* ROBERT G. MCCLOSKEY, THE AMERICAN SUPREME COURT (2d ed. 1994); MARTIN SHAPIRO, LAW AND POLITICS IN THE SUPREME COURT: NEW APPROACHES TO POLITICAL JURISPRUDENCE (1964); 1 CORWIN ON THE CONSTITUTION (Richard Loss ed., 1981); 2 CORWIN ON THE CONSTITUTION (Richard Loss ed., 1987); 3 CORWIN ON THE CONSTITUTION (Richard Loss ed., 1988); Robert A. Dahl, *Decision-Making in a Democracy: The Role of the Supreme Court as a National Policy-Maker*, 6 J. PUB. L. 279 (1957).

147. *See generally* RICHARD I. IPPOLITO, ECONOMICS FOR LAWYERS (2005); RICHARD A. POSNER, ECONOMIC ANALYSIS OF LAW (6th ed. 2002); GORDON TULLOCK, LAW AND ECONOMICS (Charles K. Rowley ed., 2005).

148. Mark G. Kelman, *Critical Legal Studies Symposium: Trashing*, 36 STAN. L. REV. 293, 294 (1984); Mark G. Kelman, *Misunderstanding Social Life: A Critique of the Core Premises of "Law and Economics,"* 33 J. LEGAL EDUC. 274, 274 (1983); Duncan Kennedy, *Cost-Benefit Analysis of Entitlement Problems: A Critique*, 33 STAN. L. REV. 387, 387 (1981); Duncan Kennedy & Frank Michelman, *Are Property and Contract Efficient?*, 8 HOFSTRA L. REV. 711, 713 (1980).

149. James Boyd White, *Economics and Law: Two Cultures in Tension*, 54 TENN. L. REV. 161, 176 (1986).

150. Harris, *supra* note 1, at 1218 n.6. *See, e.g.*, Devon W. Carbado & Mitu Gulati, *The Law and Economics of Critical Race Theory: Crossroads, Directions, and a New Critical Race Theory*, 112 YALE L.J. 1757, 1757 (2003) (book review), *reprinted in part in* A WOMAN'S PLACE, *supra* note 43, at 140-53; Patricia J. Williams, *Spare Parts, Family Values, Old Children, Cheap*, 28 N. ENG. L. REV. 913, 914 (1994) (critiquing Elizabeth M. Landes & Richard A. Posner, *The Economics of the Baby Shortage*, 7 J. LEGAL STUD. 323 (1978)); William Bratton, *Anti-Subordination and the Legal Struggle Over the "Means of Communication": Law and Economics of English Only*, 53 U. MIAMI L. REV. 973, 973-74 (1999).

151. *See supra* note 148; *see also* Harris, *supra* note 25, at 907-14 (reviewing the empirical study undertaken in the book *Whitewashing Race*, *see* BROWN ET AL., WHITEWASHING RACE, *supra*

Finally, the tendency of CRT to deconstruct and expose the racial meanings of law betrays its post-modern sensibilities and contradicts its commitment to modernist ideals of justice, truth, and dignity[152] (because as race crits and other post-modern scholars might argue a principle in one context may mean something radically different in another). However, in view of the dual vision that W.E.B. Du Bois located in the oppressed, and the dual respect and disdain that oppressed people have shown for the law,[153] scholars have encouraged race crits to inhabit the tension between its post-modern insights and its modernist ideals of justice because within that tension lie CRT's creative potential.[154]

V. LAW AND THE CONSTRUCTION OF RACE

Below I provide a brief structural history of the relationship between race and law focusing primarily on case law, and drawing on and applying some of the insights and methodological tendencies of Critical Race Theory. Specifically, I employ historical analysis and narrative to tell the story of this relationship between law and race. The narrative emphasizes the longevity of American racial ordering and practice, the breadth of that racial ordering, and its depth in regulating American life. It does so to demonstrate the ways in which law both constructs and produces races and racism and the deeply structured nature of race in U.S. society. Further, it is meant to show, focusing on the United States Supreme Court, how the law's increasing reliance on colorblind individualism works to maintain, rather than undermine the racial caste system created over several hundred years.[155] In doing so, it provides a counter-narrative to the dominant and ever popular story about race and law that

---

note 25, and arguing that although the study thoroughly contested the colorblind conception of racism as the function of individual bigoted action, empirical evidence alone is insufficient to dislodge the theory). Harris, examining the various interpretations of whether race played a role in aftermath of Hurricane Katrina, explains that empirical evidence was not enough, in part, because colorblindness is a well-funded ideology promoted over the last several decades; but primarily because "the ability to process empirical facts into a different understanding [is] . . . compromised by the divergent . . . perspectives through which the facts are viewed." *See* Harris, *supra* note 25, at 913. She argues that there are frames or frameworks, in this case racial frames, which allow us to make sense of facts; they are "what lies between the facts and our perceptions - the mediating structures that allow us to make sense of the world." *Id.* at 914. Regarding Katrina she ventures:

> ...while people of all races agree that Katrina exposed the social costs of poverty, most Whites consider race largely irrelevant in explaining what happened (or did not happen) while Blacks tend to view race as a crucial part of the story. In the aftermath of Katrina, the question that is being debated is less a matter of what happened - what is at issue is why - and here the absence of a consensus demonstrates how racial divisions in the interpretation of seemingly uncontested facts can result in entirely divergent assessments of causation. The facts in this case did not seem to lead to a new racial paradigm; indeed, initial differences in the perceptions of the salience of race seemed to persist, notwithstanding relative agreement on the facts.

Id. at 913.

152. *See generally* Harris, *supra* note 42, at 743.

153. WILLIAMS, *supra* note 96, at 35-36; *see* DU BOIS, *supra* note 130, at 3.

154. Harris, *supra* note 42, at 760, 778 (suggesting that in inhabiting this tension, CRT race-crits aspire to and attempt to make real the dreams that modernity promised).

155. DELGADO & STEFANCIC, CRT: AN INTRODUCTION, *supra* note 1, at 21-22.

suggests that the struggle for racial justice, though long and incremental, is nevertheless forward-moving, progressive, and eventually triumphant, given the American creed and precepts.[156] Instead it suggests the stagnation of racial progress because of the continuity of the underlying structures of white supremacist thought, operation, and social arrangements, though accomplished through new and changing forces and rationalizations.

A. Early Construction of Race by Law

Throughout most of American history, legislators, legal practitioners, and judges have made and interpreted various legal doctrines, rules, and procedures to define, construct, produce, and preserve white privilege and black subordination, as well as the subordination of other people of color. Throughout most of its history, American law has been decidedly race conscious and specifically white supremacist whenever it has encountered what it, itself, has often defined as Other.

For instance, in order to perpetuate a white state, judges defined whiteness through case law to determine whether a Japanese man was white for purposes of citizenship, whether a Chinese person was black or Indian for the purpose of determining whether he could testify against a white, and whether Mexicans were white and thus entitled to serve on juries.[157] American law facilitated, defined, and established white privileges by limiting the rights of Indians to their land and facilitating white appropriation of the same land.[158] It did so using slave laws, black codes, and Jim Crow laws to exploit Black labor and maintain Black subordination for the purposes of white wealth accumulation and white racial class consolidation.[159] It has constructed race for the purposes of determining who might vote, the manner in which those who presumably were entitled to vote could do so, and whether such people could actually and effectively vote.[160] It delineated a range of businesses practices af-

156. *See* DELGADO & STEFANCIC, CRT: AN INTRODUCTION, *supra* note 1; BELL, *supra* note 2, at 22.

157. Hernandez v. Texas, 347 U.S. 475, 477, 482 (1954); Ozawa v. United States, 260 U.S. 178 (1922); People v. Hall, 4 Cal. 399, 399 (Cal. 1854); *see also* Neil Gotanda, *Comparative Racialization: Racial Profiling and the Case of Wen Ho Lee*, 47 UCLA L. REV. 1689, 1695 (2000) (comparing *People v. Hall* and the *Dred Scott* case); Miguel A. Méndez, *Hernandez: The Wrong Message at the Wrong Time*, 4 STAN. L. & POL'Y REV. 193 (1993), *reprinted in part in* THE LATINO/A CONDITION, *supra* note 1, at 602-04.

158. Johnson v. M'Intosh, 21 U.S. 543, 587-88 (1823); *see also* Eric Kades, *History and Interpretation of the Great Case of* Johnson v. M'Intosh, 19 LAW & HIST. REV. 67, 67 (2001).

159. Plessy v. Ferguson, 163 U.S. 537, 559-60 (1896); Scott v. Sandford, 60 U.S. 393, 404-06 (1857).

160. TEX. REV. CIV. STAT. ANN. ART. 3107 (Vernon 1925). In "1924, the Texas statute, Art. 3093a, afterwards numbered Art. 3107 (Rev. Stat. 1925) declared 'in no event shall a Negro be eligible to participate in a Democratic Party primary election in the State of Texas.'" Smith v. Allwright, 321 U.S. 649, 658 (1944) (citing Nixon v. Herndon, 273 U.S. 536 (1927)). In *Breedlove v. Suttles*, 302 U.S. 277, 280-281 (1937), the Supreme Court upheld a Georgia provision requiring payment of a poll tax as a prerequisite for voting. Though a white man brought the case, the estab-

fecting everyone from laborers to professionals including who could be treated by a doctor or a nurse. It has used racial categories to the detriment of people of color to determine questions concerning where people can live,[161] who they can marry,[162] what schools they can attend,[163] and where they sit on a train, and in a cafeteria, or theater.[164]

From a CRT perspective, the movement from overt racial oppression sanctioned by law to law's racial neutrality or colorblindness has done little to undo the systemic and accumulated conditions of racial oppression created by and through law over several hundred years. For that matter, except for two very brief though significant periods, American law has been and remains a bulwark of white supremacy. The promulgation of the Emancipation Proclamation, and the ratification of the thirteenth, fourteenth and fifteenth amendments,[165] should have disrupted

---

lishment of poll taxes was one of the many ways in which blacks were prohibited from voting. *Breedlove* was overruled by *Harper v. Va. State Bd. of Elections*, 383 U.S. 663 (1966).

161. Kraemer v. Shelley, 198 S.W.2d 679 (Mo. 1946) (upholding restrictive covenants in housing). *Kraemer* was overturned by the United States Supreme Court in *Shelley v. Kraemer*, 334 U.S. 1, 23 (1948). See also Hurd v. Hodge, 334 U.S. 24, 33-34 (1948) (companion case invalidating racially restrictive covenants in the District of Columbia—to which the fourteenth amendment did not apply—by applying the Civil Rights Act of 1866); Barrows v. Jackson, 346 U.S. 249, 258 (1953) (damages not awardable for breach of a racially restrictive housing covenant). *But see* BELL, *supra* note 2, at 294-96 (describing subsequent history and current problems of housing discrimination). *See generally* CLEMENT E. VOSE, CAUCASIANS ONLY: THE SUPREME COURT, THE NAACP, AND THE RESTRICTIVE COVENANT CASES (1959). Three Justices did not participate in Shelley and Hurd, reportedly because they owned property subject to racially restrictive covenants. C. HERMAN PRITCHETT, CIVIL LIBERTIES AND THE VINSON COURT 142 (1954). For other discussions of housing segregation, see Calmore, *supra* note 41; DOUGLAS S. MASSEY & NANCY A. DENTON, AMERICAN APARTHEID: SEGREGATION AND THE MAKING OF THE UNDERCLASS (1993); DAVID DELANEY, RACE, PLACE, AND THE LAW, 1836-1948 (1998).

162. Jackson v. State, 72 So. 2d 114 (1954). Bell notes that "[a]ccording to one study, 38 states had miscegenation statutes at one time or another during the nineteenth century, and as late as 1951, 29 statutes were still on the books." BELL, *supra* note 2, at 256. The United States Supreme Court struck down these laws in *Loving v. Virginia*, 388 U.S. 1 (1967). *See* RACHEL F. MORAN, INTERRACIAL INTIMACY: THE REGULATION OF RACE AND ROMANCE 6 (2001).

163. Sweatt v. Painter, 210 S.W.2d 442, 444 (Tex. Civ. App. 1948), *rev'd*, 339 U.S. 629, 635-636 (1950) (ruling that a black student could attend a white university). However, while the Court granted relief in this case it did not address the constitutionality of the separate but equal doctrine. Rather this doctrine was finally overturned in *Brown*.

164. *See* Plessy v. Ferguson, 163 U.S. 537, 545, 548 (1896).

165. U.S. CONST., amend. XIII-XV. It should be noted that between 1777 and 1817, slavery had been abolished in the Northern states by, variously, constitutional provisions, constitutional interpretation, judicial decisions, and most often by gradual emancipation statutes. *See* LEON F. LITWACK, NORTH OF SLAVERY: THE NEGRO IN THE FREE STATES, 1790-1860 3-20 (1961); BELL, *supra* note 2, at 22-23; ARTHUR ZILVERSMIT, THE FIRST EMANCIPATION: THE ABOLITION OF SLAVERY IN THE NORTH 169-89 (1967). Slavery had already been abolished by judicial decision in England—but not in its colonies—in 1772. Somersett v. Stewart, 98 Eng. Rep. 499 (K.B. 1772); *see also* A. LEON HIGGINBOTHAM, JR., IN THE MATTER OF COLOR: RACE AND THE AMERICAN LEGAL PROCESS: THE COLONIAL PERIOD 313 (1978). The American national government, however, protected slavery through the fugitive slave clauses of the Articles of Confederation, the Northwest Ordinance, and the U.S. Constitution, among other provisions, although these instruments studiously avoided using the word. Articles of Confederation, art. 6 (1781); Northwest Ordinance, art. 6 (1787); U.S. CONST., art. IV, §2, cl. 3. Other provisions of the original Constitution implicitly upheld slavery. U.S. CONST., art. I, § 2, cl. 3 (slaves counted as three-fifths the number of free persons for purposes of apportionment of representation in the House of Representatives); U.S. CONST., art. I, § 9, cl. 1 (no Congressional prohibition of the importation of slaves before 1808). Generally, courts enforced the fugitive slave clause despite its increasing unpopularity in the North.

the legal construction of white racial supremacy and nonwhite racial subordination, but ultimately failed to do so. Nonetheless, wrenched from a civil war, these laws helped radically change the status of most blacks from slaves to free people and might have held the promise of providing the economic, social, and political rewards of citizenship and belonging.[166] But the promise of Reconstruction was short-lived, as the nation's political leadership capitulated to the southern elites' efforts to reassert white control over black life.[167] Various legal actors including judges, enacted laws and interpreted legal doctrine, to narrow the rewards of citizenship based on race helping to legitimate the Jim Crow era of racial segregation. Significant among these developments was the Supreme Court's interpretation of the Equal Protection Clause of the fourteenth amendment in *Plessy* permitting the segregation and subordination of black people.

The legal strategies of the NAACP and others beginning in the 1920's[168] and culminating in decisions such as *Hernandez* and *Brown,* helped to spark the civil rights movement, renewing the promise of equality in the 1960's. It was successful to the extent that it unraveled the explicit manifestations of racial ordering.[169] Crucial to this limited success was the Supreme Court's reinterpretation of equal protection as sanctioning the use of racial measures, such as affirmative action and school desegregation to potentially undo the racialized social and institutional patterns of oppression.[170]

See ROBERT M. COVER, JUSTICE ACCUSED: ANTISLAVERY AND THE JUDICIAL PROCESS 175-89 (1975).

166. *See* JOHN HOPE FRANKLIN, RECONSTRUCTION: AFTER THE CIVIL WAR 36-39 (1960); KENNETH M. STAMPP, THE ERA OF RECONSTRUCTION, 1865-1877 12 (1965); RECONSTRUCTION: AN ANTHOLOGY OF REVISIONIST WRITINGS 473-531 (Kenneth M. Stampp & Leon F. Lutwack eds., 1969); WILLIAM BROCK, CONFLICT AND TRANSFORMATION: THE UNITED STATES, 1844-1877 360-61 (1973); ERIC FONER, RECONSTRUCTION: AMERICA'S UNFINISHED REVOLUTION, 1863-1877 77-79 (1988); DAVID HERBERT DONALD, JEAN H. BAKER & MICHAEL F. HOLT, THE CIVIL WAR AND RECONSTRUCTION 536 (2001).

167. WILLIAM GILLETTE, RETREAT FROM RECONSTRUCTION, 1869-1879 3-7 (1979); C. VANN WOODWARD, REUNION AND REACTION: THE COMPROMISE OF 1877 AND THE END OF RECONSTRUCTION 3-4, 246 (1951); RECONSTRUCTION: AN ANTHOLOGY OF REVISIONIST WRITINGS, *supra* note 166, at 473-531. Whatever the reasons for the judiciary's rapid retreat from enforcing the constitutional and statutory law of Reconstruction, they did not include any overall aversion to so-called judicial activism. In this period, the courts regained the prestige and increased the power they enjoyed before the Dred Scott case. *See* STANLEY I. KUTLER, JUDICIAL POWER AND RECONSTRUCTION POLITICS 31-35 (1968); William F. Wiecek, *The Reconstruction of Federal Judicial Power,* 13 AM. J. LEGAL HIST. 333 (1969), *reprinted in part in* AMERICAN LAW AND THE CONSTITUTIONAL ORDER 237-45. (Lawrence M. Friedman & Harry N. Scheiber eds., 1988).

168. KLUGER, *supra* note 34, at 100-750; MARK TUSHNET, THE NAACP'S LEGAL STRATEGY AGAINST SEGREGATED EDUCATION, 1925-1950 xi (1987).

169. Robert Carter, *The Warren Court and Desegregation,* 67 MICH. L. REV. 237, 247 (1968) (Carter, a veteran NAACP litigator stated "[F]ew in the country, black or white, understood in 1954 that racial segregation was merely the symptom, not the disease; that the real sickness is that our society in all its manifestations is geared to the maintenance of white superiority").

170. *See, e.g.*, Swann v. Charlotte-Mecklenburg Bd. of Educ., 402 U.S. 1, 28 (1971); Brown v. Bd. of Educ. (*Brown II*), 349 U.S. 294, 300-301 (1955).

The highlights of this period were the passage of the Civil Rights Act of 1964, the Voting Rights Act of 1965, the Fair Housing Act of 1968, and the decision in *Griggs v. Duke Power Co.*[171] The latter case involved hiring practices, based on standards not related to job performance, which had the effect of disqualifying disproportionate numbers of blacks relative to whites. In *Griggs*, the Supreme Court took an approach to race (not the colorblind approach) that might have engendered actions that would result in actual changes in the conditions of black lives in terms of poverty, wealth accumulation, health, etc. This was so because it potentially rendered successful, suits brought on evidence of racial disparities and the impact of laws (demonstrated largely through statistical evidence of continuing disparities), instead of on proof of some individual's intention to discriminate.[172]

B. Colorblindness in Law Blocking Racial Progress

However, just as during and after Reconstruction,[173] courts began to narrow these laws, thereby stabilizing and legitimizing the prevailing order of white privilege and nonwhite disadvantage. So for example, though the desegregation (integration) of secondary schools in the United States had only begun in earnest in 1964 due to the Court's initial hesitancy and massive white resistance;[174] the Court had already signaled its retreat from desegregation by 1973 by allowing a school board to poten-

---

171. Civil Rights Act of 1964, 42 U.S.C.S. § 2000a (2006); Voting Rights Act of 1965, 42 U.S.C.S. § 1971 (2006); Fair Housing Act of 1968, 42 U.S.C.S. § 3601 (2006); Griggs v. Duke Power Co., 401 U.S. 424, 429-30 (1971) (construing Title VII of the Civil Rights Act of 1964).

172. *Cf.* McClesky v. Kemp, 481 U.S. 279, 329-330 (1987) (involving the Georgia sentencing system, which does not mention race (race-neutral) but has a disparate and disadvantaging impact on blacks). Georgia's capital sentencing system had been invalidated three times and Georgia had a long history of a "dual system." In the case, McCleskey, a black man, was sentenced to death for murdering a white police officer. *Id.* at 283. He challenged the death sentence in his case based on a sophisticated statistical study, finding that prosecutors in Georgia sought the death penalty in seventy percent of cases involving a black defendant and white victim, compared to only thirty-two percent of the time where both defendant and victim were white. *Id.* at 286-87. It also showed that the death penalty was assessed in twenty-two percent of cases involving black defendants and white victims, as compared to only 8 percent in cases involving both whites. *Id.* at 286. The Court rejected the use of the statistical reports to document discriminatory effect in the absence of proof of intentional discrimination. *Id.* at 297-99. The Court nevertheless stresses its "unceasing" efforts to overcome racism in the criminal justice system, *id.* at 309, but undercut this assertion near the outset of the case by limiting the examination to one of "purposeful discrimination." *Id.* 292. Thus, the Court's unceasing efforts to eradicate racial prejudice in the criminal justice system stopped at the door of proof of intent against and in an individual defendant's case without an examination of the systemic context in which the case was decided. *See* Hernandez, *supra* note 2, at 141-43 (discussing this case).

173. *See, e.g.*, The Civil Rights Cases, 109 U.S. 3, 23–26 (1883); The Slaughter-House Cases, 83 U.S. 36, 77-79 (1873).

174. *See, e.g.*, OLGETREE, *supra* note 34, at 10 (discussing the Court's "deliberate speed"); *see also* BROWN, *supra* note 30, at 167-74 (discussing the massive white resistance to Brown ranging from protest, and violence against black students attempting to attend white schools to white officials entirely eliminating public schools while publicly funding white children's private school attendance and including rapid and substantial movement of white families to the suburban areas in order to avoid sending their children to schools where the possibility of integration with blacks was slimmer).

tially escape the imposition of desegregation (integration) orders if it could show that segregation had occurred in the district owing to acts other than the board's *intentional* activities.[175] The Court held this despite the fact that the de facto practice of racial segregation in society was nearly universal.

In *Milliken v. Bradley*,[176] decided a year later, the Court dealt the "deathblow to the [nation's] ability to successfully integrate public schools."[177] It effectively made U.S. suburbs safe havens for whites who did not want their children to attend integrated schools with black children, and in doing so, the Court contributed to "white flight" to those suburbs.[178] Further, that blacks and other minorities would have to prove that some individual or institution had engaged in individual and identifiable discriminatory acts in order for the Courts to redress their grievances or remedy discrimination, was later confirmed in *Washington v. Davis*,[179] despite the pervasive and systemic presence of racism, racial segregation, and racial discrimination in the country.[180]

The Supreme Court, however, narrowed the laws, not only by focusing primarily on intentional, identifiable acts of discrimination, but also by focusing on the individual and by virtually banning the use of racial categories—arguably long before affirmative uses of these categories could affect any significant change in the social arrangements and structure of white racial power. In focusing on the individual, the courts

---

175. BROWN, *supra* note 30, at 208-10 (discussing Keyes v. School Dist. No. 1, Denver, Colo., 413 U.S. 189 (1973)). In *Keyes*, the plaintiffs proved the Denver county school board had operated to deliberately segregate schools in a core section of the city district in which over one-third of the black children attended, but the Court held the school board could rebut the presumption of system-wide segregation by proving that they had not intentionally segregated the schools system-wide. *Keyes*, 413 U.S. at 252-53.

176. 418 U.S. 717 (1974). *Milliken* was one of the first desegregation cases to be pursued in the North, where there had been few laws requiring segregation, but de facto segregation was prevalent. The plaintiffs sought a remedy for the segregation of Detroit schools, which would involve the adjoining, predominantly white suburban school districts. *Id.* The trial court had held that the State of Michigan was responsible for designing the school district system in which the city of Detroit was effectively racially segregated. *Id.* The Supreme Court held that, absent a showing that a constitutional violation in one district produced a significant desegregation effect in another, there could be no "cross-district remed[ies]." *Id.* at 744; *see* BROWN, *supra* note 30, at 213. According to Kevin Brown, the effect of the decision was that the suburbs were deemed safe havens by whites who could afford to move and who wanted to avoid integration. BROWN, *supra* note 30, at 210. Although the Keyes and Milliken cases represented the Court's initial retreat from school desegregation, Brown argues that the primary reason for the complete abandonment of desegregation throughout the 1980's and 1990's was the result of the application of the ideology of colorblind individualism to school desegregation cases.

177. BROWN, *supra* note 30, at 210-11.

178. BROWN, *supra* note 30, at 211.

179. 426 U.S. 229, 240 (1976).

180. Justice Thurgood Marshall—the first black U.S. Supreme Court Justice and the attorney that had argued *Brown* before the Court—responded to a similar idea in *Regents of the University of California v. Bakke*: "It is unnecessary in 20th-century America to have individual Negroes demonstrate that they have been victims of racial discrimination; the racism of our society has been so pervasive that none, regardless of wealth or position, has managed to escape its impact." 438 U.S. 265, 400 (1978).

were able to ignore or trivialize societal-wide racial ordering that burdened minorities as a group and individuals based on group membership, and pretend that racism was an unfortunate feature of the past only currently existing in the aberrant individual.[181] In taking a colorblind approach, the Court prohibited almost all uses of racial categories regardless of whether they were being used to subordinate blacks and other non-whites ("invidious discrimination") or to redress systemic white racial oppression of and on behalf of blacks and other nonwhites ("benign discrimination"). In essence, the Court increasingly applied the approach of colorblind individualism.[182] And, because *Brown* and its progeny had largely eliminated explicit racial subordinating laws, the focus and targets of this colorblind approach became the *remedies* and *measures* meant to address racial oppression (measures opposed mostly by whites as reverse discrimination), even as minorities have sought to expand these.

The Supreme Court has not, in so many words, declared colorblindness to be the new interpretation of equal protection, but rather it has largely accomplished it through various procedural and other standards drawing on its logic.[183] These include not only limiting remedies to intentional conduct, but also applying the Court's highest and toughest level of review, strict scrutiny now to all cases involving racial classifications, even if they are employed remedially to redress the consequences of earlier intentional discrimination. The Court's rationale is that "any official action that treats a person differently on account of his race or ethnic origin is inherently suspect," and that "[d]istinctions between citizens solely because of their ancestry are by their very nature odious to a free people."[184] The Court is motivated by a number of concerns, but seems particularly concerned about protecting "innocent people,"[185] a phrase used generally to refer to white people.[186]

---

181. *See infra* notes 193-99 and accompanying text (discussing Wygant v. Jackson Bd. of Educ., 476 U.S. 267, 276 (1986)); *see also, e.g.*, DELGADO & STEFANCIC, CRT: AN INTRODUCTION, *supra* note 1, at 7.

182. *See* BROWN, *supra* note 30.

183. Prior to 1995, the Supreme Court had been divided as to whether strict scrutiny should apply to all cases involving racial classifications. Strict scrutiny had been applied to "invidious" discrimination cases where discrimination built over centuries of practice was meant to disadvantage minorities. The Court, however, had applied a more lenient level of scrutiny to remedies that relied on race to benefit minorities and undo the past discrimination. This had been termed by the Court "benign" discrimination. However, in 1995 a solid majority of the Court held in favor of applying strict scrutiny to all uses of race-specific laws. *See* Adarand Constructors, Inc. v. Pena, 515 U.S. 200, 222 (1995).

184. *Adarand*, 515 U.S. at 223-25 (internal citations omitted). Strict scrutiny requires that the use of a racial classification be narrowly tailored to meet a compelling governmental interest.

185. *See, e.g.*, *Wygant*, 476 U.S. at 276; *Bakke*, 438 U.S. at 308 (using the phrase "innocent persons" meaning innocent third parties).

186. *See* BELL, *supra* note 2, at 123; Robert L. Hayman, Jr. & Nancy Levit, *Un-Natural Things: Constructions of Race, Gender, and Disability*, *in* CROSSROADS, DIRECTIONS, AND A NEW CRITICAL RACE THEORY, *supra* note 1, at 180; Thomas Ross, *Innocence and Affirmative Action*, 43 VAND. L. REV. 297, 297-301 (1990), *reprinted in* CRT: CUTTING EDGE, *supra* note 1, at 635-47;

But the Court has also expressed a concern for protecting subordinated racial groups from stereotypical stigmas associated with remedies such as affirmative action, claiming "it may not always be clear that a so-called preference is in fact benign."[187] The effect of the requirement to prove intent and the application of strict scrutiny in the context of equal protection defined as individual as opposed to group protection and increasingly influenced by colorblind ideology are many. They include, the Court: (1) treating racial classifications as if they are the source of racial oppression as opposed to the systemic racial ordering in the human and institutional decision-making and operation of white power in the United States; (2) treating claims of white people to maintain the unearned privileges bestowed by the racial caste system as if they were the same as nonwhite claims to change and make the social system more fair; (3) ignoring societal wide racial ordering and the racial caste system by isolating and remedying only those practices that can be specifically identified and proven as the products of intentional actions; and (4) thus severely limiting anti-discrimination or anti-oppression measures, either forward-looking or past-correcting, thereby preserving the status quo of racial inequality.

For example, in *Regents of the University of California v. Bakke*,[188] a suit bought by a white applicant denied admission to a medical school, the Supreme Court struck down an admissions program that reserved a number of seats for minority students in the entering class.[189] The Court noted, "the guarantee of equal protection cannot mean one thing when applied to one *individual* and something else when applied to a person of another color."[190] According to Justice Powell's majority opinion, then, strict scrutiny—which had hitherto been applied primarily in cases where the government had used racial identification to disadvantage minorities—should apply to any racial classification for any purpose. The Court thus struck down a program meant to undo the racial ordering of white privilege and nonwhite subordination, arguing that equality protected the individual regardless of his "race," or should be blind to "race" (biology), despite the socially-constructed meaning, both material and expressive, of race in America that rendered admissions into the medical school predominantly white in the first place.[191] In addition, Justice Powell expressed concern for "innocent [white] persons," noting that it was impermissible for them to be forced "to bear the burdens of redressing grievances not of their making,"[192] but seemed blind to the fact that that his *decision left blacks bearing the burdens of exclusion, subordination and discrimination not of their own making.*

187. *Adarand*, 515 U.S. at 226 (quoting *Bakke*, 438 U.S. at 298 (opinion of Powell, J.)).
188. 438 U.S. 265 (1978).
189. *Id.* at 289-90.
190. *Id.* (emphasis added).
191. *See supra* note 26 and accompanying text (on the meaning of race).
192. *Bakke*, 438 U.S. at 298.

Further in 1986, the Court struck down provisions of a collective bargaining agreement that in response to past integration litigation provided black teachers greater protection against layoffs than it provided to white teaches with higher levels of seniority.[193] In absence of the agreement, an agreement, again, meant to undo the racial ordering of white privilege and nonwhite subordination, black teachers would have been the first fired because they were the last hired, their hiring presumably occasioned in the first place by affirmative action measures. [194] Brought by white teachers seeking to protect their positions through the seniority plan, the Court, referring to whites, noted that the level of scrutiny did not alter "because the challenged classification operates against a group that historically had not been subject to governmental discrimination."[195] Commenting on societal-wide racial discrimination as a basis for upholding the agreement, the Court notes:

> Societal discrimination, without more, is too amorphous a basis for imposing a racially classified remedy . . . . No one doubts that there has been serious racial discrimination in this country. But as the basis for imposing discriminatory legal remedies that work against innocent people, societal discrimination is insufficient and over-expansive.[196]

Apparently it was acceptable for blacks and other nonwhites to suffer the cost of societal discrimination operating to the benefit of whites, but inappropriate and over-expansive for whites to bear any costs in eliminating this same societal discrimination that primarily and inappropriately privileged them.

And in 1989, the Court struck down a plan by the city of Richmond requiring those who received city contracts to subcontract thirty percent of the contract's value to businesses owned by minorities.[197] The city enacted the plan to remedy past discrimination in the construction industry based on a number of factors including, (1) testimony about racial discrimination in the industry, (2) the fact that, although black residents constituted almost fifty percent of the city, they received less than one percent of public contracting funds; (3) that there were almost no minority contractors in local and state contractors' associations; and (4) that "in 1977, Congress [had] made a determination that the effects of past discrimination had stifled minority participation in the construction industry nationally."[198] The Court explained that these facts did not "provide the city of Richmond with a 'strong basis in evidence for its conclu-

193. *Wygant*, 476 U.S. at 273; *see also* Ross, *supra* note 186, at 638 (discussing this case).
194. Ross, *supra* note 186, at 638.
195. *Wygant*, 476 U.S. at 273.
196. *Id.* at 276 (emphasis omitted).
197. City of Richmond v. J.A. Croson Co., 488 U.S. 469, 477, 498 (1989) (plurality opinion).
198. *Id.* at 499 (plurality opinion).

sion that remedial action was necessary.'"[199] The Court reasoned that the wrong was so "ill-defined" that "relief"—apparently incredulously and impermissibly—"could extend until the percentage of public contracts awarded to [minority-controlled businesses] in Richmond mirrored the percentage of minorities in the population as a whole."[200] And although the Supreme Court in *Grutter v. Bollinger*,[201] ultimately allowed applicants' "race" to be considered in the limited area of admissions in higher educational institutions in a case involving a white student denied law school admission;[202] it did so on a theory of diversity that abstracted and disconnected the meaning of diversity from exclusionary practices and racial/social justice concerns.[203]

However, it is in the complicated area of voting where the colorblind ideology seems to have received its greatest boost. Here, despite the fact that the 15th Amendment was initially enacted to provide the newly emerging slaves a right to vote and there has been and continues to be a long and appalling struggle to ensure meaningful enfranchisement,[204] the Court, according to Bell, has established a new constitu-

199. *Id.* at 500 (plurality opinion) (quoting *Wygant*, 476 U.S. at 277 (plurality opinion)).

200. *Id.* at 498 (plurality opinion). Justice O'Connor also seemed concerned by the fact, in a section of the opinion lacking a majority, that Richmond was 50 percent black and five of the nine council officials that passed the ordinance were black. She suggested that this presented a stronger need for the application of strict scrutiny since it was not a majority of whites acting in a way that burdened themselves but was the act of minorities doing so. *Id.* at 495-96 (opinion of O'Connor, J.). "Many commentators, the Court dissenters included, were astonished by the *Croson* decision." BELL, *supra* note 2, at 666; *see also Croson*, 488 U.S. at 551-52 (Marshall, J., dissenting); Patricia Williams, *Legal Storytelling: The Obliging Shell: An Informal Essay on Formal Equal Opportunity*, 87 MICH. L. REV. 2128 at 2129–30 (1989); Ross, *supra* note 144, at 381.

201. 539 U.S. 306 (2003).

202. The decision came on the heels of lower court decisions and statutes applying colorblind approaches to similar affirmative action cases. The effects of these laws were that fewer black and brown students gained admission into these institutions, leaving them overwhelmingly white. *See, e.g.*, California Proposition 209, *supra* note 77; Hopwood v. Texas, 78 F.3d 932, 934 (5th Cir. 1996). These events revealed not only the social embeddedness of white supremacist racial ordering but demonstrated the way in which colorblindness policies were a proxy for and a mechanism for maintaining white access and privilege.

203. *Who Gets In: A Quest for Diversity after* Grutter – *The 2004 James McCormick Mitchell Lecture*, 52 BUFF. L. REV. 531, 579, 584 (2004) (transcript of discussion by Frank Wu and Charles Daye). Other panelists included Athena Mutua, Shedon Zedeck, Margaret Montoya and David Chambers. Law school admissions continue to be a focus of CRT attention. *See, e.g.*, Dorothy A. Brown, *Taking* Grutter *Seriously: Getting Beyond the Numbers*, 43 HOUS. L. REV. 1, 5 (2006); Kevin R. Johnson & Angela Onwuachi-Willig, *Cry Me a River: The Limits of "A Systemic Analysis of Affirmative Action in American Law Schools,"* 7 AFR.-AM. L. & POL'Y REP. 1, 1-5 (2005). The decision was then justified in part on proclaiming the benefits of diversity as teaching presumably white students that "there is no minority viewpoint," and "visibly" lending legitimacy to the system by signifying that the "path to leadership [is] . . . open to talented and qualified individuals of every race and ethnicity." *Grutter*, 539 U.S. at 320, 332 (internal citations omitted).

204. This struggle of course continues up until this very day. Consider the efforts to disenfranchise minorities in Florida during the 2000 Presidential election and again in Ohio in the 2004 election. For discussions of these events, see, e.g., Calmore, *supra* note 30, at 1267-71; Hugh M. Lee, *An Analysis of State and Federal Remedies for Election Fraud, Learning from Florida's Presidential Election Debacle*, 63 U. PITT. L. REV. 159, 159-160 (2001); see also U.S. COMMISSION ON CIVIL RIGHTS, DRAFT REPORT: VOTING IRREGULARITIES IN FLORIDA DURING THE 2000 PRESIDENTIAL ELECTION (approved by the Commissioners on June 8, 2001), *available at* http://www.usccr.gov/pubs/vote2000/report/main.htm; Monique L. Dixon, *Constructive Disenfran-*

tional injury. First articulated in the case of *Shaw v. Reno*,[205] the injury seems to be that of a state by "a predominant use of race in redistricting" impermissibly sending the message "that racial identity is and should be an American citizen's most salient political characteristic."[206] The point is that redistricting should be race-blind.[207] Unsurprisingly, the case involved blocked efforts to create a voting district that would render black votes meaningful, while ignoring the pervasive historical and current realities of racial bloc voting by whites who often have been reluctant to elect black representatives or elect whites who will effectively represent black or other nonwhite interests and ignoring the racialized politics and gerrymandering of white politicians.[208] The ultimate effect of these interpretations is that law, while denying blacks and other non-

---

chisement: The Problems of Access & Ambiguity Facing the American Voter: Minority Disenfranchisement During the 2000 General Election: A Blast from the Past or a Blueprint for Reform, 11 TEMP. POL. & CIV. RTS. L. REV. 311, 311-313 (2002). On felony disenfranchisement, one of the major forms of minority disenfranchisement used in Florida, see, e.g., Afi S. Johnson-Parris, Note, *Felon Disenfranchisement: The Unconscionable Social Contract Breached*, 89 VA. L. REV. 109, 109-14 (2003); Pamela S. Karlan, *Ballots and Bullets: The Exceptional History of the Right to Vote*, 71 U. CIN. L. REV. 1345, 1345-46 (2003); Pamela S. Karlan, *Convictions and Doubts: Retribution, Representation, and the Debate over Felon Disenfranchisement*, 56 STAN. L. REV. 1147, 1147-50 (2004); Marc Mauer, *Disenfranchisement of Felons: The Modern-Day Voting Rights Challenge*, CIVIL RIGHTS JOURNAL, Winter 2002, at 40; J. Whyatt Mondesire, *Felon Disenfranchisement: The Modern Day Poll Tax*, 10 TEMP. POL. & CIV. RTS. L. REV. 435, 435-36 (2001); Elena Saxonhouse, Note, *Unequal Protection: Comparing Former Felon's Challenges to Disenfranchisement and Employment Discrimination*, 56 STAN L. REV. 1597, 1598-1601 (2004); Christopher Uggen & Jeff Manza, *Disenfranchisement and the Civic Reintegration of Convicted Felons*, *in* CIVIL PENALTIES, SOCIAL CONSEQUENCES 67, 68-72 (Christopher Mele & Teresa A. Miller eds., 2005); Note, *One Person, No Vote: The Law of Felon Disenfranchisement*, 115 HARV. L. REV. 1939, 1939-42 (2002). For a discussion specifically on the 2004 election, see, e.g., Richard L. Hasen, *Beyond the Margin of Litigation: Reforming U.S. Election Administration to Avoid Electoral Meltdown*, 62 WASH. & LEE L. REV. 937, 938-39 (2005) (discussing problems with the 2004 presidential election); Daniel P. Tokaji, *Early Returns on Election Reform: Discretion, Disenfranchisement, and the Help America Vote Act*, 73 GEO. WASH. L. REV. 1206, 1220-39 (2005) (discussing problems with the 2004 presidential election in Ohio).

205. 509 U.S. 630 (1993).

206. BELL, *supra* note 2, at 516.

207. *Id.* (explaining that the new objective of the Court was to enforce a color-blind approach); *see also* Calmore, *supra* note 30, at 1271-74 (arguing that the Supreme Court is undermining the Voting Rights Act through an imposition of colorblind injustice and commenting that "[w]hile the Supreme Court majority opinions are death to black race consciousness, they are amazingly naive or intellectually dishonest when it comes to appreciating that whiteness is also deployed in bloc voting ways"). Calmore goes on to argue that:

> The race opinions of a five-to-four majority on the Supreme Court achieve a false coherence through an incredible process of decontextualization. As a consequence of this radical decontextualization, the Supreme Court majority appears literally blinded by color as it - the same five characters each time - neither acknowledges nor recognizes the degree of racially polarized white bloc voting and therefore will not permit black bloc voting in what really amounts to self defense. The Court acts as if black bloc voting is the first move rather than the second, the initial fire rather than the return fire. Through this misrecognition of whiteness, the Court majority subjects this second move to strict scrutiny as it claims to be unable to distinguish invidious from benign racial discrimination. The second move, blacks taking race predominantly into account, violates the equal protection rights of whites, because they are deemed to suffer individual 'expressive harm.' Individual rights trump efforts to redress vote dilution, a group harm.

Id. at 1273 (internal citations omitted).

208. *See also* BELL, *supra* note 2, at 489-527 (tracing the cases and various obstacles to meaningful enfranchisement for blacks).

whites justice, has protected vested white interests accumulated over time, such as seniority systems, educational advancement, wealth, and/or or white expectations, through the racialized orderings of the society, *first on a theory of white superiority and increasingly on the theory of colorblindness.*

Until the job of racial justice is done, CRT theorists might argue, CRT will expose the whiteness of colorblindness, the white supremacist effects of colorblind laws and rulings, and the white consciousness of American society and power. And CRT theorists will do so, from inside the experiences, consciousness, and perspectives of Black, Native American, Latina/o, and Asian American people, etc., using these essentialized categories that white power created to oppress, strategically as sources of solidarity, empowerment, and analysis.

VI. CRT-RELATED SCHOLARSHIP, THE LATCRIT EXAMPLE: DEEPENING AND BROADENING THE CRT PROJECT?

The development and proliferation of other scholarship and alternative institutions such as LatCrit, Asian American legal scholarship, and critical race feminism, and the writings of sexual minorities of color to explore race, law, and other systems of oppression has raised concerns about the fragmentation of the Critical Race Theory (CRT) project. The development of this CRT-related scholarship, and for instance the separate institutionalization of the LatCrit project, also raised concerns over the multiplication of identity-based groups and thus the elevation of identity politics as opposed to the formation of a broad political movement, a criticism widely made in reference to identity politics.[209] However, though a broad political movement has yet to emerge, this "fragmentation" has actually deepened and broadened the Critical Race Theory project by providing the necessary intellectual expansion and theoretical bridges between identity politics and a politics of solidarity based on difference.[210] So, for instance, these bodies of scholarship broadened the racial lens through which the workings of white racial privilege were revealed, such as through Asian American legal scholarship's exploration of Asian-American experiences, and deepened the commitment to the project of antisubordination by focusing on other systems of social subordination, such as the patriarchal gendered oppression of women explored in the context of critical race feminism, as discussed earlier. However, in addition, they spurred the development of other theories

209. *See, e.g.*, discussion *supra* note 118; Delgado, *Blind Alleys*, *supra* note 17 at 127-28; Bryant, *supra* note 7.

210. A politics of solidarity based on difference stresses common projects and commitments and embraces difference and diversity. But it rejects assimilation of one group into another as a prerequisite to or on basis of that solidarity. It also rejects the universalization of one group's experience. *See* IRIS MARION YOUNG, JUSTICE AND THE POLITICS OF DIFFERENCE 157-158 (1990); NANCY FRASER, JUSTICE INTERRUPTUS: CRITICAL REFLECTIONS ON THE "POSTSOCIALIST" CONDITION 197-205 (1997) (examining and critiquing this idea).

through which to analyze the systems of subordination and developed concepts that emphasized coalition and praxis. Multidimensionality is one such emerging theory which is based on the insights of antiessentialism and intersectionality, insights foundational to the formation of related scholarship communities; while the concept of antisubordination praxis has stressed coalition building in addition to and through theory-informed practice and practice-informed theory. And finally, LatCrit, in particular, in institutionalizing the LatCrit conference also institutionalized the *practice* of continuing to develop critical theory, building coalitions, and engaging practice.

A. Antiessentialism and Intersectionality: Informing Formation, Leading to Multidimensionality

The depth of LatCrit institutionalization and the conscious theorizing around this institutionalization and development is unique among much of the CRT-related scholarship, and I focus on LatCrit for this reason.[211] But several of the themes I discuss here as basic to LatCrit theory have been built upon, shared and/or mutually developed by the work of scholars who identify as CRT scholars or with one or another related body of work.

People such as Robert Chang, Jerome Culp, Angela Harris, Berta Hernández-Truyol, and particularly Frank Valdes and later Elisabeth Iglesias,[212] among others, were the spirit behind the establishment and/or blossoming of LatCrit. They saw LatCrit as a first cousin of CRT and were committed to building on its strengths and avoiding its mistakes.[213] The first mistake they sought to avoid was the perceived elitism and exclusivity of CRT, while nonetheless focusing on issues germane to the Latina/o community and Latina/o identity (a focus that could be perceived as exclusive). [214] The latter was a challenge, in part because Latina/o identity as a practical matter seemed to embrace people representing a host of different nationalities with different cultural perspec-

211. LatCrit also is now an incorporated non-profit organization. While not heavily institutionalized, there are at least two journals committed to the promotion of Asian American legal scholarship focusing primarily on Asian Pacific American communities. These are the Asian Law Journal at University of California Boalt Hall School of Law and the Asian Pacific American Law Journal at University of California Law School. Adrien Wing has done the most work in theorizing and pulling together materials that represent critical race feminist insights by publishing two readers, one on Critical Race Feminism, *see* CRITICAL RACE FEMINISM, *supra* note 37, and another on Global Critical Race Feminism, *see* GLOBAL CRITICAL RACE FEMINISM, *supra* note 44.

212. I believe few people would argue with my characterizing Francisco "Frank" Valdes as the father of LatCrit. However, for many of us, Elizabeth "Lisa" Iglesias is the mother of LatCrit. From what I can figure out Lisa became a central part of LatCrit by around LatCrit's second annual conference. She remained immersed in LatCrit very much guiding its efforts with Frank and bringing to the fold people like myself. She left LatCrit sometime after Latcrit VIII (I was out of the country during that year). But the parting and its aftereffects seemed anything but amenable and for some of us represented a low moment for LatCrit as a community committed to life affirming practice, scholarship and safe space.

213. Valdes, *Ethnicities*, *supra* note 104, at 3-7.

214. *Id.*; Valdes, *Under Construction*, *supra* note 104, at 1090.

tives and historical experiences; and who were raced differently, as black, white, and/or mestizo, among other differences. Valdes envisioned a larger pan-Latina/o identity, but one which both respected and explored the diversities within the group while building solidarity among the differently positioned individuals and groups within the larger pan-ethnic collectivity.[215] From this perspective, to exclude white people, black people or lesbian people or just about anybody else committed to social justice, had the potential of excluding someone who might identify as Latina/o. Further, the founders of LatCrit understood that those who identified as non-Latina/o might have analogous experiences that would also contribute to the building of LatCrit theory.

In this way, LatCrit faced the issue of antiessentialism in Latina/o identity concretely in its experience of politically promoting group formation for the purpose of producing knowledge and building community.[216] The anti-essentialism critique had engendered substantial prior debate in CRT, feminist, and other intellectual circles.[217] It recognizes that there is no single Latina/o essence, no coherent collective identity or single experience that could reflect the common interests of the people constituting the group "without acknowledging the intra-group differences, [218] and rejects the idea of essentialism in Blackness, Asian Americanness, women and the like. For example, CRT scholars joined other women of color who had long challenged the category of "woman" in feminist writings as largely reflecting the interest and priorities of white middle class women in the United States and thus not representative of all women's experience; a quintessential anti-essentialist critique. [219]

Although LatCrit and other CRT related scholarship took an anti-essentialist approach to the categories of blacks, Latina/os, Asian-Americans, and/or women with regard to the dynamics and definition of the groups, their formation was simultaneously premised on the pragmatic idea that these essentialized group identities, having arisen in part

215. Valdes, *Ethnicities*, *supra* note 104, at 9.

216. Asian-American legal scholarship faced similar challenges. It focused on the unique and varied experiences of Asian Americans in U.S. society. Asian Americans, like Latina/os, consist of people from widely varying Asian communities and nationalities. Like Latina/os they experience a racialized identity that often flattens misunderstands their differing ethnic diversities. Their experiences often include an American view of them as foreign, thereby delegitimizing and limiting their human potential while simultaneously holding them up as model minorities, as against other minorities. ROBERT S. CHANG, DISORIENTED: ASIAN AMERICANS, LAW, AND THE NATION-STATE 1-2 (1999); Aoki, *supra* note 1, at 1477; Chang, *supra* note 1, at 1303; Chang, *supra* note 26, at 87-96; Gotanda, *supra* note 2, at 28; Saito, *supra* note 1, at 294-95.

217. *See, e.g.*, Elizabeth M. Iglesias & Francisco Valdes, *Introduction: Expanding Directions, Exploding Parameters: Culture and Nation in LatCrit Coalitional Imagination*, 33 U. MICH. J.L. REFORM 203, 227-232 (2000).

218. Elizabeth M. Iglesias, *Toward Progressive Conceptions of Black Manhood – LatCrit and Critical Race Feminist Reflections Thought Piece,* May 2001, *in* PROGRESSIVE BLACK MASCULINITIES, *supra* note 107, at 55.

219. *See, e.g.*, Harris, *supra* note 42, at 755.

in response to oppression based on essentialism, had been and could be used strategically, politically, and consciously to fight oppression.[220]

The antiessentialist insight was complemented by intersectional theory, which allowed for a more nuanced understanding of intra-group difference. It simultaneously demonstrated the links between different systems of subordination such as racism, sexism, and heterosexism as located in the particular social positions of racial subgroups, such as Latinas or black women. [221] Intersectional theory, first articulated as a theory by Kimberlé Crenshaw, drew upon black feminist thought which had consistently argued that black women were not only oppressed by the white supremacist system of racism but were also oppressed by the patriarchal practices and system of sexism.[222] The theory thus explored the experiences of black women at the intersection of racism and sexism, rejecting a single-axis framework (race or gender) for understanding the arguably doubly burdened conditions of black women.[223] As such, Crenshaw's concept of intersectionality, found a ready home and a site for its further development in LatCrit[224] Asian American legal scholarship,[225] and Critical Race Feminism,[226] as it had in CRT.[227] Each of these bodies of scholarship has made substantial contributions to further

220. OMI & WINANT, *supra* note 26, at 53-60 (suggesting that oppressed people come together strategically for survival because they are oppressed based upon essentialized categories); *see also* Robert S. Chang & Natasha Fuller, *Performing LatCrit: Introduction*, 33 U.C. DAVIS L. REV. 1277, 1291-92 (2000) (citing Sumi Cho & Robert Westley, *Critical Race Coalitions: Key Movements that Performed the Theory*, 33 U.C DAVIS L. REV. 1377, 1414 (2000)) (noting that: "unbounded anti-essentialist theory can be disabling to community organizing, and 'once set in motion, antiessentialism unmodified has no limiting principles to prevent minority groups from being deconstructed until all that remains are disunited and atomized individuals themselves.' Cho and Westley criticize the second wave's fascination with anti-essentialism as 'the ahistorical pursuit of the 'theoretical' that represents an abdication of political engagement and the relinquishment of the full promise of anti-subordinationist intellectual production.' Anti-essentialist theory is understood here to be antithetical to effective political organizing. If this is right, those within the field of CRT may be working at cross purposes; perhaps CRT needs to reform itself and embrace what Professor Cho describes elsewhere as 'essential politics.'" (citations omitted)).

221. 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, *reprinted in part in* POWER, PRIVILEGE AND LAW, *supra* note 1, at 465-72 [hereinafter Crenshaw, *Demarginalizing Intersection*]; *see also* K.L. Broad, *Critical Borderlands & Interdisciplinary, Intersectional Coalitions*, 78 DENV. U. L. REV. 1141, 1143 (2001); Crenshaw, *Mapping the Margins*, *supra* note 18, at 1299; Mari J. Matsuda, *When the First Quail Calls: Multiple Consciousness as Jurisprudential Method*, 11 WOMEN'S RTS. L. REP. 7, 7–9 (1989);.

222. *See* Crenshaw, *Mapping the Margins*, *supra* note 18, at 1252.

223. *Id.*

224. *See, e.g.*, Berta Esperanza Hernández-Truyol, *Women's Right as Human Rights–Rules and Realities and the Role of Culture: A Formula for Reform*, 21 BROOKLYN J. INT'L L. 605, 610-11 (1996); DELGADO & STEFANCIC, *supra* note 1; *see also* discussion *supra* note 212 (role of Valdes and Iglesias).

225. *See. e.g.*, Peter Kwan, *Jeffrey Dahmer and the Cosynthesis of Categories*, 48 HASTINGS L.J. 1257, 1288 (1997).

226. *See, e.g.*, CRITICAL RACE FEMINISM, *supra* note 37, at 177-79.

227. *See, e.g.*, Hutchinson, *supra* note 19, at 9-12; Matsuda, *supra* note 221, at 9. Much of the discussion on multidimensionality is taken from and appears in my chapter on progressive black masculinities. *See* PROGRESSIVE BLACK MASCULINITIES, *supra* note 105, at 21-23.

developing the idea of intersectionality, including using it as a basis for the emerging theory of multidimensionality.

B. Multidimensionality: An Emerging Theory?

The theory of multidimensionality captures three separate ideas. First, it recognizes that an individual has many dimensions, some of which are embodied human traits such as skin color, sex, ear-lobe length, eye color; and others which are expressive, such as being Methodist or Catholic, a cat owner or dog owner, etc. Second, multidimensionality identifies some of these dimensions as materially relevant,[228] meaning that a particular society has taken some dimensions such as color, sex, or a particular religious belief (but not ear-lobe length or owning a cat or a dog) and constructed meanings about the groups that possess them. It then allocates or denies both material and status-related resources through systems of racism, sexism, and anti-Semitism, for example. These systems operate on the individuals who belong to groups that inhabit or express a particular trait producing a host of experiences, rewards or demerits. Said differently, multidimensionality captures the way society disadvantages people or benefits them primarily on the basis of their possessing a particular trait. Thus, the second idea of multidimensionality is a focus on systems of subordination and privilege, such racism, sexism, and heterosexism.[229]

Third, multidimensionality recognizes that these systems intersect, inter-relate, and are mutually reinforcing so that for example, racism is patriarchal and patriarchy is racist. In addition, however, the intersection of two or more systems of disadvantage or privilege often produces unique categories and experiences. So for example, intersectional theory might suggest that black men are privileged by gender and oppressed by race.[230] But this might not sufficiently explain racial profiling as a phenomenon that happens most frequently to black men.[231] Multidimensionality, however, might better capture the idea that black men are sometimes oppressed because they are "blackmen" one word, one position, one socially, multidimensionally constructed oppressed group of people—they are both black and men.[232] The positionality of blackmen—one word—could be further analyzed by looking at the class or another materially relevant status of those who are, in this example, racially profiled. In this sense, it replaces an approach that is additive, that is, one that says, for instance, that poor black men are poor + black +

228. Valdes, *Under Construction*, *supra* note 107, at 1094.

229. Hutchinson, *supra* note 144, at 1199.

230. PROGRESSIVE BLACK MASCULINITIES, *supra* note 107. However, Stephanie Phillips has pointed out to me that these interpretations may not be correct interpretations of the intersectional theory. She argues instead that there is no meaningful difference between the theory of intersectionality and multidimensionality.

231. *Id.* at 21-22.

232. *Id.* at 23.

men, to one that allows the analyst to explore the way that axes of classism, racism, and gender oppression mutually construct unique positions for individuals and groups, while simultaneously demonstrating that these same forces are implicated across identity categories.

And finally, multidimensionality is an approach. In recognizing, for example, that racial oppression is also gendered, sexualized, and classed, a multidimensional approach conceptually links the struggles of racial justice, gender justice, and class justice; and has the potential to link the differing groups fighting for these different types of justice. As such a multidimensionality approach is an approach to building solidarity and potentially coalitions.[233]

C. Antisubordination Praxis

Ideas about solidarity, multidimensionality and coalition-building are meant both to inform and to serve the principle and project of antisubordination. In LatCrit, the commitment to antisubordination stated explicitly at the beginning of the movement meant that LatCrit embraced participation of people from other groups that had been historically oppressed. Thus, LatCrit embraced those who were fighting also against heterosexism, and other sexually-based oppressions. This effort was made easier by Valdes' participation, as he was known affectionately in some quarters as the Queer Theory Philosopher after the publication of a path-breaking piece on sexual identity.[234] At the same time, LatCrit's anti-subordination commitment[235] and its kinship with CRT led the organizers to consciously pursue both a theory and the practice of coalition building as well as a theory of praxis—theory—informed practice and practice—informed theory, as explored by Eric Yamamoto, among others.[236] So for example, LatCrit developed a number of programmatic devices that have been employed in LatCrit conference planning to help balance the demands of facilitating their commitment to both anti-essentialism and coalition building. One such device was the practice at each conference of rotating centers, in which a non-Latina/o group's experiences and insights were "centered" in a conference workshop for exploration. In this way, LatCrit hoped to expand its theory, deepen its knowledge about various groups' histories, and having brought these groups into the conference, potentially to build coalitions with them.[237]

---

233. Valdes, *Ethnicities*, *supra* note 104, at 1094.

234. Francisco Valdes, *Queers, Sissies, Dykes and Tomboys: Deconstructing the Conflation of "Sex," "Gender," and "Sexual Orientation" in Euro-American Law and Society*, 83 CAL. L. REV. 3 (1995).

235. Iglesias & Valdes, *supra* note 37, at 1265-67.

236. Eric K. Yamamoto, *Critical Race Praxis: Race Theory and Political Lawyering Practice in Post-Civil Rights America*, 95 MICH. L. REV. 821, 829 (1997).

237. Iglesias & Valdes, *supra* note 37, at 1267. Another device not related to coalition building was the practice of "streaming a topic," such as class. Streaming required that LatCrit would return to a topic over a period of years in order to build upon previous and new insights.

The idea of coalition building is also at the center of LatCrit, CRT and other related scholarship's commitment to critical race praxis, as developed by Eric Yamamoto.[238] This concept is often termed "antisubordination praxis," to encompass the many justice projects that inform the antisubordination commitment. Yamamoto had been involved in justice struggles that brought different but conflicting minority groups together.[239] Thus, the idea of coalition and alliances are a part of the idea. But, a critical race praxis or antisubordination praxis also refers to the idea that critical theorizing should and needs to be informed by practice, by active engagement with developments on the ground while practice should and needs to be informed by theorizing and theories about what is happening, all for the benefit of oppressed communities. Yamamoto's notion of critical race praxis is summarized as:

> [Combining] critical, pragmatic, socio-legal analysis with political lawyering and community organizing to practice justice by and for racialized communities. Its central idea is that racial justice requires antisubordination practice. In addition to ideas and ideals, justice is something experienced through practice . . . It requires in appropriate instances, using critiquing, and moving beyond notions of legal justice pragmatically to heal disabling wounds and forge intergroup alliances. It also requires, for race theorists, enhanced attention to theory translation and deeper engagement with frontline practice; and for political lawyers and community activists, increased attention to a critical rethinking of what race is, how civil rights are conceived, and why law sometimes operates as a discursive power strategy.[240]

While this definition raises a number of provocative issues, two seem particularly important. The first is the recognition that justice is experienced through practice and that racial justice in particular requires an antisubordination practice.[241] These ideas capture and reinforce CRT

---

238. Yamamoto, *supra* note 236, at 829. For a fuller examination of the idea of critical race praxis, particularly its focus on interracial healing or community building by CRT-related scholars, see, e.g., Keith Aoki & Margaret Chon, *Nanook of the Nomos: A Symposium on Critical Race Praxis*, 5 MICH. J. RACE & L. 35, 36 (1999); see also Gitanjali S. Gutierrez, Note, *Taking Account of Another Race: Reframing Asian-American Challenges to Race Conscious Admission in Public Schools*, 86 CORNELL L. REV. 1283, 1312-1318 (2001) (explaining critical race praxis as the pursuit of interracial justice through antisubordination practice, political lawyering, and education. Antisubordination practice "seeks to disrupt the use of law as an instrument for perpetuating hierarchical power relations. Political lawyering calls for "lawyers to play a more active role in working with their minority clients to shape and guide antidiscrimination litigation" and "could involve bringing lawsuits as part of a comprehensive impact strategy." The "educative function of litigation seeks to increase "awareness of the court and parties alike about the interdependence of the legal rights of opposing minority litigants as well as the distinctions between their historical and current experiences of racial subordination" (internal quotation marks and citations omitted)); Reginald Leamon Robinson, *Human Agency, Negated Subjectivity, and White Structural Oppression: An Analysis of Critical Race Practice/Praxis*, 53 AM. U.L. REV. 1361, 1365-66 (2004) (defining critical race praxis in similar ways as antisubordination praxis but critiquing the race consciousness implicit in the idea).

239. Yamamoto, *supra* note 236, at 829.

240. Aoki & Chon, *supra* note 238, at 36 (summarizing Yamamoto, *supra* note 236).

241. *See* Yamamoto, *supra* note 236, at 880.

notions that context is important and that what is considered fair will be contextual on the one hand, and on the other that different groups will have to work together to accomplish justice, the justice for healing their intergroup wounds or otherwise. Second, Yamamoto insists that CRT theorists translate theory into practical use for activist and legal practitioners and that these practitioners think critically and theoretically about what their work means and the ways it may be limited.[242] This is best accomplished by these groups working together, theorist and activist, to mutually inform one another's work; and more importantly, to do so on behalf of and in conjunction with communities of color. Critical race and related scholarship theorists believe that engaged community work and practice as well as attention to this work should dictate and inform the contours of CRT theorizing, grounding it and potentially rendering it more useful for community advancement. This idea is reminiscent of their experiences of having grounded their critique of CLS scholarship in the African American civil rights experience.[243] However, antisubordination praxis requires something more than knowing an experience, it requires active engagement.

These commitments and ideas, for instance, were institutionalized in annual LatCrit conferences. The conferences were organized around four functions, namely, "(1) the production of knowledge; (2) the advancement of social transformation; (3) the expansion and connection of anti-subordination struggles; and (4) the cultivation of community and coalition,"[244] both within and beyond the confines of legal academia in the United States. Further LatCrit established "early guideposts," that in some ways mirrored but greatly expanded CRT tenets, evidencing a commitment to antisubordination praxis and multidimensional/antiessentialist theorizing. These were to:

[1] Recognize and Accept the Political Nature of Legal "Scholarship" Despite Contrary Pressures.

[2] Conceive Ourselves as Activist Scholars Committed to Praxis to Maximize Social Relevance.

[3] Build Intra-Latina/o Communities and Inter-Group Coalitions to Promote Justice Struggles.

---

242. *Id.* at 828-30.

243. *See, e.g.*, Matsuda, *supra* note, 141 at 344-45. But antisubordination requires more than just knowledge of an experience, it requires engagement. In a time of backlash; however, where there is increasing ideological pressure to be blind to racial and racialized events, such as Hurricane Katrina. *See infra* notes 300-16 and accompanying text (noting that being engaged may be simply naming and re-naming, hearing, and re-telling the hearing of our experiences from our collective perspective). But even this telling and retelling must be done in community or in some sort of minority public sphere.

244. Iglesias & Valdes, *supra* note 37, at 1262; *see* Valdes, *Ethnicities*, *supra* note 104, at 1095; *see* Valdes, *supra* note 19, at notes 161, 175.

[4] Find Commonalities While Respecting Differences to Chart Social Transformation.

[5] Learn from Outsider Jurisprudence to Orient and Develop LatCrit Theory and Praxis.

[6] Ensure a Continual Engagement of Self-Critique to Stay Principled and Grounded.

[7] Balance Specificity and Generality in LatCritical Analysis to Ensure Multidimensionality.[245]

In addition, LatCrit institutionalized a host of projects in order to facilitate the production of critical scholarship and antisubordination practice.[246]

In these and other ways, LatCrit, CRT related scholarship and others, such as queer legal theory, critical work on whiteness and on Native Americans, have contributed valuable insights to the CRT project. The insights into various social groups supply the knowledge base of CRT and inform efforts to build coalitions among various groups. The information and coalitions are meant to service the goal of antisubordination praxis, a goal to which each of the groups is committed. In this sense, what could be understood as a process of fragmentation can be seen as a period of intense contextualized study that deepened the CRT project of exposing the connections between race and law from a variety of perspectives. These efforts also have broadened the CRT project by bringing more people to the table who understand and see the connections between their and others' subordination and who might join coalitions and the CRT project of fighting against racial oppression and all other forms of subordination. This move from "color" to consciousness informed by the understanding and appreciation of difference arguably allows the multidimensionality of oppression to be attacked from a number of vantage points, ventures, and enterprises.

VII. CLASSCRITS?

Nevertheless, several scholars including critical race scholars themselves have criticized CRT scholarship for its focus on the words, discourse, and discursive patterns that support race, gender, and sexual consciousness as opposed to the material determinants of these social sys-

245. Iglesias & Valdes, *supra* note 37, at 1263; Valdes, *Ethnicities*, *supra* note 104, at 1095.

246. These include publishing essays from every conference and workshop resulting in published symposia of LatCrit scholarship, securing coordinators and funding for future conferences several years in advance, incorporating LatCrit into a not-for-profit corporation, establishing the website www.latcrit.org, securing organizational NGO status for the purposes of participating in international foray, founding CLAVE, a peer-reviewed periodical, and establishing numerous programs including a Student Scholar program. *See* Iglesias & Valdes, *supra* note 37, at 1325.

tems of subordination.[247] Specifically these scholars, in a variety of ways, have called for analyses of the class system in U.S. society and the way in which race, gender, and other forms of oppression mutually construct and are constructed by it. [248] Although these calls may critique CRT for ignoring "material factors," which Delgado defines as "issues that turn on tangible events in the social or physical world,"[249] they more often refer to and include 1) calls for exploring the economics of race, a call that narrows the concept of materiality to ideas about the economy as a significant determining factor in racial outcomes, or 2) calls for specifically exploring the class system as the product of economic relations and economic ordering—the product of the production and distribution of goods and services.[250]

While much of CRT scholarship seems focused on discourse, race as a function of ideas, and race as culture, individual CRT and LatCrit scholars have consistently focused on the class/materialist elements of race, such as John Calmore's focus on housing,[251] Enrique Carrasco's focus on development,[252] and Kevin Johnson's focus on immigration;[253]

---

247. Harris, *supra* note 42, at 777-78 (suggesting that race-crits "return to the vexed question of the relationship between race and class" and noting that "[t]he current interest in theories of culture has all but crowded out materialist work on race and racism"); *see, e.g.*, Delgado, *Blind Alleys*, *supra* note 17, at 122.

248. *See, e.g.*, Delgado, *Blind Alleys*, *supra* note 17, at 125-28 (calling for CRT to return to the materialist traditions of the first generation of CRT scholars such as Derrick Bell). Delgado also calls specifically for an examination of the relationship between race and class. *Id.* at 151; *see also* Carbado & Gulati, *supra* note 150, at 1757 (suggesting that CRT use the insights being developed within the field of law and economics to analyze workplace discrimination); Harris, *supra* note 42, at 777-78. *See generally* Jennifer Hochschild, *Symposium: Going Back to Class? The Reemergence of Class in Critical Race Theory*, 11 MICH. J. RACE & L. 99 (2005).

249. Delgado, *Two Ways to Think about Race*, *supra* note 17, at 2280 n.20 (defining material factors as those "that turn on tangible events in the social or physical world").

250. I am defining class very broadly in order to leave open for discussion and later consideration new theories that better define the idea. At the same time, I am adopting the multiple meanings that I believe Delgado seeks to capture in talking about the material determinants of racism. For instance, though Delgado defines material factors as those "that turn on tangible events in the social or physical world[,]" he explains that he, at times, uses "the terms 'material' or 'economic determinism' as synonymous with 'racial realism.'" at n.38. He encompasses the ideas of economic relations by suggesting that it is racial realist, referring specifically to Derrick Bell, who understands that "racism is a means by which our system allocates privilege, status, and wealth." Delgado, *Blind Alleys*, *supra* note 17, at 123. Further, he captures what I consider in part class dynamics when he notes that racism serves "the 'political and economic ends' of 'powerful whites.'" *Id.* at n.16 (citing Bell, *supra* note 65). These ideas are partially captured in his comments: "[W]hile text, attitude, and intention may play important roles in our system of racial hierarchy, material factors such as profits and the labor market are even more decisive in determining who falls where in that system." Delgado, *Blind Alley*, *supra* note 17, at 123.

251. *See generally* Calmore, *supra* note 41; John O. Calmore, *A Call to Context: The Professional Challenges of Cause Lawyering at the Intersection of Race, Space, and Poverty*, 67 FORDHAM L. REV. 1927, 1948-49 (1999); John O. Calmore, *Race/ism Lost and Found: The Fair Housing Act at Thirty*, 52 U. MIAMI L. REV. 1067 (1998).

252. Enrique R. Carrasco, *Critical Race Theory and Post-Colonial Development: Radically Monitoring the World Bank and the IMF*, *in* CROSSROADS, DIRECTIONS, AND A NEW CRITICAL RACE THEORY, *supra* note 1, at 366-75; Enrique R. Carrasco, *Opposition, Justice, Structuralism, and Particularity: Intersections between LatCrit Theory and Development Studies*, 28 U. MIAMI INTER-AM. L. REV. 313, 327-35 (1997); Enrique R. Carrasco, *Law, Hierarchy, and Vulnerable Groups in Latin America: Towards a Communal Model of Development in a Neoliberal World*, 30 STAN. J.

while others have written one or more class-related articles. [254] Further, LatCrit has periodically, but repeatedly, focused on class, economic inequality, and the economics of race in their conferences,[255] as have CRT scholars in symposia such as the Washington and Lee Symposium "CRT and the Next Frontier."[256] In addition, promising new work in the area of class and race has emerged in law and in other fields.[257] However, a systematic analysis of class, particularly as a product of economic ordering, as well as its relationship to race has not yet emerged, *even though critical race scholars have argued for years that the class system in the U.S. mutually constructs race, gender, and other forms of oppression.*

Below I briefly proffer a number of reasons that may explain why CRT and related scholarship has not yet developed a more systematic analysis of class. I then posit that the critical study of law and class or classcrits, while open to further definition, is necessary to examine the ways in which class mutually constructs race, further grounds analyses of racism, and guides antisubordination praxis.

A. Tendencies, Tensions and Fears as Obstacles to Critical Class Analysis

There are three central tendencies within CRT and related scholarship that may explain why this scholarship has failed to adequately engage in critical class analyses. Each tendency involves a tension between material/class and discourse-focused examinations. The discursive sides

---

INT'L L. 221, 221-45 (1994); Enrique R. Carrasco & Randolph Thomas, *Encouraging Relational Investment and Controlling Portfolio Investment in the Aftermath of the Mexican Financial Crisis*, 34 COLUM. J. TRANSNAT'L L. 539, 545-55 (1996).

253. Kevin R. Johnson, *An Essay on Immigration Politics, Popular Democracy, and California's Proposition 187: The Political Relevance and Legal Irrelevance of Race*, 70 WASH. L. REV. 629, 633-642 (1995). *See generally* Kevin R. Johnson, *Fear of an "Alien Nation": Race, Immigration, and Immigrants*, 7 STAN. L. & POL'Y REV. 111(1996); Kevin R. Johnson, *"Melting Pot" or "Ring of Fire"?: Assimilation and the Mexican-American Experience*, 85 CAL. L. REV. 1259 (1997), *reprinted in* 10 LA RAZA L.J. 173 (1998); Kevin R. Johnson, *The Case Against Race Profiling in Immigration Enforcement*, 78 WASH. U. L. REV. 675 (2000).

254. *See* Kevin R. Johnson, *Roll Over Beethoven: "A Critical Examination of Recent Writing about Race,"* 82 TEX. L. REV. 717, 722-26 (2004) (responding to Delgado's critiques in Delgado, *Blind Alleys, supra* note 17 and critiquing him for overlooking much scholarship that addresses the concerns he raises and exploring the plethora of articles that have a class/materialist focus).

255. Two of the eleven LatCrit conferences held thus far specifically focused on class and economic issues. *See LATCRIT: Latina & Latino Critical Theory*, http://personal.law.miami.edu/~fvaldes/latcrit/latcrit/index.html (last visited Oct. 22, 2006). These were LatCrit V titled "Class in LatCrit: Theory and Praxis of Economic Inequality," and Latcrit X titled "Critical Approaches to Economic Injustice." *Id. See also* Johnson, *supra* note 254 (listing numerous examples of the essays discussing the relationship between class and race).

256. Brown, *supra* note 146, at 1499.

257. *See, e.g.*, MELVIN L. OLIVER & THOMAS M. SHAPIRO, BLACK WEALTH/WHITE WEALTH: A NEW PERSPECTIVE ON RACIAL INEQUALITY 35-38 (2d ed. 2006); MASSEY & DENTON, *supra* note 161, at 4. Legal scholars too have begun to investigate these issues from a variety of perspectives. *See generally* IAN AYRES, PERVASIVE PREJUDICE? UNCONVENTIONAL EVIDENCE OF RACE AND GENDER DISCRIMINATION (2001); EMMA COLEMAN JORDAN & ANGELA P. HARRIS, ECONOMIC JUSTICE: RACE, GENDER, IDENTITY, AND ECONOMICS (2005); A WOMAN'S PLACE, *supra* note 43; HUMAN RIGHTS AND THE GLOBAL MARKETPLACE: ECONOMIC, SOCIAL AND CULTURAL DIMENSIONS (Jeanne M. Woods & Hope Lewis eds., 2005).

of these tensions, however, are complementary in ways that reinforce the discursive turn in CRT scholarship. The first tension lies in the inherent nature of the legal profession, which while structuring material conditions does so through language and discourse. The second tension lies in CRT's embrace of post-modernist theories, which focuses on discourse as a significant site of power, in part because humans primarily understand the material world through ideas and language.[258] These theories, however, recognize that humans are material beings in a material world.[259] The third tension is found in CRT's embrace of the idea that race is socially constructed, which given its development in CRT lends itself toward a focus on consciousness,[260] even though consciousness is shaped by and shapes materiality. These tensions are compounded by additional political and academic issues that may further reinforce the tendencies toward discourse analyses as opposed to the materiality of economic relations, economic ordering, and class.

The first tension is implicated in the nature of law itself. Legal work primarily consists of manipulating language and discourse in the process of structuring arguments. Though these arguments are sometimes meant to effectuate particular material arrangements and may both unleash and justify the use of coercive power, these events are accomplished through legal discourse. This occupational feature may mean that legal academics in particular, who may be removed from the rough and tumble of legal practice, are more comfortable mining text, language, and discourses for their meaning, rather than tracking their material effects or responding with concrete strategies to a particular material event. Nevertheless, law through discourse structures material conditions in part by allocating and regulating resources. It decides who the winners and losers are, on what terms trade will be conducted and what kinds of property will be protected.

While many of these issues are explored as individual concepts through the traditional liberal legal discourse of contract, property, trade, and corporate law, or even welfare and poverty law; and are increasingly examined through the overlay onto law of the standard neo-classical eco-

258. *See* Shane Phelan, *(Be)Coming Out: Lesbian Identity & Politics*, 18 SIGNS: J. WOMEN IN CULTURE & SOC'Y 765, 768-69 (1993) (describing the poststructuralist and postmodernist project as one challenging and seeking to unravel the historicity of truth claims rather than engaging in argumentation about whether something is true or not). Phelan describes poststructualism as being more modernist in tone, "retaining some measure of confidence in the Enlightenment ideas of reason and freedom while indicting modern Western societies for betraying these categories even as they ostensibly serve them." *Id.* at 768. In this sense it uses deconstruction or engages genealogy to note gaps between what the West says it is or promises and what it does. On the other hand, postmodernist such as Derrida seek "continually to disrupt modern categories even as they rely on them." *Id.* at 768. Phelan notes one can be both a poststructuralist and a post modernist but "no affinity necessarily exists between the two." *Id.*

259. John Lye, *Some Post-Structuralist Assumptions* (1996-97), http://www.brocku.ca/english/courses/4F70/poststruct.html.

260. Delgado, *Two Ways to Think about Race, supra* note 17, at 2285; Delgado, *Blind Alleys, supra* note 17, at 123.

nomic discourse as in the Law and Economics tradition, it might be helpful to begin to examine the web of historically-decided economic policies that constitute the background and hegemonic economic code of law. Litowitz describes this code, worldview or web of social policies as including "private ownership of property, employment at will, inheritance, freedom of contract, limited liability for business organizations, patriarchy, and a regime of negative rights that ensures that individuals must secure their own health care, day care, and other benefits."[261] Here, one might add white supremacy as an economic imperative of the code.[262] And, one might question who and which social groups may have pushed and/or been best served by these policies and their combination. So for instance, while notions of private property may have a history that predates the founding of the United States, arguably the fact that the founding fathers belonged to the property owning class may be a significant factor in its being enshrined in the U.S. Constitution to protect such ownership. In any event, this web of policies, I believe, help to mark the boundaries of the reproduction of class within the United States that operates both independently and mutually with other subordinating structures to limit the material well being of many people including racial and other minorities. This code, perhaps to the chagrin of its author, provides us an entry into class.[263]

The focus on language and discourse as an inherent feature of modern legal work is reinforced by the post structuralism/modernism of CRT that suggests that power is located in discourses. That is, power is located in rule-governed discourses that shape the features of a society and individual identity and are constructed over time (historically deter-

261. Litowitz, *supra* note 68, at 540.

262. While this may seem like a political imperative rather than an economic one, it is not clear if these two things can be separated in the American social order. Slavery was a racialized economic system; it was both economic and political. *See* Matsuda, *supra* note 141, at 334; *see also*, Anthony Farley, *Accumulation*, 11 MICH. J. RACE & L. 51, 51 (2005) (suggesting that race and class are created in the same moment of primal accumulation in which the differences between the haves and the have-nots is marked on the body, marked racially). Slavery and Jim Crow later did not simply block access to equal services, but also to equal access to resources, jobs, and the like. *See* WOODWARD, *supra* note 32, at 7-16. Similarly Colorblindness may do the same by leaving intact and allowing white communities to build on the historical advantages locked in through slavery, segregation, and discrimination. *See* Daria Roithmayr, *Locked in Inequality: The Persistence of Discrimination*, 9 MICH. J. RACE & L. 31, 33-34 (2003); *reprinted in* 12 VA. J. SOC. POL'Y & L. 197 (2004) (discussing the manner in which advantages get locked-in over time using the economic model that in part explains the advantages monopolies maintain long after the anticompetitive behavior has ceased). In addition, an easily managed or "efficient" workplace may require homogeneity in personnel, or in America, whiteness. *See* Carbado & Gulati, *supra* note 150, at 1762. Pleasing the affluent customer may mean hiring salespeople with whom this particular type of customer may be most comfortable. *See generally id.* In the United States, white people are proportionately more affluent than say black people.

263. Litowitz certainly rejects the notion of class in the Marxian sense as the fount of all oppression. *See* Litowitz, *supra* note 68, at 534. He also rejects the idea that law is an instrument of class. *Id.* But he more generally seems to reject notions of class because the modern economy has begun to "blur class divisions." *Id.*; *see also infra* note 293-95 and accompanying text. *See generally* THOMAS L. FRIEDMAN, THE WORLD IS FLAT: A BRIEF HISTORY OF THE TWENTY-FIRST CENTURY (2005) (making a similar point).

mined) through social action.[264] Though these discourses appear as instruments of some sort of specific knowledge, to in fact reveal knowledge,[265] there are multiple and overlapping discourses and all manner of people have some power to contest these discourses—these sites of power. This is so even though some people have more power to influence the content and development of discourse than others.[266] For lawyers and non-lawyers alike, these insights might easily be seen as a description of law. Law appears as a specialized body of knowledge, that often establishes, or rather reveals the boundaries of appropriate action, and perhaps—though less apparent—thought. Further, just about anyone can bring a suit or contest the discourse, and it is clearly a site of power. It is a site of power not only because it is backed by coercion but also because the language of law as socio-historically developed, itself is socially meaningful and powerful—"I know my rights; this is my property." At this level the post-structuralist description of discourse as power not only seems to describe law but also reinforces the centeredness of discourse in law.

Delgado captures this idea when he suggests that CRT scholars favor analyses that emphasize "texts, narratives, ideas, and meanings" because they understand these as giving rise to racial discrimination by conveying messages "that people of other racial groups are unworthy," and thus locating racial power in discourse.[267] He criticizes these types of analyses as leading writers to focus on and "analyze hate speech, media images, census categories, and such issues as intersectionality and essentialism."[268] But he argues that "while text, attitude, and intention may play important roles in our system of racial hierarchy, material fac-

264. Here I am drawing particularly on Michel Foucault's work and understanding him as the central poststructuralist theorist. *See* MICHEL FOUCAULT, MADNESS AND CIVILIZATION: A HISTORY OF STORY OF INSANITY IN THE AGE OF REASON (Richard Howard trans., 2001); Litowitz, *supra* note 68, at 534 (describing this work in part); *see also* Roger Jones, *Post Structuralism, by Roger Jones*, http://www.philosopher.org.uk/poststr.htm (last visited Oct. 20, 2006) (describing Foucault as the central poststructuralist theorist).

265. Litowitz, *supra* note 68, at 534 (drawing on Michel Foucault's and explaining the way in which discourses appear as instruments of knowledge and noting, "psychoanalysis favors the male experience (the Oedipal drama) and thereby marginalizes women, just as the discourse of medicine tends to medicalize the behavior of those who stand outside established social categories (e.g., homosexuals were considered 'sick')"). *See generally* FOUCAULT, *supra* note 264.

266. FRASER, *supra* note 210, at 154. The idea of multiple sites of power seems to be in some tension with the idea of a single hegemonic order. Litowitz, in arguing for a reconsideration of Gramsci's theory of hegemony suggests that the critical legal studies in using Gramsci, never resolved the tension between Gramsci's understanding of hegemony as a tool of the dominant economic class and other ideas of hegemony as a product of structure. *See* Litowitz, *supra* note 68, at 549. He rejects the idea that law reflects a dominant class and appears to reject the idea of class in general given this moment of advanced capitalism but he argues there is a dominant or hegemonic code that relates to economic issues socially constructed over time. *Id.* He suggests that someone fighting for welfare rights is not fighting against a class but rather a code, a worldview. *Id.*

267. Delgado, *Blind Alleys*, *supra* note 17, at 123.

268. *Id.*

tors such as profits and the labor market are even more decisive in determining who falls where in that system."[269]

The strong focus of post structuralism/modernism on discourse and CRT's embrace of it, however, may squeeze out other poststructuralist insights about materiality.[270] So for instance, post structuralism understands human beings as "material beings, "embodied and present in the physical world, entrenched in the material practices and structures of their society -- working, playing, procreating, living as parts of the material systems of society."[271] Thus while humans may dream of crossing the Atlantic, and may do so by engaging materiality using a boat or plane, it is unlikely they can swim across it, even if they think they can. In this sense, the idea of discursive power does not negate the reality of material constraints, and equally important, it does not negate the reality of coercive power—violence—as a creator, consequence and tool of power. Coercive power is simply another axis of power in addition to discursive power.[272] And the idea of discursive power, itself, though said to reside in many places and posited as contested and contestable rather than as a "one-way imposition by a dominant" group,[273] does not negate the idea of dominant groups or classes (which may have greater power to say what is authoritative). Rather, these dominant groups may be powerful and their views may be hegemonic, but they are neither monolithic nor are their views the only game in town.[274] Further, the very notion of discursive power also points to lopsided or unequal relationships that are all too often the products and producers of unequal and in-egalitarian material relations.[275] In this sense, post-structuralism may remind us that "it's more than the economy, stupid." But, it does not negate economic conditions even if human understanding of these conditions is mediated through language. Thus while post-structuralism's analyses of discourse may reinforce the centeredness of the discourse of law, it, in and of itself, need not preclude a focus on economic structures

269. *Id.*

270. *See generally* FRASER, *supra* note 210 (indicating a need for both).

271. *See* Lye, *supra* note 259. This is so even though they suggest that human beings can only appreciate the material world though ideas and language.

272. Litowitz, *supra* note 68, at 518-19.

273. *Id.* at 534; *see also* Jones, *supra* note 264.

274. Litowitz, *supra* note 68, at 550-51 (noting that postmodernists recognize that there are dominant groups and arguing that just because there are multiple discourses at the local level does not negate that there is a hegemonic discourse at the meta level of law). *See generally* Mutua, *Theorizing Progressive Black Masculinities, in* PROGRESSIVE BLACK MASCULINITIES *supra* note 107 (discussing ideal American masculinity as hegemonic, but not the only idea of masculinity that exists in society); Patricia Hill Collins, *A Telling Difference: Dominance, Strength, and Black Masculinities, in* PROGRESSIVE BLACK MASCULINITIES, *supra* note 107, at 73 (noting that while there is a dominant hegemonic masculinity, all masculinities across racial, ethnic, and other groups contain hegemonic ideas such as, white men encounter a hegemonic masculinity that dictates what white men should be and do and black men encounter ideas that tell them what they, as black men, should be and do).

275. *See* FRASER, *supra* note 210, 11- 33 (noting that it is difficult to have dignity and respect for all identities in the absence of an egalitarian society).

and relations as decisive determinants of class status and significant features in the construction of race.

If the practice of law is centered on discourse and CRT's post structuralism emphases on examining discourse reinforces the tendency toward a focus on discursive maneuvers, then the development of CRT's basic premise that race is socially constructed, has sometimes been understood in ways to further bolster the discursive turn. Delgado argues that it is the idea of the social construction of race that is at the crux of the discursive turn. He suggests it leads to a focus on "race" instead of race.[276] By this he means that CRT has tended to focus on defining race, thinking about how it works, theorizing its processes as located in "[i]deal factors -- thoughts, discourse, stereotypes, feelings, and mental categories[,]"[277] rather than focusing on how race works in the real world, and the way it functions as an ordering principle in a world of power, resources, and privilege.[278] The idea of race as a social construction suggests that race is not some immutable biological characteristic but rather a socio-historical process through which society assigns meanings to different types of human bodies.[279] These meanings are unstable and changing but deeply structured over time not only by consciousness but also by material constraints, economic imperatives, institutionalized practices, and other social ordering features of society.[280] That is, it is shaped by materiality as much as by consciousness. And yet, somehow the idea that race is socially constructed has developed in a way to suggest that race is simply a product of consciousness unrelated to materiality—a consciousness that seems to hover above and beyond social life and practice.[281]

So, for instance, some CRT scholars see unconscious racism as the biggest obstacle to racial justice.[282] This is true to the extent that unconscious racism allows people to continue to act in racist ways while simultaneously believing that racism no longer exists. This may engender in them apathy or hostility to racial justice activism. But this view ignores the ways in which race is economically, socially, and institutionally grounded and reproduced even where individuals are present and consciously committed to eradicating racism. Others seem to suggest that as race is simply a product of consciousness and not real, that race con-

276. Delgado, *Blind Alleys*, *supra* note 17, at 1 36.
277. Delgado, *Two Ways to Think about Race*, *supra* note 17, at 2280.
278. THE LATINO/A CONDITION, *supra* note 1, at 17.
279. OMI & WINANT, *supra* note 26, at 55.
280. *Id.* at 56 (noting that race is a matter of both structure an cultural representation).
281. Much in Delgado's piece leads me to think his understanding of the social construction of race as an "idea" focused and as a product of consciousness is related to Lawrence's notion of "unconscious racism." Charles R. Lawrence, III, *The Id, the Ego, and Equal Protection: Reckoning With Unconscious Racism*, *in* CRT: KEY WRITINGS, *supra* note 1, at 235-57. I in someone ways share this intuition, and thus see this idea as a quirky function of the development of the idea of social construction rather than something inherent.
282. Brown, *supra* note 146, at 1489-91.

sciousness should be eliminated as a strategy for addressing racial injustice. Instead, they emphasize spiritual transformation.[283] While I am sympathetic to this type of thinking, it does not provide a theory about the process of transformation, given embodiment. That is, it does not contemplate the limitations and compromises that material (human) embodiment imposes on spiritual consciousness as manifested in the world or the way constraints are further materialized in social systems and institutional structures. Nor does it suggest the process for transforming them.[284] In addition, to say race is not biological is not to say it is not real.[285] Further, to say race is a social construction, again, is not to say that race is not real. Money is the preeminent social construction as a medium of exchange. No one seems to suggest that money is not real or rather that the consequences of having or not having money in the context of modern economies is somehow unreal.[286] These ideas detach the consciousness of social construction from the materiality of social construction and they mistakenly uproot racism from its material basis as well as detach race from class.

Nevertheless the idea and prominence of racial consciousness or unconsciousness as developed within the idea of the social construction

---

283. *See* Robinson, *supra* note 238, at 155-58.

284. Here, the process of transforming these material constraints on consciousness or calling into existence a different reality requires something more than consciousness or thought but may well entail a combination of thought, speech, and action or action and speech that changes the conditions of consciousness. As such it might well require or be well served by race consciousness that produces the actions that would help dismantle the structural and material conditions that reinforce the conscious conceptions. *See generally* NEALE DONALD WALSCH, CONVERSATIONS WITH GOD: AN UNCOMMON DIALOGUE (Book 1) (1996).

285. OMI & WINANT, *supra* note 26, at 55.

286. One can argue that money is a social construction. It is certainly not biological but it is a function of human interaction. Money is a medium of exchange. Further, expressions of money, such as the one-dollar bill and a twenty-dollar bill can be said to be the same, in the way that people, are viewed the same, as humans. That is, these bills are made out of the same sort of paper, with roughly equivalent amounts of ink, thus their value from this perspective is the same. However, people do not value the twenty-dollar bill the same as they do the one-dollar bill, nor do they say the bills have equal value. This difference is real, not as a matter of biology or substance, but is real because a zillion transactions every day make them real and make the difference between the bills real. This difference is buttressed not only a billion people believing that the difference between the dollar bills is real, and acting like they are different everyday (consciousness made real,) but is also buttressed and made real by a variety of systems and institutions such as the U.S. economy and the global economy, the federal reserve bank, the stock market, the real estate agent, and the grocery store down the street, etc.

In addition, people's belief systems about the dollar in general could rapidly change, leading to its abandonment as a system or medium of exchange. However, such an outcome would likely be precipitated by some enormous event, and not because a small group of people, marginal and outside the mainstream, with limited control over the production of money and its accumulation wished or believed it so. Additionally, the value or perception of say, twenty dollars does not remain the same over time, twenty dollars thirty years ago was more valuable. And finally, and more important to analysis of class to the extent that access to money is allowed to determine a person's access to status, goods and services, the systems by which it produced and allocated creates classes, races, genders etc., designed to perform both and different economic and psychological functions. Race is structural and made real in a similar way that money is. Further, capitalism, the system we now inhabit, guides this process buttressed and ordered in part by law and makes access to money in its various forms the key to accessing resources, the absence of which helps delineate class and race.

of race, further fortifies the discursive turn in CRT. This turn, to the extent it focuses on discourse, may occlude a focus on materiality and therefore on economic and class issues.

But political and other analytical pressures also exist that may have hindered the development of class analyses within CRT and strengthened the tendency toward the discursive turn. For example, hesitancy to take on class may result from fear of analyzing or critiquing capitalism, the reigning economic order, given the political environment that champions unfettered capitalism as a panacea for all ills despite its apparent tendency to concentrate wealth in the hands of a few. This tendency uninterrupted by policy decisions to curb it or disrupt its lopsided material distributions, has increasingly created and cemented vast economic inequalities in the social system, widening and hardening the gap between the rich and the poor.[287] However, engagement with class issues as a potential critique of this situation may be seen as politically difficult and unpopular given the politics of triumphant which celebrate the failure of communism and conclude the inadequacy of Marxist analyses with their particular emphasis on class divisions.[288]

This fear may be bolstered by economic analyses which understand current economic arrangements as necessary and natural and which inform us everyday that the maintenance and creation of the rich are indispensable and beneficial to the rest of us, including the middle, working, and poor classes,[289] while suggesting that policies which put the well-

287. The London Financial Times published a three-part series discussing growing economic inequality in the United States and the stagnation of wages. *See* Krishna Guha, Edward Luce, & Andrew Ward, *Anxious Middle: Why Ordinary Americans Have Missed Out on the Benefits of Growth*, FIN. TIMES (London), Nov. 2, 2006 (commentary) (part one of three); Krishna Guha, Edward Luce, & Alim Remtulla, *Seeking Shelter: Why Democrats are in Retreat From Their Free Trade Record*, FIN. TIMES (London), Nov. 3, 2006 (commentary) (part two of three); Alan Beattie, *Sans Safety Net, 'Creative Destiny' Spurs America*, Nov. 6, 2006 (commentary) (part three of three); *see also* Lawrence Summers, *The Global Middle Cries out for Reassurance*, FIN. TIMES (London), Oct. 30, 2006 (commentary); ROBERT PERUCCI & EARLY WYSONG, THE NEW CLASS SOCIETY 3-34 (1999) (suggesting that class mobility is stagnating); JEFF FAUX, THE GLOBAL CLASS WAR: HOW AMERICA'S BIPARTISAN ELITE LOST OUR FUTURE – AND WHAT IT WILL TAKE TO WIN IT BACK 49-75 (2006) (noting the increasing gap between rich and poor and commenting that while racism, sexism and other forms of oppression seem to be decreasing, class inequalities are growing); Robert D. Atkinson, *Inequality in the New Knowledge Economy*, *in* THE NEW EGALITARIANISM 52, 54-55 (Antony Giddens & Patrick Diamond eds., 2005) (discussing growing inequality that began in the late 1970's and explaining that this phenomenon covers not only inequitable distribution of income but also the distribution of jobs with the share of jobs with middle income also shrinking). The result is an economy that is crating a U-shaped job-market with more job increases found at the top and the bottom. *Id.*

288. Litowitz, *supra* note 68, at 534-5.

289. Recall for instance, President Ronald Reagan's supply–side economics, derisively termed "trickle-down economics," which put the rich at the center of tax-reduction for the so-called benefit of greater production. Marjorie E Kornhauser, *The Morality Of Money: American Attitudes Toward Wealth and the Income Tax*, 70 IND. L.J. 119, 154-56 (1994). The rich, the theory suggested would take the money saved from lower taxes and invest it, stimulating the economy and therefore providing overall growth and benefits that would trickle down to the rest of us. *Id.* Instead, the rich got richer and the United States deficit soared, later restraining investment in social spending because the deficit had to be reduced. *Id.*; *see, e.g.*, Susan Pace Hamill, *An Evaluation of Federal Tax Policy*

being of the vast majority humanity at the center of economic ordering and development are destructive. And it may be further bolstered by the legal code of economic policy choices, described earlier, that protects the regime of "private ownership of property, employment at-will, inheritance, freedom of contract, limited liability for business organizations, patriarchy, and a regime of negative rights that ensures that individuals must secure their own health care, day care, and other benefits."[290]

At the same time, a hesitancy to take on class may arise from deep within CRT itself. Race in America cuts across class. So for instance, the working class is and has historically been divided by race.[291] But class also divides race,[292] a fact CRT recognizes in its antiessentialist notions, which call for racial solidarity and unity as a political commitment rather than as an essential fact. However, a class analysis may nonetheless require race crits to acknowledge their own class standing and interest as part of an educated elite. This interrogation might suggests among other things, that their interest in being and remaining a part of this elite class is in tension with the empowerment of the racialized underclass, which empowerment is both the producer and product of a radically transformed social order. Said differently, in order for the masses of racialized underclasses to live well, the current system which privileges the educated elite will need to be transformed. Empowerment of these classes is both the prerequisite and desired outcome of this transformation. Seen from another perspective, race crits may not look at class because they are less affected by the oppressive mechanisms of class, even as these accompany racial oppression.

There may also be a number of other analytical obstacles to CRT taking on class analyses. For instance, Litowitz argues that central insights of Marxism (as the most important theory of class relations) are no longer adequate explanations of or salient factors in advanced capitalism

---

Based on Judeo-Christian Ethics, 25 VA. TAX REV. 671, 729-734 (2006) (providing brief discussions of trickle down economics and explaining trickle down theory, noting that there is no reliable proof that tax cuts for the rich will stimulate growth and explaining that although George W. Bush calls himself a Christian that his tax policies fail to meet a Christian ethics that would require those who benefit the most from society to contribute—as part of sacrifice—more for its upkeep and requiring that revenues raised be sufficient to ensure vulnerable people an opportunity to live up to their God given potential); *see also* Audrey Farlane, *The New Inner City: Transformation, concentrated Affluence and the Obligations of the Police Power*, 8 U. PA. J. CONST. L. 1, 4-5 (2006) (discussing recent city urban planning goals of attracting a more affluent population on the theory that they will bring needed resources to the city by which poorer communities will benefit but finding that the poorer communities disappear, get relocated and do not benefit from this focus).

290. Litowitz, *supra* note 68, at 549 (referring to the code also as a worldview and explaining that "social reform involves subversion of a dominant rationality").

291. *See, e.g.*, Herbert G. Gutman, *Black Coal Miners and the Greenback-Labor Party in Redeemer, Alabama: 1878-1879, The Letters of Warren D. Kelley, Willis Johnson Thomas, "Dawson," and Others*, 10 LAB. HIST. 506, 506 (1969); Herbert G. Gutman, *Reconstruction in Ohio: Negroes in the Hocking Valley Coal Mines in 1873 and 1874*, 3 LAB. HIST. 243, 243-44 (1962). *See generally* ROEDIGER, *supra* note 116.

292. I owe this line of argument to Robert Chang.

where class lines have become blurry.[293] For example, he poses the following situation and question. "A female computer consultant for IBM speculates on the stock market at night and owns a studio apartment that she rents to a janitor. To what class does she belong?"[294] While a definitive answer may not arise under traditional Marxist theory of class, rethinking class in light of modern theories, current practices, and concrete economic developments may provide a sophisticated analysis of this scenario.[295] Equally important however, is that the demonstration of blurry class lines does not render class as an economic feature or an analytical or political tool non-existent. In fact critical race analysis may be helpful in this regard. So for instance, race, Latinos, blacks, as groups, as political projects, as analytical categories are "blurry"—in-essential, non-monolithic, unstable, socially constructed, and difficult to define. Yet, they exist because we create and recreate them structurally, politically, intellectually and discursively. Though a slightly different process is at work in the case of class, and perhaps gender, these social processes and phenomenon exists at multiple levels.

And finally, CRT may have failed to adequately engage class analysis because, as Angel Harris suggests, we may simply have lost the language for talking about class given the declining currency of Marxist theory.[296] Or it may be that in light of the law and economics tradition, CRT scholars believe it necessary and are reluctant to master another field such as economics. In any case, this failure is puzzling given CRT's suggestion that class mutually constructs race on the one hand, and on the other, given its success in contributing to the examination of other subordinating systems, such as gender oppression, (intersectionality)[297] and potentially gay, lesbian, and queer oppression (multidimensionality).[298]

B. Why ClassCrits?

My thoughts on what class is and what a critical class analysis might reveal are provisional. However, it appears important for CRT to engage and develop a critical class analysis because mapping out the different economic classes, the power relations between them, and the way law operates on them, as it has in the context of the systems of race, gender, and sexuality is crucial to understanding the ways in which class mutually supports these other systems. In addition, a class analysis may further ground examinations of racism and racist practices and potentially guide CRT's antisubordination praxis.

---

293. Litowitz, *supra* note 68, at 534.

294. *Id.* at 534; *see also* FRIEDMAN, *supra* note 263.

295. For new ideas about class, see, e.g., PERUCCI & WYSONG, *supra* note 287, at 3-34; FAUX, *supra* note 287, at 49-75 (discussing the existence of an international elite class).

296. Harris, *supra* note 42, at 777.

297. *See supra* note 18.

298. *See supra* note 19.

Classes arguably arise out of the division of labor and the stratification of society in producing, distributing, and perhaps even in consuming goods and services.[299] A critical class analysis, then, presumably not only interrogates such concerns as supply and demand, wages and productivity and ultimately how goods, services, wealth, and surplus are created and allocated, but also interrogates the power relations and power differentials between different economic groups in society and the ways they are being maintained or changed. From a critical perspective, particularly in the context of CRT, there is a special concern about oppressive class relations and ways to eliminate them as well as a focus on context. That is, a critical class analyses rejects abstract notions of efficiency, insisting instead on inquiries as to who the players are, who the beneficiaries are, whose needs are accommodated,[300] and whose needs are not, why some allocations are deemed efficient and others not, and, which is skeptical of claims of neutrality and efficiency where the crystallized protections of a privileged group status have not been investigated in the given context. In this vein, critical class theory might have something to say or to ask about the groups produced in the modern social economic processes and the current moment in which segmented joblessness and stagnating wages for the overwhelming majority of the population exist in the face of increasing productivity, a growing economy, staggering profits, and incomprehensible executive pay as mediated by the corporation and justified by an array of people institutionally and otherwise placed.[301]

In addition, a class analysis may further ground examinations of racism and racist practices because it focuses on the economic allocations of material resources accumulated over the last several hundred years, much of which has been racialized. It is unlikely that these racialized allocations (whether productive or distributive) will be eliminated without addressing the structural and economic foundations of class. Nor is it clear that class and the economic harms of lower class status can be eliminated without addressing both the material and psychological seductions embodied and structured by race. Ultimately, the best way to address these issues may not simply be calling for the edification and respect of different groups but the difficult work of building coalitions to demand concrete economic and material changes for those oppressed by race and class as well as those oppressed by class across race. Four

299. These are relatively typical elements in the definition of class. *See* FRASER, *supra* note 210, at 17 (making a similar point).

300. *See, e.g.*, Martha T. McCluskey, *Illusion of Efficiency in Worker's Compensation "Reform,"* 50 RUTGERS L. REV. 657, 666-67, 716-50 (1998) (unraveling the neoeconomic rhetoric of efficiency versus redistribution distinction makes redistribution claims seem suspect).

301. James Bernstein, *Q&A: Working Harder and Taking Home Less*, NEWSDAY, Aug. 29, 2006; Rex Nutting, *Bernanke Overstates Wage Growth: Earnings are not Catching up with Productivity, Government Figures Show*, MAKETWATCH, Jul. 19, 2006; *see also* Steven Greenhouse & David Leonhardt, *Real Wages Fail to Match A Rise in Productivity*, N.Y.TIMES, Aug. 28, 2006.

questions raised in the context of Hurricane Katrina might exemplify these points.[302]

In the wake of Hurricane Katrina, many commentators questioned whether the government's slow, seemingly disinterested response and inadequate rescue of the thousands of mostly African American people stranded in the flooded city of New Orleans was the result of racism.[303] That is, they wondered whether the government's pitiful response was not just evidence of incompetence but rather occurred because the victims were black.[304] This is a typical race consciousness question and race questions are increasingly formulated in this manner.[305] It gets at what Rory calls the sadism of the American social order. It is about the motivation and lack of it of white Americans, white officials providing justice to black Americans. It is about conscious and unconscious racism. It is about the current and momentary intent of individuals, about the intentional conduct of current players. It is often countered by efforts to shame the perpetuators, but has the effect of making invisible those affected by such policies.[306] This shaming, in turn, reinforces calls for what Fraser has called a politics of recognition, a politics that advances the edification and respect of different positioned groups.[307] A politics of recognition focuses on cultural solutions, which might include the celebration of a subordination group's culture or heroes such as in black history month, Martin Luther King Day, or advocate for participation in gay parades? But a politics of recognition, while necessary, may not address the structure of racial segmentation, stratification, and caste.

Assuming the legitimacy of the first question, a second question nonetheless arises. It is: Why did so many black people stay behind or find themselves stranded or left behind in New Orleans? Presumably, these people did not stay behind because they were black but rather because they were poor. They lacked the means required to leave New Orleans and to sustain themselves while away from their homes.[308] Why they were so poor that they could not leave in the face of a serious hurri-

---

302. H.R.J. Res. 437, 109th Cong. (2005); Elizabeth Fussell, *Understanding Katrina: Perspectives from the Social Sciences* (2005), http://understandingkatrina.ssrc.org/Fussell/ (detailing arguments highlighted here). For example, 27.3% of New Orleans households did not have cars, compared to 9.4% of the U.S. population as a whole. *See id. See generally* MICHAEL ERIC DYSON, COME HELL OR HIGH WATER: HURRICANE KATRINA AND THE COLOR OF DISASTER (2006).

303. DYSON, *supra* note 302, at 17-33 (explaining the ways in which race played a role in the failure of the federal government to respond to the Katrina hurricane victims and noting that malicious intent may not have been involved but indifference as part of a southern, national, and historical script where black grief and pain is involved are among the factors at work).

304. *Id.*

305. This may be the result of an understanding of racism as largely the product of intentional individual action, a concept that ignores race as a caste system and a limited concept of race that is promoted by colorblindness ideology. *See* Harris, *supra* note 42, at 750-52.

306. Delgado, *Two Ways of Thinking about Race*, *supra* note 17, at 2295.

307. FRASER, *supra* note 210, at 11-66.

308. DYSON, *supra* note 302, at 5-6 (discussing poverty in New Orleans and noting that one in four people in New Orleans lacked access to a car).

cane when most others in the area were able to leave, in the richest country in the world, is the class question. Though this question may ultimately implicate American sadism, it has much more to do with what Rorty calls American selfishness.[309] That is, it is no doubt about economics, but it is primarily about the policy choices and practices that structure the production and maldistribution of resources that render some groups well off and others much less so.[310]

The third question was why the majority of the poor people in New Orleans appeared to be overwhelmingly black when they constitute only a little over a third of the metropolitan area of 1.3 million people.[311] These are the questions of class and race, the way class and race are related and have developed over time through structural elements, and policies developed throughout the United States and in elements and policies specific to Louisiana.[312] Unraveling this question would presumably demonstrate the way in which racism is a co-constituent and feature of the economic order in that it further delineates and allocates certain economic functions and distributions to particular racialized and classed groups. As such, race itself is a system for allocating resources but one that is part and parcel of, or intertwined with the class system, even as it operates independently and relatedly as a belief system that provides a ready justification for the racialized results of various distributions, as well as providing cohesion to the overall system. Race provides cohesion or stability, in part, by seducing the more privileged in the class delineated by race (white working class v. black working class) to support the system in order to maintain both material and psychological privileges.[313] From this perspective, race arguably developed as part of the U.S. economic system of slavery, to legitimate both it and other forms of exclusion, and then was used, consciously and institutionally for

309. *See generally* RICHARD RORTY, ACHIEVING OUR COUNTRY: LEFTIST THOUGHT IN TWENTIETH-CENTURY AMERICA (1998).

310. *See generally id.*

311. *See New Orleans Demographics, Population (LA)*, CityRating.com, http://www.cityrating.com/citystats.asp?city=New+Orleans&state=LA (last visited Oct. 21, 2006) (discussing the metropolitan area with a population of 1.3 million and a black people constituting 37.5% of the population); *see also* Dominguez, *supra* note 28 (discussing the term of "third world" as applied to hurricane victims and discussing who was seen and unseen in this diverse metropolitan area). Other questions might include whether it was only poor blacks stranded by Hurricane Katrina. Though the majority of those stranded in New Orleans seemed to be black, were there others? Were there poor whites also affected, and if so why were they less visible? And who might be served by the invisibility of poor whites? For instance, as Dyson notes, some of the regions hardest-hit by Katrina are suffer from extreme poverty. DYSON, *supra* note 302, at 5. The include Mississippi as the poorest state in the nation and Louisiana the second poorest. *Id.* Further, more than 90,000 people in these areas as well as parts of Alabama where Katrina hit made under $10,000 a year. *Id.*

312. DYSON, *supra* note 302, at 1-14 (discussing race and poverty and explaining that those left behind in New Orleans had been socio-economically left behind years ago). That is, they suffered from poverty and New Orleans in particular like many urban other urban areas in the US, suffered from concentrated poverty. *Id.*

313. *See, e.g.*, BELL, *supra* note 2, at 8 (citing EDMUND S. MORAN, AMERICAN SLAVERY, AMERICAN FREEDOM 8 (1975)) (discussing how the development and maintenance of a poor black subclass enabled and enables poor white to identify with wealthy whites).

years to structure, perpetuate, and legitimate the resulting order of what bell hooks calls the "white supremacist capitalist patriarchy."[314] That racism can be addressed without addressing the economic class configurations that structure race is doubtful. Similarly, the kind of mass-movement needed to challenge dominant—economic-class[315] power would require dealing with the racial issues that divide those who must work for their living as opposed to those, along with those with specialized skills, who live off the work of others.[316]

And finally, there is the question of the evacuation order, a policy that seemed blind to the classed realities of poverty in New Orleans and yet help to make visible the racialized maldistribution of resources in New Orleans. Though the government was said to have had an evacuation plan for New Orleans that recognized that over a hundred thousand people might be stranded and included government efforts to locate and move people out of the city who were unable to do so themselves, this was not the evacuation plan carried out. Rather, the evacuation plan, bungled no doubt, consisted primarily of an order for people to vacate New Orleans, an order that which while recognizing that some people might get stuck, assumed that most people had adequate and enough wealth to evacuate. It was a policy and plan based on assumptions of adequate wealth where for many no such wealth existed and in a country where poverty levels and gaps between the rich and the poor, though ignored, are increasing. That a policy of assumed wealth often disadvantages a racialized population reflects the way that classed policies have racial impact.

In a similar way, consider the famous case of *San Antonio Independent School District v. Rodriguez*.[317] Here the Supreme Court upheld a school financing system that allowed one district's schools to be vastly under funded as compared to another.[318] The under-funded school district consisted mostly of Mexican Americans.[319] The Court's decision did not rely on race but rather on wealth and assumptions of adequate

314. bell hooks, KILLING RAGE: ENDING RACISM 78 (1995).

315. *See, e.g.*, Manning Marable, INTRODUCTION: BLACK STUDIES AND THE RACIAL MOUNTAIN, *in* DISPATCHES FROM THE EBONY TOWER: INTELLECTUALS CONFRONT THE AFRICAN AMERICAN EXPERIENCE 1, 19 (Manning Marable ed., 2000) (explaining a transformationist idea as one seeks to "transform the existing power relationships and the racist institutions of the state, the economy and society," and noting that this requires "the building of a powerful protest movement, based largely among the most oppressed classes and social groups, to demand the fundamental restructuring of the basic institutions and patterns of ownership within society[]").

316. Here I am relying on Jeff Faux description of the global elite class as comprising those with specialized skills, those who control large corporations and those who live off their capital earnings as opposed to living off of their work. *See* FAUX, *supra* note 287, at 64.

317. 411 U.S. 1 (1973).

318. *Rodriguez*, 411 U.S. at 5-6.

319. *Id.* at 12.

wealth.[320] The effect was that the racialized class of poor people would continue to be so because they would attend inferior schools that would render inferior jobs and thus maintain them as a relatively poor racialized group. In this sense, race crits as lawyers should analyze class because law makes decisions on the basis of wealth that have racial implications and makes decisions on the basis of race that have wealth and class implications.

These last three questions relating to race, class, and race- and class-based policies and laws, may also better guide CRT's antisubordination praxis. That is, it may counsel the kind of activities that are more effective in eliminating both racial and class subordination. These activities may be those that do not simply call for the respect of subordinated groups but seek to challenge, change, or transform the social arrangements that structure the subordination itself. Fraser argues that the valuation and respect of different groups is more likely in an egalitarian society, a society where significant material inequalities between groups do not exist.[321] She contends that though a politics of recognition, which call for respect of differently positioned groups is important, a "politics of redistribution" that seeks to change the processes and conditions that create and structure the devaluation of differently-positioned individuals and groups is a more effective strategy for accomplishing the edification of different groups in a pluralist society.[322] Rorty agrees, noting that sadism may be more effectively undermined by policies and activities that attack the selfish maldistribution of resources that create classes and help to racialized groups.[323] The difference in strategy may be the difference between creating an outcry to force an incompetent and potentially racist politician to resign his post after the hurricane or celebrating the cultural mix needed to create a rich culture such as New Orleans, in distinction to creating a movement to push for policies such as new and affordable housing built on higher-ground, plans to locate, track and help return poor evacuees back to the city, improved health care, adequate living wages, greatly expanded public transportation system and/or increased credit or other opportunities to buy cars, etc. Surely the first is helpful, but the latter is indispensable.

CONCLUSION

I have suggested that critical race theory rose in response to, and as a challenge to the ascendance of a colorblind ideology, an ideology that seeks to ignore the structural, persistent and current manifestations of

320. I owe the insight of "assumed wealth" to Leslie Bender & Daan Braveman, editors of *Power, Privilege and Law. See* POWER, PRIVILEGE AND LAW, *supra* note 1, at 381-98 (discussing Braveman's analysis of *San Antonio Independent School District v. Rodriguez*).

321. FRASER, *supra* note 210, at 27-33.

322. *Id.*

323. *See generally* RORTY, *supra* note 309.

racism that inure from the social historical process of race formation in the United States. The goal of colorblind ideology, while ostensively to promote equality as symmetry among the races, in fact cements the privileges of whiteness. CRT's insights about colorblindness, as well as a number of its other themes and methodologies result in part from a process of conflict inspired growth which provided CRT scholars the opportunity to experience the meaning and impact of their insights as part of their development. These conflicts also however, in some ways spur the development of CRT related scholarship. This scholarship while potentially fragmenting the movement has in fact contributed much to the theoretical project of CRT. This scholarship has expanded the CRT knowledge base of different groups, historical experiences, it has expanded CRT's knowledge around different kinds of oppression, and it has aided CRT in the development of new theories. In doing so this scholarship has also strengthened CRT's commitment to coalition building and antisubordination praxis.

Classcrits becomes another frontier for CRT. The expectation is that classcrits will expand CRT knowledge about another type of subordination and set of power relations that, like gender and sexuality, is differently constructed than race, even as it mutually reinforces race. In addition, classcrits becomes another way, in the face of the expansion of colorblind ideology, to get, in part, at the materiality and economic foundations of race. That is, while colorblind advocates continue to attempt to reduce race to who did what, the question reappears in the form of why the poor are disproportionately black or why blacks are disproportionately poor. And though a cultural reply that blames the victim is expected, the tools will be developed to challenge these replies in ever more sophisticated economic and class terms. And finally, though strategies that attempt to overcome white cultural domination and promote black and other non-white edification continue, these efforts will be complemented by coalitions built across race and across types of oppression to challenge the material, classed and economic structures and resource allocations that shape them.

BACK FROM THE MARGINS: AN ENVIRONMENTAL NUISANCE PARADIGM FOR PRIVATE CLEANUP COST DISPUTES

RONALD G. ARONOVSKY†

ABSTRACT

The law governing private cleanup cost disputes is in disarray. For over twenty years, owners of contaminated property voluntarily cleaned up pollution caused by others in reliance on a federal law right under CERCLA to recover from those who contributed to the contamination their fair share of cleanup costs. That changed with the U.S. Supreme Court's decision in Cooper Industries v. Aviall Services, Inc. Aviall *held that liable persons under CERCLA (such as the current owners of contaminated property) who voluntarily incur cleanup costs cannot sue for contribution under CERCLA, calling into question any future role for federal law in most private cleanup cost disputes. Current state law fails to offer a meaningful alternative to federal law for allocating cleanup costs at many of the nation's hundreds of thousands of contaminated sites. This need not be so.*

Private nuisance is a flexible doctrine that is potentially well-suited for resolving cleanup cost allocation disputes. Doctrinal limitations and inter-state inconsistencies, however, currently shackle private nuisance law and prevent its application to many common problems involving contaminated property. CERCLA has pushed nuisance and other state law theories to the margins of private cleanup cost litigation. After Aviall, *private nuisance could emerge to play a major role in resolving private cleanup cost disputes, but only if states seize the opportunity to re-examine and modernize the law of private nuisance in soil and groundwater contamination cases. This article analyzes the deficiencies that currently riddle private nuisance law and proposes an environmental nuisance paradigm. The paradigm would expand private nuisance law to address contamination created by prior property owners and encourage site cleanup through a rebuttable presumption that chemical contamination constitutes a continuing nuisance that remains actionable until the contamination is abated.*

---

† Associate Professor of Law, Southwestern University School of Law. J.D., 1980, University of California, Berkeley. A.B., 1977, University of California, Berkeley. The author thanks Dean Bryant G. Garth for his continued support and encouragement, Austen L. Parrish, Clifford Rechtschaffen, Laurel S. Terry and Dennis T. Yokoyama for their helpful and constructive comments on prior drafts of this article, and Elika Jon Decker, Amanda Kent, Maia Spotts and Erin Stratte for their excellent research assistance.

TABLE OF CONTENTS

INTRODUCTION

The law governing private cleanup cost disputes is in disarray. The U.S. Supreme Court's decision in *Cooper Industries, Inc. v. Aviall Services, Inc. (Aviall)*,[1] called into question the availability of federal cleanup cost contribution rights under the Comprehensive Environmental Response, Compensation and Liability Act (CERCLA).[2] In *Aviall*, the Court held that a potentially responsible party (PRP)[3] who voluntarily spent millions of dollars in cleanup costs could not use CERCLA's contribution provision[4] to recover from another PRP who caused much of the contamination its fair share of cleanup costs.[5] In so doing, *Aviall* cast doubt on the future role of federal law in private cleanup cost disputes.[6]

---

1. Cooper Indus., Inc. v. Aviall Servs., Inc., 543 U.S. 157 (2004).

2. Pub. Law No. 96-510, 94 Stat. 2767 (1980) (codified as amended at 42 U.S.C.A. §§ 9601-9675 (West 2006)). CERCLA is sometimes referred to as the "Superfund" statute. *See infra* notes 42-63 and accompanying text.

3. Individuals or entities falling within one of the four categories of "covered persons" set forth in (CERCLA sections 107(a)(1)-(4)), are often referred to as PRPs. *See infra* notes 46-51 and accompanying text.

4. 42 U.S.C.A. § 9613(f)(1).

5. *Aviall*, 543 U.S. at 167. *See infra* notes 77-83 and accompanying text; *see also* Ronald G. Aronovsky, *Federalism and CERCLA: Re-Thinking the Role of Federal Law in Private Cleanup Cost Disputes*, 33 ECOLOGY L. Q. 1, 45-50 (2006) (discussing the *Aviall* decision).

6. As used in this article, the phrase "private cleanup cost disputes" refers to disputes between private parties regarding the allocation of contaminated property cleanup cost responsibilities. These disputes may concern reimbursement (by a direct action for damages or a derivative action for contribution) by a PRP of its fair share of cleanup costs already expended by a current landowner or

The absence of a federal cleanup cost contribution right, of course, would not implicate major national environmental policy concerns if PRPs could recover cleanup costs from other PRPs under state law. Often, however, no such state law remedy is available.[7] It should be. Private nuisance is *potentially* well suited for resolving contaminated property cleanup cost allocation disputes. Doctrinal limitations and interstate inconsistencies, however, currently shackle private nuisance law by restricting its application to neighboring land use disputes and shielding from liability a defendant's continued failure to abate long-standing contamination.

Indeed, the inadequacies of nuisance law (and other state law theories) helped inspire CERCLA's passage in 1980 to provide for the remediation of contaminated sites.[8] For over two decades, CERCLA pushed nuisance and other state law theories to the margins of private cleanup cost litigation as the owners of contaminated property and other PRPs voluntarily incurred cleanup costs in reliance on CERCLA cleanup cost contribution rights. *Aviall* proved this reliance was misplaced. The states should seize the opportunity created by *Aviall* to modernize the law of private nuisance by eliminating anachronistic limitations that currently prevent it from playing a vital role in common cleanup cost allocation disputes.

States should not ignore this opportunity. The underlying problem is significant: there are hundreds of thousands of contaminated sites across the United States[9] that will cost hundreds of billions of dollars to remediate.[10] Voluntary cleanup of these sites is essential because state and federal government agencies lack the resources either to directly remediate the pollution at these sites or force private parties to clean them up by prosecuting thousands of regulatory cleanup obligation enforcement lawsuits.[11] Cleanup cost recovery rights are crucial to promoting voluntarily remediation[12] because at most sites more than one PRP contributed to site contamination.[13]

---

other PRP. They may also concern the allocation of future cleanup cost responsibilities, including claims for (a) damages covering future cleanup costs, (b) declaratory relief allocating future cleanup cost responsibilities, or (c) injunctive relief directing one or more PRPs to remediate contamination.

7. *See infra* notes 150-360 and accompanying text.

8. *See infra* notes 35-41 and accompanying text.

9. *See infra* notes 20-23 and accompanying text.

10. *See infra* notes 24-25 and accompanying text.

11. *See* Aronovsky, *supra* note 5, at 42-43, 56-57. As used in this article, the phrase "voluntary cleanups" refers to remediation undertaken by a private party short of a court order, i.e., at the party's own initiative or at the direction of a regulatory agency.

12. *See* James B. Brown & Michael V. Sucaet, *Environmental Cleanup Efficiency: Private Recovery Actions for Environmental Response Costs*, 7 T.M. COOLEY L. REV. 363, 387 (1990) ("A liable party is more likely to expend cleanup funds if it is reasonably certain that response costs expended on contamination may be recovered. Moreover, a PRP armed with knowledge of legal rights against others is more likely to initiate a prompt cleanup."); *cf.* Andrew R. Klein, *Hazardous Waste Cleanup and Intermediate Landowners: Reexamining the Liability-Based Approach*, 21

Cleanup cost contribution is particularly important for owners of contaminated property. Current landowners typically bear regulatory responsibility to clean up their property, including pollution caused by others.[14] Cleanup cost contribution rights thus can affect both the incentive of current property owners to voluntarily take the lead in site cleanup short of government enforcement action and the fairness of cleanup cost responsibility allocation.

This article proposes an environmental nuisance paradigm[15] that would allow private nuisance to play a significant role in allocating the costs of cleaning up the nation's polluted soil and groundwater. This paradigm, if adopted by state courts or legislatures, would transform the ancient law of nuisance to meet the realities of modern contamination disputes in the following respects. First, it would permit private nuisance claims against the former owners of contaminated property, not just neighboring owners. Second, it would eliminate the antiquated doctrine of *caveat emptor* as a defense to former owner nuisance liability while continuing to allow market-related factors to affect nuisance remedies. Third, it would revitalize the continuing nuisance doctrine by basing it on a defendant's continued failure to abate the soil or groundwater contamination caused by the defendant. Finally, it would create a rebuttable presumption that soil and groundwater contamination constitutes a continuing nuisance by placing on any party contending that the contamination was a permanent nuisance the burden of proving that the contamination could not reasonably be abated.

Under this paradigm, private nuisance claims would compliment any private cleanup cost remedies remaining under CERCLA after *Aviall*. The inadequacy of current state law requires a uniform federal remedy to ensure nationwide availability of contribution rights for PRPs who voluntarily incur cleanup costs.[16] Such a CERCLA "safety net" remedy,

---

HARV. ENVTL. L. REV. 337, 346 (1997) (noting that "[j]oint and several liability, in particular, can lead to PRPs facing exposure far out of proportion to any damage they actually caused").

13. *See, e.g.*, ENVTL. LAW INST., AN ANALYSIS OF STATE SUPERFUND PROGRAMS: 50 STATE STUDY, 1998 UPDATE 33 ("Most hazardous substance sites have more than one potentially responsible party."); *see also* Key Tronic Corp. v. United States, 511 U.S. 809, 819 n.13 (1994) ("CERCLA is designed to encourage private parties to assume financial responsibility of cleanup by allowing them to seek recovery from others."); Atl. Research Corp. v. United States, 459 F.3d 827, 836 (8th Cir. 2006) ("Contribution is crucial to CERCLA's regulatory scheme.").

14. For example, CERCLA section 107(a)(1) imposes response cost liability on the current owner of contaminated property even if the current owner did not cause any of the site contamination. 42 U.S.C.A. § 9607(a)(1) (West 2006).

15. As used in this article, the phrase "environmental nuisance paradigm" refers to a model framework for private nuisance claims regarding the allocation of remediation or cleanup cost responsibilities for soil or groundwater contamination of plaintiff's property.

16. In Aronovsky, *supra* note 5, from which some background information in this article is drawn, I argue that while significant litigation and remedial efficiency and flexibility interests could be served if state law provided the primary source for rules of decision in private cleanup cost disputes (as described *infra* at notes 133-48 and accompanying text), a federal "safety net" cleanup cost remedy for all PRPs is necessary because of the inconsistencies and doctrinal limitations of current state law. Aronovsky, *supra* note 5 at 68-79. This article addresses how private nuisance law, as

however, would not necessarily provide a rule of decision superior to private nuisance law in contaminated property disputes. To the contrary, the proposed paradigm for private nuisance law could: (a) facilitate prompt site remediation through the presumption that the contamination can be abated; (b) create incentives for informal cleanup cost dispute resolution by expanding current landowner cleanup cost rights; (c) promote efficiency by permitting all contaminated property related claims be resolved under a common body of state law; and (d) encourage flexibility by allowing PRPs to pursue technically sound, cost-effective cleanups.[17] The paradigm's benefits thus extend far beyond addressing the aftermath of *Aviall*. At some point, the uncertainty created by *Aviall* about the availability of PRP cost recovery rights under CERCLA will be resolved by federal legislation or case law.[18] Whether or not all PRPs ultimately come to enjoy a federal right to cleanup cost contribution, states should adopt the proposed environmental nuisance paradigm to promote litigation and remedial efficiency and flexibility.

This article argues that states should adopt this proposed paradigm by statute or case law. Part I describes the contaminated property problem facing the United States, the emergence of CERCLA in the 1980s as the primary rule of decision in private cleanup cost disputes, and the uncertain future role of federal law in contaminated property litigation after *Aviall*.[19] Part II evaluates current nuisance law as applied to private cleanup cost disputes (including an analysis of its doctrinal and practical limitations) and argues that states should seize the opportunity caused by post-*Aviall* uncertainty to claim a significant role for private nuisance law in cleanup cost disputes. Part III proposes the means to accomplish this goal through adoption of the proposed environmental nuisance paradigm, and analyzes how the paradigm would expand cleanup cost remedies for the current owner of contaminated property, encourage prompt and efficient site remediation, and promote the informal resolution of cleanup cost disputes.

---

revitalized by adoption of the proposed environmental nuisance paradigm, could serve these efficiency and flexibility goals and offer a meaningful (if not superior) alternative to a CERCLA "safety net" remedy for current owners of contaminated property.

17. For example, the right to recover cleanup costs under a private nuisance theory would not be dependent on a current landowner plaintiff proving that claimed costs were incurred in a manner consistent with the National Contingency Plan (NCP), 40 C.F.R. pt. 300 *et seq.*, (2005), a set of U.S. Environmental Protection Agency (EPA) regulations setting forth the often costly and time-consuming procedures required for conducting a cleanup under CERCLA. *See* 42 U.S.C.A. § 9607(a)(4)(B) (private party may only recover response costs that are consistent with NCP). *See infra* notes 54, 63, 140-45 and accompanying text.

18. *See infra* notes 88-121, 133-48 and accompanying text.

19. As used in this article, the phrase "primary rule of decision" refers to the body of law most frequently relied on by private parties to resolve disagreements regarding the allocation of cleanup costs. For a discussion of the emergence of federal law as the primary rule of decision in private cleanup cost disputes, see Aronovsky, *supra* note 5, at 9-35.

I. BACKGROUND

A. The Scope of the Problem

Soil and groundwater contamination is an enormous national problem.[20] Estimates of the number of contaminated sites throughout the country reach into the hundreds of thousands,[21] from seriously contaminated National Priorities List (NPL)[22] sites to properties presenting relatively discrete remediation issues. All levels of government—federal, state and local—are involved in regulating the investigation and cleanup of contaminated sites. The United States Environmental Protection Agency (EPA) is actively involved at only a small percentage of the nation's contaminated sites;[23] state and local government authorities serve as the lead agency at the vast majority of sites.[24] It will take decades to remediate these sites, costing hundreds of billions of dollars.[25]

---

20. *See* Aronovsky, *supra* note 5, at 7-9.

21. The EPA has estimated that approximately 77,000 contaminated sites have been discovered throughout the United States, and that approximately 217,000 additional sites across the country will eventually be identified. EPA, CLEANING UP THE NATION'S WASTE SITES: MARKETS AND TECHNOLOGY TRENDS viii (2004), *available at* http://www.clu-in.org/download/market/2004market.pdf [hereinafter EPA, CLEANING UP THE NATION'S WASTE SITES]. The EPA estimates that there are "more than 450,000 brownfields" in the United States. EPA, Brownfields Cleanup and Redevelopment, http://www.epa.gov/swerosps/bf/about.htm (last visited Aug. 31, 2006); *accord* U.S. GEN. ACCOUNTING OFFICE, COMMUNITY DEVELOPMENT: LOCAL GROWTH ISSUES—FEDERAL OPPORTUNITIES AND CHALLENGES 118 (2000), *available at* http://www.gao.gov/new.items/rc00178.pdf. CERCLA section 101(39)(A) defines "brownfield" as property, "the expansion, redevelopment, or reuse of which may be complicated by the presence or potential presence of a hazardous substance, pollutant, or contaminant." 42 U.S.C.A. § 9601(39)(A) (West 2006).

22. *See* 42 U.S.C.A. § 9605(a)(8)(B) (requiring that EPA list "national priorities among the known or threatened releases of hazardous substances throughout the United States" and "revise the list no less than annually"). As of October 2006, 1,243 sites were listed on the NPL, with another 61 sites proposed for addition to the NPL. *See* EPA, NPL Site Totals by Status and Milestone, http://www.epa.gov/superfund/sites/query/queryhtm/npltotal.htm (last visited Nov. 19, 2006).

23. According to the EPA, the "vast majority" of contaminated sites will be cleaned up under state (rather than federal) regulatory authority oversight. *See* EPA, State and Tribal Response Programs, http://www.epa.gov/swerosps/bf/state_tribal.htm (last visited Nov. 7, 2006) (quoting S. REP. NO. 107-2 (2001)) [hereinafter EPA, State and Tribal Response Programs]; *see also* EPA, CLEANING UP THE NATION'S WASTE SITES, *supra* note 21, at viii (estimating that 90% of contaminated sites will likely be managed under state or underground storage tank programs).

24. EPA, State and Tribal Response Programs, *supra* note 23; *see also* Richard L. Revesz, *Federalism and Environmental Regulation: A Public Choice Analysis*, 115 HARV. L. REV. 553, 596–98 (2001) (describing state regulatory agency role at most contaminated sites); Roger D. Schwenke, *Applying and Enforcing Institutional Controls in the Labyrinth of Environmental Requirements - Do We Need More Than the Restatement of Servitudes to Turn Brownfields Green?*, 38 REAL PROP. PROB. & TR. J. 295, 297 (2003) ("[S]tates are responsible for the bulk of environmental enforcement activities, including contamination detection, site remediation, and notification requirements."); Philip Weinberg, *Local Environmental Laws: Forging a New Weapon in Environmental Protection*, 20 PACE ENVTL. L. REV. 89, 107 (2002) (discussing use of local environmental controls to fill state and federal regulatory gaps).

25. *See, e.g.*, EPA, CLEANING UP THE NATION'S WASTE SITES, *supra* note 21, at viii (estimating that 294,000 sites will need cleanup during the next three decades at a cost of $209 billion).

At most of the nation's sites more than one PRP contributed to site contamination.[26] A site may have been contaminated by the activities of successive owners or operators of that property. For example, several companies may have operated businesses disposing of hazardous wastes on a parcel of industrial property. Similarly, hundreds or thousands waste generators may have sent hazardous substances to a landfill. At other sites, contamination on one parcel may have been caused by the migration (e.g., in groundwater or surface water, or through the air) of contaminants from a neighboring parcel. An environmental regulatory agency will often name the current owner as a respondent on a cleanup order at multi-PRP sites, requiring the current owner to comply with regulatory requirements, or risk administrative or judicial enforcement proceedings as well as severe penalties for non-compliance.[27] Private cleanup cost disputes arise when the current owner or another PRP at a multi-PRP site incurs cleanup costs in excess of what the PRP perceives to be its fair share, but is unable to informally obtain reimbursement or contemporaneous cost contribution from other PRPs.

B. Private Cleanup Cost Disputes Before CERCLA

Before CERCLA was enacted in 1980, federal law did not provide a private cleanup cost dispute remedy.[28] Instead, state law provided the sole rule of decision governing private disputes regarding the allocation of responsibility for soil or groundwater contamination.[29] Four common law theories provided the primary potential bases for private cleanup cost claims[30]: negligence,[31] nuisance,[32] trespass,[33] and strict liability for ultra-hazardous activity.[34]

---

26. ENVTL. LAW INST., AN ANALYSIS OF STATE SUPERFUND PROGRAMS: 50 STATE STUDY, 1998 UPDATE 33 ("Most hazardous substance sites have more than one potentially responsible party.").

27. *See, e.g.*, 42 U.S.C.A. § 9606(b)(1) (West 2006) (imposing $25,000 per day fines for non-compliance absent "sufficient cause"); 42 U.S.C.A. § 9607(c)(3) (authorizing treble cost punitive damages for failure "without sufficient cause" to "properly provide removal or remedial action upon order of the President" under CERCLA sections 104 or 106).

28. *See* Aronovsky, *supra* note 5, at 9-12 (discussing legal framework of pre-CERCLA private cleanup cost disputes); *see also* Robert B. Mckinstry, Jr., *The Role of State "Little Superfunds" in Allocation and Indemnity Actions Under the Comprehensive Environmental Response, Compensation and Liability Act*, 5 VILL. ENVTL. L.J. 83, 86 (1994) ("Prior to 1980, no federal legislation existed which addressed past disposals of hazardous wastes; all existing laws were directed only at regulating current activity."); Steven T. Singer, *An Analysis of Common Law and Statutory Remedies for Hazardous Waste Injuries*, 12 RUTGERS L.J. 117, 147 (1980) (before CERCLA, no federal statute "would support a claim for damages from toxic chemicals").

29. *See, e.g.*, Theodore Baurer, *Love Canal: Common Law Approaches to a Modern Tragedy*, 11 ENVTL. L. 133 (1980) (discussing common law theories potentially applicable to the Love Canal site); Singer, *supra* note 28, at 122–38 (reviewing available common law theories of recovery).

30. Cleanup cost claims (e.g., claims for breach of contract, waste, misrepresentation) also can arise as a consequence of contractual privity. *See* Klein, *supra* note 12, at 374 ("The cause of action that most directly protects 'vertical' landowners in hazardous waste litigation is fraud."); *see also* Ronald G. Aronovsky, *Liability Theories in Contaminated Groundwater Litigation*, 1 J. ENVTL. FORENSICS 97, 111-113 (2000) (discussing common law theories).

31. *See, e.g.*, Ewell v. Petro Processors of La., Inc., 364 So.2d 604, 606 (La. Ct. App. 1978) (defendants negligently permitted toxic waste to leak from disposal pits onto plaintiff's property); P.

With growing awareness in the 1970s about the extent of the nation's soil and groundwater contamination problems, the perception grew that existing state law could not adequately address site remediation and the allocation of cleanup responsibilities.[35] Tort law varied dramatically from state to state. Moreover, for reasons equally applicable today, common law tort theory often fit poorly with the realities of private cleanup cost disputes. First, nuisance, trespass, and negligence all required proof of culpability,[36] while strict liability for ultra hazardous activity required an unpredictable multi-factored analysis of whether the defendant's activity was abnormally dangerous.[37] Second, common law tort theories typically required proof that the defendant's conduct caused plaintiff's damage (e.g. cleanup costs)—a potentially significant evidentiary problem at older contaminated sites.[38] Third, a range of liability defenses could prevent recovery under common law tort theories.[39] The

---

Ballantine & Sons v. Pub. Serv. Corp. of N.J., 91 A. 95, 96 (N.J. 1914) (tar products escaping from defendant's plant leached through soil to groundwater that migrated to plaintiff's well).

32. *See, e.g.*, Assoc. Metals & Minerals Corp. v. Dixon Chem. & Research, Inc., 197 A.2d 569, 580 (N.J. Super. Ct. App. Div. 1963) (escape of sulfur dust to neighboring property); Helms v. E. Kan. Oil Co., 169 P. 208, 208-09 (Kan. 1917) (migration of refinery oil and other hazardous materials to neighboring property).

33. *See, e.g.*, Curry Coal Co. v. M.C. Arnoni Co., 266 A.2d 678, 683 (Pa. 1970) (sludge dumped on ground surface seeped into mine); Burr v. Adam Eidemiller, Inc., 126 A.2d 403, 405-08 (Pa. 1956) (water from defendant's spraying slag pile contaminated plaintiff's underground water supply); Elsey v. Adirondack & St. Lawrence R.R. Co., 161 N.Y.S. 391, 393 (Sup. Ct. 1916) (subsurface migration of pollutants from defendant's railroad embankment to plaintiff's property).

34. *See, e.g.*, Cities Serv. Co. v. State, 312 So. 2d 799, 803 (Fla. Dist. Ct. App. 1975) (breach of waste reservoir damaging public waters); Atlas Chem. Indus., Inc., v. Anderson, 514 S.W.2d 309, 314-15 (Tex. App. 1974), *aff'd*, 524 S.W.2d 681 (Tex. 1975) (contaminants released into stream that crossed plaintiff's land), *abrogated by* Neely v. Cmty. Prop., Inc., 639 S.W.2d 452 (Tex. 1982). *See generally* Alexandra B. Klass, *From Reservoirs to Remediation: The Impact of CERCLA on Common Law Strict Liability Environmental Claims*, 39 WAKE FOREST L. REV. 903, 917-18, 934 (2004) (discussing pre-CERCLA strict liability environmental contamination cases).

35. A House of Representatives report on the CERCLA legislation concluded that: "Existing state tort laws present a convoluted maze of requirements under which a victim is confronted with a complex of often unreasonable requirements with regard to theories of causation, limited resources, statutes of limitations and other roadblocks that make it extremely difficult for a victim to be compensated for damages." H.R. REP. NO. 96-1016, at 63-64 (1980), *reprinted in* 1980 U.S.C.C.A.N. 6119, 6140-41.

36. *See infra* notes 151, 153, 174 and accompanying text.

37. *See* RESTATEMENT (SECOND) OF TORTS § 519 (1977) (engaging in an activity which is abnormally dangerous subjects the actor to strict liability for harm caused to his neighbors resulting from the abnormally dangerous character of the activity, even though the actor has exercised the utmost care and has acted without negligence). The Restatement identifies six factors as guidelines to determine whether an activity is abnormally dangerous:

> (a) existence of a high degree of risk of some harm to the person, land or chattels of others; (b) likelihood that the harm that results from it will be great; (c) inability to eliminate the risk by the exercise of reasonable care; (d) extent to which the activity is not a matter of common usage; (e) inappropriateness of the activity to the place where it is carried on; and (f) extent to which its value to the community is outweighed by its dangerous attributes.

Id. § 520.

38. *See* Gary Milhollin, *Long-Term Liability for Environmental Harm*, 41 U. PITT. L. REV. 1, 6 (1979) ("Whether the theory is negligence, nuisance or strict liability, the plaintiff must prove that it was the defendant's act which caused the harm.").

39. These potential liability defenses included the doctrine of *caveat emptor*, discussed *infra* at notes 222-46 and accompanying text.

statute of limitations in particular often barred private cleanup cost claims at many sites where contamination occurred decades ago.[40] In sum, by the end of the 1970s, the nation had begun to appreciate the magnitude of its soil and groundwater contamination problem, and the inadequacy of the existing state law framework for addressing the problem. In December 1980, Congress passed CERCLA, which for the next twenty-five years provided a uniform, nationally applicable rule of decision for private cleanup cost disputes.[41] That is, until the U.S. Supreme Court decided the *Aviall* case in 2004.

C. CERCLA

1. CERCLA Liability

The liability scheme established by CERCLA in 1980 dramatically departed from state law governing cleanup cost disputes.[42] CERCLA liability does *not* require proof of causation.[43] It imposes status-based, strict,[44] and retroactive[45] liability on four categories of "covered per-

40. *See infra* notes 247-59 and accompanying text. The "discovery rule" (tolling the statute of limitations until plaintiff knows or has reason to know of her claim) may preserve otherwise time-barred claims involving older contamination problems. *See, e.g.*, McKelvey v. Boeing N. Am., Inc., 86 Cal. Rptr. 2d 645, 651 (Cal. Ct. App. 1999) (under California law, the period of limitations runs "without regard to whether the plaintiff is aware of the specific facts . . . , provided that he has a 'suspicion of wrongdoing,' which he is charged with once he has 'notice or information of circumstances to put a reasonable person on inquiry.'" (citations omitted)).

41. *See* Aronovsky, *supra* note 5 at 12-35 (discussing evolution of pre-*Aviall* PRP cleanup cost claim case law under CERCLA).

42. *See* Atl. Research Corp. v. United States, 459 F.3d 827, 830 (8th Cir. 2006) (CERCLA, enacted to encourage timely cleanup of hazardous waste sites and place cleanup costs on those responsible for contamination, "effectively transformed centuries of real property and tort liability law by making those who contaminate a site strictly liable for the costs of subsequent cleanup by others."); *see also* Ronald G. Aronovsky & Lynn D. Fuller, *Liability of Parent Corporations for Hazardous Substance Releases Under CERCLA*, 24 U.S.F. L. REV. 421, 425-26 (1990) (describing enactment of CERCLA).

43. *See, e.g.*, Kalamazoo River Study Group v. Menosha Corp., 228 F.3d 648, 656–57 (6th Cir. 2000). However, a defendant may show a lack of causation as part of a divisibility affirmative defense. *Id.*; *see also infra* notes 60, 71 and accompanying text. CERCLA section 107(b) provides a "covered person," including the current landowner who did not contribute to site contamination, with only three affirmative defenses to liability: (1) act of God; (2) act of war; or (3) act of a third party. 42 U.S.C.A. §§ 9607(b)(1)-(4) (West 2006). Equitable defenses (e.g., laches, estoppel) do not bar CERCLA liability. *See, e.g.*, W. Props. Serv. Corp. v. Shell Oil Co., 358 F.3d 678, 692–93 (9th Cir. 2004) ("[E]quitable defenses such as laches are not available as a bar to section 107(a) liability"). CERCLA also exempts certain activities from liability. *See, e.g.*, 42 U.S.C.A. § 9607(j) (federally permitted releases); 42 U.S.C.A. § 9607(i) (federally registered pesticide discharges); 42 U.S.C.A. § 9607(d) (persons acting pursuant to the NCP or following orders given by an on-site response coordinator appointed under the NCP). In 2002, Congress added two narrow affirmative defenses to CERCLA liability, found in sections 107(o) (*de micromis* generators of waste at NPL sites before April 1, 2001) and 107(q) (certain owners or operators of properties contiguous to up-gradient contamination sources).

44. *See, e.g.*, New York v. Shore Realty Corp., 759 F.2d 1032, 1042 (2d Cir. 1985) ("Congress intended that responsible parties be held strictly liable . . . .").

45. *See, e.g.*, United States v. Alcan Aluminum Corp., 315 F.3d 179, 188-89 (2d Cir. 2003) (holding that CERCLA liability is retroactive and that application of retroactive CERCLA liability is constitutional), *cert. denied*, 540 U.S. 1103 (2004).

sons:"[46] (1) the current owner and operator[47] of contaminated property;[48] (2) anyone who owned or operated contaminated property when it was polluted;[49] (3) anyone who arranged to dispose of hazardous substances on another's property;[50] and (4) anyone who transported a hazardous substance to the contaminated property.[51] CERCLA applies to the release or threatened release of "hazardous substances," a broadly defined term covering a wide variety of pollutants.[52] Petroleum, however, is expressly excluded from CERCLA's definition of "hazardous substances."[53] The national contingency plan (NCP),[54] a set of EPA regulations, sets forth the procedure for responding to a release of hazardous substances under CERCLA.[55]

2. Cost Recovery

CERCLA provides regulatory agencies with several enforcement remedies. EPA may issue a unilateral administrative order requiring a PRP (i.e., a CERCLA covered person[56]) to undertake specific remedia-

46. CERCLA section 107(a)(1)-(4) identifies the four categories of covered persons. 42 U.S.C.A. § 9607(a)(1)-(4).

47. CERCLA defines the "owner or operator" of a facility as "any person owning or operating such facility," excluding "a person who, without participating in the management of a vessel or facility, holds indicia of ownership primarily to protect his security interest in the . . . facility." 42 U.S.C.A. § 9601(20)(A).

48. Section 107(a)(1) imposes liability on the current owner or operator of a "facility." 42 U.S.C.A. § 9607(a)(1). Section 101(9)(B) broadly defines "facility" to include "any site or area where a hazardous substance has been deposited, stored, disposed of or placed, or otherwise come to be located [except] any consumer product in consumer use or any vessel." 42 U.S.C.A. § 9601(9)(B). Liability under section 107(a)(1) applies to a current owner of contaminated property without regard to whether the current owner caused any site contamination. 42 U.S.C.A. §§ 9601(9)(B), 9607(a)(1).

49. Section 107(a)(2) imposes liability on "any person who at the time of disposal of any hazardous substance owned or operated any facility at which such hazardous substances were disposed of." 42 U.S.C.A. § 9607(a)(2). Section 101(21) defines "person" to include "an individual, firm, corporation, association, partnership, consortium, joint venture, commercial entity, United States Government, State, municipality, commission, political subdivision of a State, or any interstate body." 42 U.S.C.A. § 9601(21). Section 101(29) incorporates the definition of "disposal" used in section 1004 of the Solid Waste Disposal Act, 42 U.S.C.A. § 6903(3) (West 2006), which provides that: "[t]he term 'disposal' means the discharge, deposit, injection, dumping, spilling, leaking, or placing of any solid waste or hazardous waste into or on any land or water so that such solid waste or hazardous waste or any constituent thereof may enter the environment or be emitted into the air or discharged into any waters, including ground waters." 42 U.S.C.A. § 9601(29).

50. 42 U.S.C.A. § 9607(a)(3). Courts have expansively interpreted the scope of "arranger" liability to include such persons as those who arrange to dispose of hazardous substances at landfills, toll formulators, and persons who send material containing hazardous substances to recyclers with the knowledge that some of the material will not be returned. *See* Aronovsky, *supra* note 5, at 13 n.48.

51. 42 U.S.C.A. § 9607(a)(4).

52. 42 U.S.C.A. § 9601(14)

53. *Id.*

54. National Oil and Hazardous Substances Pollution Contingency Plan (NCP), 40 C.F.R. pt. 300 (2005).

55. 42 U.S.C.A. § 9605. *See* 40 C.F.R. § 300.1 (the purpose of NCP is "to provide the organizational structure and procedures for preparing for and responding to discharges of oil and releases of hazardous substances, pollutants, and contaminants").

56. *See supra* notes 46-51 and accompanying text.

tion tasks.[57] The government may also conduct the cleanup itself,[58] and then sue under section 107(a), CERCLA's direct cost recovery provision, to recover its costs of responding to the release or threatened release of hazardous substances (response costs).[59] Under section 107(a)(4)(A), a PRP is jointly and severally liable[60] for "all costs of removal or remedial action incurred by the United States Government or a State or an Indian tribe not inconsistent with the national contingency plan."[61] Section 107(a)(4)(B) permits a private party who incurs response costs to bring a cost recovery action,[62] providing that PRPs are liable for "any other necessary costs of response incurred by *any other person* consistent with the national contingency plan."[63]

---

57. 42 U.S.C.A. § 9606(a).

58. 42 U.S.C.A. § 9604(a)(1).

59. 42 U.S.C.A. § 9607(a)(4)(A). CERCLA refers to cleanup costs incurred in response to the release or threatened release of hazardous substances as "response costs." *See* 42 U.S.C.A. § 9601(25) (defining "response"); 42 U.S.C.A. § 9607(a)(4) (referring to "costs of response").

60. 42 U.S.C.A. § 9607(a)(4)(A). CERCLA is silent regarding the scope of response cost liability. By the mid-1980s, courts generally concluded that liability to the government under CERCLA section 107(a) was joint and several. *See* John M. Hyson, *"Fairness" and Joint and Several Liability in Government Cost Recovery Actions Under CERCLA*, 21 HARV. ENVTL. L. REV. 137, 150–60 (1997) (discussing evolution of CERCLA joint and several liability case law). The courts similarly came to hold that liability under section 107(a) for costs incurred by private parties was joint and several. *See, e.g.*, Bedford Affiliates v. Sills, 156 F.3d 416, 423-24 (2d Cir. 1998) (explaining section 107(a)(4)(B) claim imposes joint and several liability at sites with indivisible harm); Rumpke of Ind., Inc. v. Cummins Engine Co., 107 F.3d 1235, 1240 (7th Cir. 1997) (explaining section 107(a)(4)(B) claim imposes joint and several liability and may be maintained by current landowner who did not add contamination to site).

61. 42 U.S.C.A. § 9607(a)(4)(A). Defendants can argue as an affirmative defense that liability should be apportioned severally if the harm at the contaminated site is divisible. *See, e.g.*, United States v. Alcan Aluminum Corp., 964 F.2d 252, 270-71 (3d Cir. 1992) (explaining several liability appropriate if PRP proves divisibility and reasonable basis for apportionment).

62. Claims for personal injury or property damages caused by hazardous substance contamination are not available under CERCLA. *See, e.g.*, Artesian Water Co. v. Gov't of New Castle County, 659 F. Supp. 1269, 1285 (D. Del. 1987) ("Congress in enacting CERCLA clearly manifested an intent not to provide compensation for economic losses or for personal injury resulting from the release of hazardous substances.").

63. 42 U.S.C.A. § 9607(a)(4)(B) (emphasis added). The phrase "any other person" refers to persons other than the United States, States, or Indian tribes who have a right of cost recovery under section 107(a)(4)(A). Post-*Aviall* decisions have disagreed about whether the term "other" in the phrase "any other person" also excludes PRPs, i.e., "covered persons" falling within the categories of liable parties set forth in § 9607(a)(1)-(4). *Compare, e.g.*, *Atl. Research*, 459 F.3d 827 at 835-36. ("'any other person' means any person other than the statutorily enumerated 'United States Government or a State or an Indian tribe.'" (citation omitted)), *and* Consol. Edison Co. v. UGI Utils., 423 F.3d 90, 99-100 (2d Cir. 2005) (holding that a PRP who voluntarily incurred cleanup costs could state a section 107(a)(4)(B) direct cost recovery claim, reasoning that it would be inappropriate to impose an "innocence" condition on the "any other person" language of section 107(a)(4)(B)), *with* Aviall Servs., Inc., v. Cooper Indus., Inc., No. 3:97-CV-1926-D, 2006 U.S. Dist. LEXIS 55040, at *21-25 (N.D. Tex. Aug. 8, 2006) (granting Cooper's motion for summary judgment following remand from U.S. Supreme Court, holding that Aviall as PRP could not maintain a section 107(a)(4)(B) claim and reasoning after "examining CERCLA holistically" that the phrase "any other person" in section 107(a)(4)(B) refers to any person other than section 107(a)(1)-(4) covered persons as well as section 107(a)(4)(A) entities). *See also* Aronovsky, *supra* note 5, at 82, n.350.

3. The Section 107 / Section 113 Conundrum

In 1986, Congress passed the Superfund Amendments and Reauthorization Act (SARA).[64] The SARA amendments included two express CERCLA contribution provisions: (1) section 113(f)(1), which provides that any person may seek contribution during or after an EPA administrative order judicial enforcement action[65] or cost recovery action,[66] and (2) section 113(f)(3)(B), which provides that any person may seek contribution after settling its CERCLA liability with the government.[67] Section 113(f) was intended to "clarify" and "confirm" CERCLA contribution rights.[68] Section 113(f) did not, however, expressly address the most common CERCLA "contribution" plaintiff: a current landowner (or other PRP) who voluntarily incurs cleanup costs but has neither been sued in an EPA judicial enforcement action or cost recovery action,[69] nor settled its CERCLA liabilities with the government.[70]

---

64. Superfund Amendments and Reauthorization Act of 1986, Pub. L. No. 99-499, 100 Stat. 1613 (1986).

65. CERCLA section 106 authorizes the federal government to (a) issue unilateral administrative orders directing parties to investigate and remediate site contamination and (b) initiate a judicial enforcement action in the event of non-compliance. 42 U.S.C.A. § 9606. Few private plaintiffs would have contribution rights triggered by a section 106 action. Section 106 actions may only be brought by the federal government, which is actively involved at only a handful of the nation's contaminated sites. *See supra* note 23 and accompanying text; *see also* Larry Schnapf, *Impact of* Aviall *on Real Estate and Corporate Transactions*, 20 TOXICS L. REP. (BNA) 607, 610 (2005) ("[S]tates bring over 70 percent of enforcement actions and the vast majority of contaminated sites are remediated under [state Superfund] programs . . . ."). Even at sites where the EPA plays an active role, there would be no reason to initiate a section 106 action unless a PRP was not in compliance with a section 106 administrative order or otherwise was not adequately responding to site contamination.

66. 42 U.S.C.A. § 9613(f)(1). Section 113(f)(1) provides that:

> [a]ny person may seek contribution from any other person who is liable or potentially liable under section 9607(a) of this title, during or following any civil action under section 9606 of this title [EPA administrative order enforcement action] or under section 9607(a) [cost recovery action] of this title. Such claims shall be brought in accordance with this section and the Federal Rules of Civil Procedure, and shall be governed by Federal law. In resolving contribution claims, the court may allocate response costs among liable parties using such equitable factors as the court determines are appropriate. Nothing in this subsection shall diminish the right of any person to bring an action for contribution in the absence of a civil action under section 9606 of this title or section 9607 of this title.

42 U.S.C.A. § 9613(f)(1).

67. 42 U.S.C.A. § 9613(f)(2)-(3). Section 113(f)(2) contemplates a settlement by a "person who has resolved its liability to the United States or a State " while section 113(f)(3)(B) allows the settling PRP to pursue CERCLA contribution against non-settling PRPs. 42 U.S.C.A. § 9613(f)(2)-(3).

68. S. REP. NO. 99-11, at 4 (1985) (objective of proposed new contribution provision was to "clarif[y] and confirm[] the right of a person held jointly and severally liable under CERCLA to seek contribution from other potentially liable parties, when the person believes that it has assumed a share of the cleanup or cost that may be greater than its equitable share under the circumstances"); H.R. REP. NO. 99-253, at 79 (1985), *reprinted in* 1986 U.S.C.C.A.N. 2835, 2861.

69. A PRP, rather than a government agency or non-PRP private party, usually takes the lead in site cleanup; accordingly, a PRP seeking contribution likely would not first have been sued for cost recovery under section 107(a). A PRP may be sued under section 107(a) by a non-PRP (1) in the rare circumstance where an "innocent" private party incurs cleanup costs, 42 U.S.C.A. § 9607(a)(4)(B); *cf.* City of Bangor v. Citizens Comm'ns Co., 437 F. Supp. 2d 180, 222-23 (D. Me.

This statutory interpretation issue, however, did not appear to trouble the courts or PRPs. After the SARA amendments, the issue before the courts was not whether a PRP could recover cleanup costs from other PRPs. Rather, litigation arose across the country about which provision of CERCLA provided the basis for such an action—the direct cost recovery provision of section 107(a)(4)(B) or the contribution provisions of section 113(f)(1).[71] By the time the Supreme Court decided *Aviall* in December 2004, each court of appeals to have addressed this section 107 / section 113 conundrum had held that a plaintiff who was a liable party under CERCLA could not bring a direct action for cost recovery under CERCLA section 107(a)(4)(B).[72] These courts assumed, expressly or

---

2006) (holding that a PRP may bring a section 107(a) action and noting that "it is hard to imagine many cases in which purely 'innocent parties' would ever be motivated to initiate an action under section 107"); or (2) where a government agency seeks to recover its direct cleanup costs or its costs of overseeing a private cleanup. 42 U.S.C.A. § 9607(a)(4)(A); United States v. E.I. Dupont De Nemours & Co., 432 F.3d 161, 179 (3d Cir. 2005) (en banc) (federal government oversight costs recoverable under section 107(a)(4)(A)).

70. The last sentence of section 113(f)(1) provides that "[n]othing in this subsection shall diminish the right of any person to bring an action for contribution in the absence of a civil action under [sections 106 or 107(a)]." 42 U.S.C.A. § 9613(f)(1). CERCLA and its legislative history, however, say nothing about what "contribution" rights (e.g., direct or implied contribution rights under CERCLA section 107(a), state contribution law contribution rights) were left "undiminished" by section 113(f)(1).

71. *See* Aronovsky, *supra* note 5, at 24-33 (discussing the section 107/section 113 conundrum). This issue mattered to CERCLA defendants, who argued that claims by one CERCLA-liable party against another sounded in contribution and thus should be brought under section 113(f)(1). They reasoned it would be unfair for a liable plaintiff to impose a joint and several liability direct cost recovery section 107(a) claim, *see supra* note 60, which would place on defendants (a) the burden of showing the divisibility of environmental harm caused by the defendant's conduct, *see, e.g.*, Pneumo Abex Corp. v. High Point, Thomasville & Denton R.R. Co., 142 F.3d 769, 776 & n.4 (4th Cir. 1998) (instructing district court under section 113(f) to allocate costs according to appropriate equitable factors but, unlike a joint and several liability action under section 107(a), not to impose a divisibility of harm allocation burden on defendants); and (b) the risk of any "orphan shares" of liability. *See* Aronovsky, *supra* note 5, at 24-26. An "orphan share" is the equitable share of cleanup cost liability attributable to a PRP that is unable to pay, such as a PRP who cannot be located or who is insolvent, deceased or bankrupt. *See, e.g.*, Pinal Creek Group v. Newmont Mining Corp., 118 F.3d 1298, 1303 (9th Cir. 1997) (holding that a PRP cannot assert a joint and several section 107(a) claim against other PRPs, because "those defendant-PRPs would end up absorbing all of the cost attributable to 'orphan shares'—those shares attributable to PRPs who either are insolvent or cannot be located or identified").

72. As of December 2004, the First, Second, Third, Fourth, Sixth, Seventh, Eighth, Ninth, Tenth, and Eleventh Circuits each had held that a PRP could *not* bring a section 107(a) cost recovery action against another PRP, and that a PRP seeking to recover cleanup costs from another PRP under CERCLA was limited to a contribution action for several liability under section 113(f). *See, e.g.*, United Techs. Corp. v. Browning-Ferris Indus., Inc., 33 F.3d 96, 100-01 (1st Cir. 1994); Bedford Affiliates v. Sills, 156 F.3d 416, 425 (2d Cir. 1998); New Castle County v. Halliburton NUS Corp., 111 F.3d 1116, 1120 (3d Cir. 1997); Pneumo Abex Corp. v. High Point, Thomasville & Denton R.R. Co., 142 F.3d 769, 776 (4th Cir. 1998); Centerior Serv. Co. v. Acme Scrap Iron & Metal Corp., 153 F.3d 344, 356 (6th Cir. 1998); Rumpke of Ind., Inc. v. Cummins Engine Co., 107 F.3d 1235, 1239–40 (7th Cir. 1997) (section 107(a) action also available to current landowner PRP who did not contribute to site contamination; all other PRPs limited to section 113(f) contribution action); Dico, Inc. v. Amoco Oil Co., 340 F.3d 525 (8th Cir. 2003); Pinal Creek Group v. Newmont Mining Corp., 118 F.3d 1298, 1306 (9th Cir. 1997); United States v. Colo. & E. R.R. Co., 50 F.3d 1530, 1536 (10th Cir. 1995); Redwing Carriers v. Saraland Apartments, 94 F.3d 1489, 1513 (11th Cir. 1996). The Fifth Circuit's three-judge panel in *Aviall* concluded that "a PRP cannot file a section 107(a) suit against another PRP; it must pursue a contribution action instead." Aviall Servs., Inc. v. Cooper Indus., Inc., 263 F.3d 134, 137 (5th Cir. 2001), *reh'g en banc*, 312 F.3d 677 (2002), *rev'd*, 543 U.S. 157 (2004).

implicitly, that the PRP plaintiff instead could and should sue for contribution under section 113(f).[73] As a result, a current owner of contaminated property or other PRP who had taken the lead cleaning up a multi-PRP site typically filed a section 113(f)(1) cleanup cost contribution action, and CERCLA contribution law became the primary rule of decision in private cleanup cost disputes.[74] Owners of contaminated property, other PRPs, and government agencies alike came to rely on the nationwide availability of a CERCLA cleanup cost remedy.[75] CERCLA contribution rights provided an incentive for PRPs to voluntarily comply with regulatory agency cleanup orders. Similarly, businesses factored the potential for obtaining cleanup costs from PRPs in deciding whether to acquire or develop brownfield or other real property that was or might be contaminated.[76] All this was changed when the U.S. Supreme Court decided *Aviall*.

This decision was supplanted by the Fifth Circuit's *en banc* decision in *Aviall* that ultimately was reversed by the Supreme Court. *Id.*; *see also* Amoco Oil Co. v. Borden, Inc., 889 F.2d 664, 672 (5th Cir. 1989) ("When one liable party sues another to recover its equitable share of response costs, the action is one for contribution, which is specifically recognized under CERCLA" (citing section 113(f)). *But see* OHM Remediation Servs. v. Evans Cooperage Co., 116 F.3d 1574, 1582 n.1 (5th Cir. 1997) ("We express no opinion . . . whether a PRP may seek to hold other parties jointly and severally liable under section 107(a) for response costs."); *Aviall*, 2006 U.S. Dist. LEXIS 55040, at *11-14 (granting Cooper's motion for summary judgment following remand from the U.S. Supreme Court, holding that Aviall as PRP could not maintain a section 107(a)(4)(B) claim, reasoning that the Fifth Circuit's *en banc* decision in *Aviall* was not binding on the district courts after the U.S. Supreme Court's *Aviall* decision so that in light of *OHM* no binding Fifth Circuit decision had yet addressed whether a PRP could bring a section 107(a) action against another PRP).

73. *See* Aronovsky, *supra* note 5, at 26-35. These courts implicitly or explicitly assumed that a section 113(f)(1) contribution would be available to a PRP, regardless of whether the PRP first had been sued under CERCLA section 106 or 107(a) or had settled its CERCLA liabilities with the government. *See, e.g.*, *Atl. Research*, 459 F.3d at 832-33 (describing as judicial "[t]raffic-directing" the pre-*Aviall* cases that limited PRPs to a section 113(f) contribution action, noting that "[i]n the pre-*Aviall* analysis, section 113 was presumed to be available to all liable parties including those which had not faced a CERCLA action"); City of Waukesha v. Viacom, Inc., 221 F. Supp. 2d 975, 979 (E.D. Wis. 2002) (explaining absence of reported decisions on the section 113(f)(1) standing issue may reflect the common understanding among the bench and bar that such an action was available to any PRP); Adobe Lumber, Inc. v. Hellman, 415 F. Supp. 2d 1070, 1076 (E.D. Cal. 2006) (permitting PRP plaintiff to proceed with section 107(a) action and noting that "the Ninth Circuit's pre-*Aviall* precedents assumed a cost recovery suit was not a prerequisite for a § 113(f) contribution action").

74. *See* Aronovsky, *supra* note 5, at 33-35. Sometimes plaintiffs would join state law claims to fill gaps in the CERCLA statutory framework. For example, state law cleanup cost claims might be asserted under an alternative state law liability scheme or to provide a vehicle for recovering petroleum remediation or other cleanup costs that could not be recovered under CERCLA. *See, e.g.*, William W. Watts, *Common Law Remedies in Alabama for Contamination of Land*, 29 CUMB. L. REV. 37, 39 (1999) (describing use of common law theories for cost claims regarding Alabama sites as alternative to CERCLA); Michael B. Hingerty, *Property Owner Liability for Environmental Contamination in California*, 22 U.S.F. L. REV. 31, 37-42, 63-82 (1987) (describing California common law and statutory theories potentially applicable to landowners). Similarly, claims for property damage (e.g., diminution in value, stigma, lost use of property) or personal injury could not be asserted under CERCLA and thus could only be brought under state tort law.

75. *See* Aronovsky, *supra* note 5, at 33-36.

76. *See id.* at 54 n.248.

4. The *Aviall* Decision

Aviall Services, Inc., the buyer of contaminated properties from Cooper Industries, Inc., brought a section 113(f) contribution claim against Cooper to recover costs Aviall incurred after it was directed by a state regulatory agency to clean up contamination to which both companies contributed.[77] In December 2004, the U.S. Supreme Court held that section 113(f) permitted a PRP to bring a contribution action only if (1) the PRP plaintiff had first been sued by the federal government under CERCLA section 106 to enforce a CERCLA administrative order, or by a government or private party under CERCLA section 107(a) for cost recovery; or (2) the PRP plaintiff had first settled her CERCLA liability with the government.[78] Because Aviall had neither been sued under CERCLA nor settled with the government, the Court held that Aviall could not maintain a section 113(f) contribution action.[79]

The Court, however, did not go so far as to hold that Aviall was barred from stating any response cost claim under CERCLA. Aviall had urged the Court in the alternative to find that Aviall could state a direct cost recovery claim under section 107(a)(4)(B) as "any other person" incurring response costs.[80] The Court chose not to decide the issue,[81] and instead remanded the case to the Fifth Circuit to address whether Aviall had waived any section 107(a)(4)(B) claim[82] and, if not, whether a PRP could assert a direct cost recovery claim under section 107(a)(4)(B).[83]

---

77. *Aviall*, 543 U.S. at 163-64. For a detailed discussion of the *Aviall* litigation, see Aronovsky, *supra* note 5, at 35-49.

78. *Aviall*, 543 U.S. at 167-68. In 2000, the District Court had granted Cooper's motion for summary judgment, holding that the first sentence of section 113(f)(1) limited contribution claims to plaintiffs (unlike Aviall) who had been first been sued under sections 106 or 107(a). Aviall Servs. Inc. v. Cooper Indus., Inc., No. Civ.A.397CV1926D, 2000 WL 31730, at *2-4 (N.D. Tex. Jan. 13, 2000). The court declined to exercise supplemental jurisdiction over Aviall's state law claims. *Id.* at *5. In 2001, a divided three-judge panel of the Fifth Circuit affirmed the district court's ruling. *Aviall*, 263 F.3d at 134. Rehearing the case *en banc* in 2002, the Fifth Circuit by a 10-3 vote reversed the district court's decision. *Aviall*, 312 F.3d at 677 (en banc).

79. *Aviall*, 543 U.S. at 171.

80. *Id.* at 168.

81. *Id.* at 168-70. In a dissenting opinion joined by Justice Stevens, Justice Ginsburg argued that the Court should have proceeded to decide whether Aviall as a PRP could state a claim under section 107(a)(4)(B). *Id.* at 170, 173-74 (Ginsburg, J., dissenting).

82. The Fifth Circuit ultimately concluded that no such waiver had occurred and remanded the case to the District Court for further proceedings. *See infra* note 88.

83. *Aviall*, 543 U.S. at 168-70. The Court noted that among the issues that might be considered on remand was whether Aviall "may pursue a section 107 cost recovery action for some form of liability other than joint and several." *Id.* at 169-70. The Court also declined to decide whether Aviall had an implied right to contribution under section 107(a) or federal common law. *Id.* at 170-71. On February 15, 2005, the Fifth Circuit ordered the case remanded to the district court with instructions to permit Aviall to amend its complaint to assert a section 107(a) claim without prejudice to Cooper's defense that such an amendment would fail to state a claim for which relief may be granted. *Aviall*, 2006 U.S. Dist. LEXIS 55040, at *7-8. On August 8, 2006, the district court granted Cooper's motion for summary judgment, holding that Aviall as PRP could not maintain a direct section 107(a)(4)(B) claim because the phrase "any other person" in section 107(a)(4)(B) referred to persons other than section 107(a)(1)-(4) covered persons as well as section 107(a)(4)(B)

Aviall caused a sea change in contaminated property law. It upset years of reliance by the regulated community that a PRP could always obtain cleanup cost contribution under CERCLA.[84] It created widespread uncertainty about the availability of federal cleanup cost recovery rights by leaving unresolved whether the lower federal courts in their pre-*Aviall* decisions had correctly barred PRPs from bringing direct cost recovery actions.[85] *Aviall* also effectively extended an invitation, which this article accepts, to evaluate how state law may enter center stage in private cleanup cost disputes by overcoming the current inconsistencies and doctrinal limitations of common law tort theory.

D. The Aftermath of Aviall

Was the *Aviall* decision a reasonable interpretation of a muddled statute? Perhaps. Should the bench, bar, and regulatory agencies have assumed for nearly twenty years that it was beyond peradventure that a PRP who had not first been sued under CERCLA could bring a section 113(f)(1) contribution action? Perhaps not. Nevertheless, the *Aviall* decision stunned environmental cleanup cost dispute stakeholders—they simply did not see it coming.[86]

1. Uncertain Future of PRP Contribution Claims

By declining to decide whether a PRP could maintain a section 107(a)(4)(B) action, the Supreme Court created profound uncertainty about whether most PRPs could recover from other PRPs their fair share of cleanup costs under CERCLA. Every court of appeals to have addressed the issue before *Aviall* had held that a PRP could not maintain a section 107(a)(4)(B) action—but did so under the express or implied assumption that the PRP instead could bring a section 113(f) contribution action.[87]

Following *Aviall*, federal courts across the country began to grapple with whether appellate decisions barring section 107(a)(4)(B) claims by PRPs retained their precedential value after *Aviall*. Some district courts (including the district court upon remand in *Aviall*) relied on pre-*Aviall* case law to hold that a PRP could not recover costs under section

---

governmental and Indian tribe plaintiffs, *id.* at *24, 29, and that there was no implied contribution right under section 107(a)(4)(B), *id.* at *36.

84. *See* E.I. DuPont de Nemours & Co. v. United States, 460 F.3d 515, 523 (3d Cir. 2006) (noting the "understanding at the time [before the *Aviall* decision was that] . . . a PRP that voluntarily cleaned up a contaminated site *sua sponte* could seek contribution from other PRPs without waiting for an enforcement action, a Government or innocent landowner cost recovery suit, or a settlement of liability. [¶] In *Cooper Industries*, the Supreme Court significantly altered this understanding."); *see also supra* note 72 and accompanying text.

85. *See supra* notes 71-73 and accompanying text.

86. *See supra* note 72; *see also* Aronovsky, *supra* note 5, at 33-58.

87. *See* Aronovsky, *supra* note 5, at 50-54.

107(a)(4)(B).[88] Other district courts after *Aviall* held that a PRP could assert a direct[89] or implied[90] claim under CERCLA section 107(a), or permitted amended pleadings adding a section 107(a) claim pending guidance from its circuit court of appeals.[91] The United States government, which had filed an *amicus curiae* brief in *Aviall* arguing that a PRP who had not first been sued under CERCLA was barred from bringing a section 113(f)(1) contribution action,[92] contended in CERCLA cases

88. *See, e.g.*, *Aviall*, 2006 U.S. Dist. LEXIS 55040, at *24, 29, 36 (granting Cooper's motion for summary judgment following remand by U.S. Supreme Court based on the "plain meaning" of the statute and pre-*Aviall* case law); Amcal Multi-Housing, Inc. v. Pacific Clay Products, No. EDCV-06-280-SGL, 2006 WL 3016326, at *12 (C.D. Cal. Oct. 10, 2006) (holding that a PRP "cannot bring a free-standing section 107 implied contribution claim [in light of pre-*Aviall* Ninth Circuit precedent], and cannot bring a section 107 cost recovery claim as it has failed to allege sufficient facts to bring it within one of the statutorily defined defenses to PRP designation status."); ITT Indus. v. Borgwarner, Inc., No. 1:05-CV-674, 2006 WL 2460793, at *3-5 (W.D. Mich. Aug. 23, 2006) (granting motion to dismiss because *Aviall* did not undermine prior Sixth Circuit precedent barring PRPs from bringing section 107(a) actions); Columbus McKinnon Corp. v. Gaffey, No. H-06-1125, 2006 WL 2382463, at *4 (S.D. Tex. Aug. 16, 2006) (granting defendant's motion to dismiss a PRP's section 107(a) claim because "under the prevailing law at this time [plaintiff] as a PRP does not have a viable claim for cost recovery under § 107(a) of CERCLA"); Spectrum Int'l Holdings, Inc. v. Universal Coops., Inc., Civ. No. 04-99 (MJD/AJB), 2006 WL 2033377, at *5 (D. Minn. July 17, 2006) (granting defendant's motion for summary judgment based on pre-*Aviall* Eighth Circuit precedent that a PRP is barred from asserting a section 107(a) claim). *See also* Aronovsky, *supra* note 5, at 51 n.244 (identifying pre-May 2006 cases holding that, notwithstanding *Aviall*, PRPs could not recover response costs under section 107(a)).

89. *See, e.g.*, Glidden Co. v. FV Steel & Wire Co., Nos. 05C1355, 05C1356, 2006 WL 2724049, at *4-5 (E.D. Wis. Sept. 21, 2006) (reversing bankruptcy court ruling that PRP claimants could not assert section 107(a) claims against bankruptcy estate); City of Martinsville v. Masterwear Corp., No. 1:04-cv-1994-RLY-WTL, 2006 WL 2710628, at *2, 3 (S.D. Ind. Sept. 20, 2006) (holding current landowner could maintain section 107(a) claim under Seventh Circuit's "innocent landowner exception," discussed in Aronovsky, *supra* note 5, at 30-33, permitting section 107(a) claims by landowners who did not contribute to site contamination); City of Bangor v. Citizens Comm'ns Co., 437 F. Supp. 2d 180, 222-23 (D.Me.2006) (holding that a PRP may bring a section 107(a) action and noting that "it is hard to imagine many cases in which purely 'innocent parties' would ever be motivated to initiate an action under section 107."); *see also* Aronovsky, *supra* note 5, at 52 n.245 (identifying pre-May 2006 cases holding that, after *Aviall*, PRPs could recover costs under section 107(a)).

90. *See, e.g.*, Raytheon Aircraft Co. v. United States, 435 F. Supp. 2d 1136, 1141-1151 (D. Kan. 2006) (holding that a PRP who incurred cleanup costs pursuant to an administrative order on consent should recover those costs under section 113(f)(3)(B), and as to other cleanup costs a PRP who cannot state a claim under section 113(f) nonetheless has an implied contribution claim under section 107(a)); Aggio v. Aggio, No. C 04-4357 PJH, 2005 WL 2277037, at *6 (N.D. Cal. Sept. 19, 2005) (holding that PRP has an implied right of contribution under section 107(a)).

91. *See, e.g.*, Gen. Motors Corp. v. United States, No. Civ.A. 01-CV-2201, 2005 WL 548266, *4-5 (D.N.J. Mar. 2, 2005) (dismissing plaintiff PRP's section 113(f)(1) claim in light of *Aviall* but granting leave to amend complaint to add section 107(a) claim because defendant would suffer no prejudice until Third Circuit decided whether to revisit holding in *New Castle County v. Haliburton NUS Corp.*, 111 F.3d 1116, 1121 (3d Cir. 1997), that a PRP may not bring a section 107(a) cost recovery action).

92. *See* Aronovsky, *supra* note 5, at 43-45. The United States is itself a PRP facing claims by PRPs at sites across the country for billions of dollars in cleanup costs. *Id.* at 44. Aviall had argued to the Supreme Court that the federal government in its capacity as a multi-site PRP could avoid paying its fair share of cleanup costs at sites nationwide if PRPs could not sue the United States or other PRPs for cost recovery under section 107(a) or contribution under section 113(f). *Id.* at 44 n.212. *See also infra* notes 106-12 and accompanying text.

across the country (both as a defendant and as *amicus curiae*) that PRPs could not recover cleanup costs under section 107(a).[93]

The issue remained muddled as courts of appeals began to weigh in on the debate and conflicts arose among the circuits. The Second Circuit, the first court of appeals after *Aviall* to address the issue, chose to evaluate PRP section 107(a) claims on a case-by-case basis, focusing on whether the PRP plaintiff "voluntarily" incurred claimed cleanup costs.[94] In *Consolidated Edison Co. v. UGI Utilities*,[95] the Second Circuit held that a PRP who incurred cleanup costs before entering into a voluntary cleanup agreement with a state agency—without first having been sued, allocated response costs by a court, or made to participate in an administrative proceeding—could bring a section 107(a)(4)(B) action.[96] Shortly thereafter, in *Schaefer v. Town of Victor*,[97] the Second Circuit held that a contaminated landfill owner could assert a CERCLA section 107(a) claim[98] against defendants who disposed of waste at the landfill even

93. *See, e.g.*, Brief of Appellee United States of America, Atl. Research Corp. v. United States, No. 05-3152, 2005 WL 3568541 (8th Cir. Dec. 5, 2005) (argument by appellee United States that district court properly dismissed CERCLA claims of PRP plaintiff under sections 107(a) and 113(f)); Brief of the United States as Amicus Curiae, Metro. Water Reclamation Dist. of Greater Chicago v. N. Am. Galvanizing and Coatings, Inc., No. 05-3299, 2006 WL 1354188 (7th Cir. May 1, 2006) (brief of Department of Justice and Environmental Protection Agency arguing, in light of a "reexamination" of government's position after *Aviall*, a PRP cannot bring a section 107(a)(4)(B) claim because (1) the phrase "any other person" in section 107(a)(4)(B) refers to persons other than section 107(a) "covered persons" as well as state, federal and tribal governments referenced in section 107(a)(4)(A); (2) section 113(f) provides the exclusive authorization for CERCLA contribution claims; and (3) there should be no exception permitting current landowners who did not contribute to site contamination to assert section 107(a)(4)(B) claims); Brief for the Federal Appellees, E.I. DuPont de Nemours & Co. v. United States, No. 04-2096 (3d Cir. Apr. 22, 2005) (PRP could not assert a cleanup cost contribution claim against the government pursuant to section 107(a) because (1) a cleanup cost claim by a PRP is necessarily a contribution claim, *id.* at 24–26; (2) section 107(a)(4)(B) standing alone does not provide an express right to contribution, *id.* at 27–28, 48–50; (3) there is no implied right to a contribution under section 107(a)(4)(B), *id.* at 28–47; and (4) the federal government's sovereign immunity bars any federal common law cleanup cost contribution claim against it, *id.* at 51–52); *see also Atl. Research*, 459 F.3d at 836-37 (8th Cir. 2006) (holding that PRP may assert section 107(a) claim against the United States, noting that a contrary ruling would result "in an absurd and unjust outcome" because "the government could insulate itself from responsibility for its own pollution by simply declining to bring a CERCLA cleanup action or refusing a liable party's offer to settle"). *But see* E. I. DuPont de Nemours and Company v. United States, 460 F.3d 515, 541 n. 31 (3d Cir. 2006) (holding that PRP DuPont may not bring section 107(a) action against the United States despite argument that such a ruling could allow the federal government to avoid contribution liability at sites where it is a PRP, observing that "DuPont does not, however, provide evidence that the EPA actually uses its enforcement discretion to avoid subjecting other federal agencies to potential liability in a later contribution suit").

94. Consol. Edison Co. v. UGI Utils., 423 F.3d 90, 100-02 (2d Cir. 2005).

95. 423 F.3d 90 (2d Cir. 2005).

96. *Consol. Edison*, 423 F.3d at 100-02. The court found it inappropriate to impose an "innocence" condition on the "any other person" language of section 107(a)(4)(B). *Id.* at 99. The court distinguished its pre-*Aviall* decision in *Bedford Affiliates v. Sills*, 156 F.3d 416, 424 (2d Cir. 1998) (holding that a CERCLA claim by a PRP who had entered into CERCLA consent decrees was "a quintessential claim for contribution" which could be brought only under section 113(f)), on the ground that the *Bedford Affiliates* plaintiff had incurred response costs only *after* having entered into a consent decree). *Id.* at 100-03.

97. 457 F.3d 188 (2d Cir. 2006).

98. *Schaefer*, 457 F.3d at 201-02. The court went on to hold, however, that Schaefer's CERCLA claim was barred by the statute of limitations. *Id.* at 210.

though the owner had entered into a series of consent orders regarding his site.[99] The court reasoned that because the owner had started incurring response costs *before* entering into the consent orders, his response costs were not incurred *solely* due to a court or administrative order imposing liability.[100]

The Eighth Circuit took a different approach by directly addressing the section 107(a) issue in *Atlantic Research Corp. v. United States*.[101] The district court had granted the United States' motion to dismiss Atlantic's section 107(a) claim based on *Dico, Inc. v. Amoco Oil Co.*,[102] a pre-*Aviall* Eighth Circuit decision holding that a PRP could not bring an action under section 107(a).[103] The Eighth Circuit reversed, concluding that *Dico's* "analytic is undermined by *Aviall*"[104] and holding that a PRP who voluntarily incurred cleanup costs for which it may be held liable may bring a direct cost recovery action under section 107(a).[105]

The Third Circuit came to the opposite conclusion in *E. I. DuPont de Nemours and Company v. United States*.[106] DuPont voluntarily undertook to clean up a site it owned in New Jersey that formerly had been

---

99. *Id.* at 192.

100. *Id.* at 201-02. By drawing such fine distinctions, the Second Circuit created new uncertainty regarding PRP cleanup cost claims. After *Consolidated Edison* and *Schaefer,* district courts in the Second Circuit wrestled with whether a PRP incurred response costs in a sufficiently "voluntary" manner to permit their recovery under a section 107(a) claim. *Compare, e.g.*, Seneca Meadows, Inc. v. ECI Liquidating, Inc., 427 F. Supp. 2d at 284-88 (W.D.N.Y. 2006) (holding that a PRP who incurred cleanup costs in connection with consent order with state agency could assert a several liability section 107(a)(4)(B) claim because the PRP had not been sued, admitted liability or fault, or been threatened with an imminent judicial or administrative liability finding), *and* City of New York v. N.Y. Cross Harbor R.R. Terminal Corp., No. 98CV7227ARRRML, 2006 WL 140555, at *4 n.6 (E.D.N.Y. Jan. 17, 2006) (holding that although plaintiff if sued would be held liable under section 107(a), it could maintain a section 107(a) claim because it conducted a voluntary investigation and cleanup without first having been sued or made to participate in an administrative proceeding), *with* Niagara Mohawk Power Corp. v. Consol. Rail Corp., 436 F. Supp. 2d 398 at 402-04 (N.D.N.Y. 2006) (dismissing section 107(a) claim because costs were incurred pursuant to consent orders, and dismissing section 113(f) claim because consent orders did not constitute settlements triggering contribution rights under section 113(f)(3)(B)).

101. 459 F.3d 827 (8th Cir. 2006).

102. 340 F.3d 525 (8th Cir. 2003); *see also supra* note 72.

103. *Atl. Research*, 459 F.3d at 830.

104. *Id.*

105. *Id.* at 836-37. The court concluded that the phrase "any other person" in section 107(a)(4)(B) meant any person other than the statutorily enumerated United States, States, or Indian tribes referenced in section 107(a)(4)(A), and that pre-*Aviall* restrictions of section 107(a)(4)(B) to "innocent" plaintiffs represented nothing more than judicial "traffic directing" light of the pre-*Aviall* analysis that "§ 113 was presumed to be available to all liable parties, including those which had not faced a CERCLA action." *Id.* at 832. The court noted that section 107(a)(4)(B) does not compel full recovery of response costs incurred by a PRP plaintiff, observing that "CERCLA, itself, checks overreaching liable parties: If a plaintiff attempted to use § 107 to recover more than its fair share of reimbursement, a defendant would be free to counterclaim for contribution under § 113(f)." *Id.* at 835. The court also held that "a right to contribution may be fairly implied from the text of section 107(a)(4)(B)," *id.* at 835, stating that "[c]ontribution is crucial to CERCLA's regulatory scheme." *Id.* at 836. In light of these holdings, the court did not address Atlantic's argument that it had an implied right to contribution as a matter of federal common law. *Id.* at 836 n.9.

106. 460 F.3d 515 (3d Cir. 2006); *see also supra* note 93 and accompanying text.

owned and allegedly contaminated by the United States.[107] DuPont brought a CERCLA response cost action against the United States.[108] The Third Circuit affirmed by a 2-1 vote the district court's order granting the United States' motion for judgment on the pleadings on DuPont's CERCLA claims.[109] The Third Circuit concluded that *Aviall* did not give it cause to reconsider its pre-*Aviall* precedents holding that a PRP could only seek contribution under section 113(f),[110] reasoning further that CERCLA's settlement scheme was inconsistent with an interpretation of section 107(a)(4)(B) that would permit direct cost recovery actions by PRPs.[111] Accordingly, the Third Circuit held that DuPont, as a PRP, could not bring a section 107(a) cost recovery action and that, because DuPont had neither settled its CERCLA liabilities with the government nor been sued under CERCLA, DuPont also could not bring a section 113(f) contribution claim in light of *Aviall*.[112]

Similarly, both the Fifth and the Tenth Circuits indicated in post-*Aviall* decisions that they would not permit a PRP to bring a section 107(a) claim. The Fifth Circuit in *Elementis Chromium L.P. v. Coastal States Petroleum Co.*[113] baldly stated in *dicta* that a PRP could not bring a section 107(a) claim.[114] In *Elementis*, the court held that a district court erred by imposing joint and several contribution liability on third-party defendants in a CERCLA contribution action.[115] In reaching this decision, the court cited a pre-*Aviall* Eleventh Circuit decision[116] for the proposition that "when one liable party sues another liable party under CERCLA, the action is not a cost recovery action under § 107(a), and the imposition of joint and several liability is inappropriate."[117]

---

107. *DuPont*, 460 F.3d at 525.

108. *Id.*

109. *Id.* at 543-44.

110. In dissent, Judge Sloviter contended that *Aviall* "clearly undermined" the Third Circuit's pre-*Aviall* precendents barring section 107(a) claims by PRPs because those decisions assumed that all PRPs could assert a section 113(f) contribution action. *Id.* at 546-47. Noting that "[t]here is nothing in the relevant language of § 107 that compels the result the majority reaches," *id.* at 546, the dissent concluded that "permitting parties who voluntarily incur cleanup costs to bring suit under § 107 comports with the fundamental purposes of CERCLA," *id* at 548, because "[v]oluntary cleanups are vital to fulfilling CERCLA's purpose." *Id.* at 549.

111. *Id.* at 541.

112. *Id.* at 543-44.

113. 450 F.3d 607 (5th Cir. 2006).

114. *Elementis*, 450 F.3d at 612-13.

115. *Id.* at 613.

116. Redwing Carriers, Inc. v. Saraland Apartments, 94 F.3d 1489 (11th Cir. 1996).

117. *Elementis*, 450 F.3d at 613 quoting *Redwing Carriers*, 94 F.3d at 1513). *Elementis* did not decide whether a PRP could sue for cost recovery under section 107(a), nor did it even reference *Aviall*. *See Atl. Research*, 459 F.3d at 834 n.7 (dismissing *Elementis* language as an "isolated quotation" regarding an issue that the Fifth Circuit was not asked to decide); *Aviall*, 2006 U.S. Dist. LEXIS 55040, at *14 (granting Cooper's motion for summary judgment following remand from U.S. Supreme Court, holding that Aviall as PRP could not maintain a section 107(a)(4)(B) claim, but observing that *Elementis* "did not squarely decide whether a private PRP can bring a cost recovery action against another PRP under § 107(a)"). Nevertheless, this passing sentence in the *Elementis* opinion underscored post-*Aviall* uncertainty about PRP cleanup cost rights. For example, the Third Circuit in *DuPont*, 460 F.3d at 542 n.32, in the course of holding that a PRP may not bring a section

In *Young v. United States*,[118] the Tenth Circuit affirmed a district court order granting summary judgment in favor of the defendant on the plaintiffs' section 107(a) claim. The district court reasoned that the plaintiffs could not maintain a section 107(a) claim because they were PRPs.[119] The Tenth Circuit affirmed, concluding that plaintiffs' claimed response costs were unnecessary and not consistent with the NCP.[120] As a result, the court observed that it did not need to "determine whether Plaintiffs are PRPs under section 107(a) and thus unable to assert a cost-recovery claim under the rule in this Circuit that a Plaintiff-PRP must proceed under the contribution provisions of CERCLA section 113(f) when the Plaintiff-PRP sues another PRP for response costs."[121] In short, after *Aviall* the availability of PRP cleanup cost rights under CERCLA likely will remain uncertain until the Supreme Court or Congress resolves the issue.

2. Impact on Voluntary Cleanups

The uncertainty *Aviall* created presented a series of stark choices to owners of contaminated property (as well as other PRPs) faced with agency remediation directives.[122] The landowner could cooperate by voluntarily (i.e., short of judicial enforcement action) complying with a regulatory cleanup directive. A landowner conducting a voluntary cleanup, however, risks incurring cleanup costs without a well-settled

---

107(a) action, cited *Elementis* for the proposition that "at least one other Circuit Court has agreed with our interpretation of § 107(a) . . . ." The district court in Columbus McKinnon Corp. v. Gaffey, No. H-06-1125, 2006 WL 2382463, at *4 (S.D. Tex. Aug. 16, 2006), granted defendant's motion to dismiss a PRP's section 107(a) claim, relying in part on the "May 2006 statement of the Fifth Circuit in *Elementis*" to conclude that "under the prevailing law at this time [plaintiff] as a PRP does not have a viable claim for cost recovery under § 107(a) of CERCLA." Defendants in other cases also pointed to *Elementis* as evidence of a circuit split regarding the availability of a section 107(a) claim for a PRP plaintiff. For example, *Elementis* was cited by the United States in *Atl. Research*, 459 F.3d at 834 n.7 as support for the proposition that a PRP may not maintain a section 107(a) action, and by UGI Industries, Inc. (the defendant in *Consolidated Edison*) in connection with its petition to the U.S. Supreme Court for a writ of certiorari as evidence of a circuit split requiring the Court to decide whether a PRP could assert a section 107(a) claim. Reply Brief of Petitioner UGI Utilities, Inc. on Petition for Writ of Certiorari to the United States Court of Appeals for the Second Circuit at 4, UTI Utilities, Inc. v. Consol. Edison Co. of New York, Inc., No. 05-1323, 2006 WL 1621796, at *4 (U.S. June 12, 2006) ("The Fifth Circuit now also has ruled that, contrary to the decision below, PRPs have no §107(a) cost recovery claim, but PRPs continue to litigate that issue across the Nation."); *see also* David Ledbetter, Kathy Robb, Andrew Skroback, Cooper Industries v. Aviall*: The Aftermath*, 26 ANDREWS ENVTL. LITIG. REP. 5 n.9 (July 14, 2006) ("Elementis Chromium cannot be reasonably read as foreshadowing of how the 5th Circuit will rule if it revisits the Section 107 issue. Frenzied parsing and speculation concerning this language say much, however, about the huge stakes in play concerning the Section 107 issue.").

118. 394 F.3d 858 (10th Cir. 2005).

119. *Young*, 394 F.3d at 860.

120. *Id.* at 863.

121. *Id.* at 862. *See* Raytheon Aircraft Co. v. United States, 435 F. Supp.2d 1136, 1145 (D. Kan. 2006) (quoting the statement from *Young* that "a PRP is unable to assert a cost recovery claim "); *Columbus McKinnon*, 2006 WL 2382463, at *4 (citing *Young* for the proposition that the Tenth Circuit after *Aviall* "implicitly recognized that the law in that circuit remained that a PRP could not maintain a § 107 claim").

122. *See* Aronovsky, *supra* note 5, at 50-57.

CERCLA right to recover from other PRPs their fair share of cleanup costs.[123] In the alternative, the landowner could refuse to comply with agency directives and instead require the agency to initiate a judicial enforcement action.[124] Such a strategy could result in a CERCLA lawsuit that would trigger section 113(f)(1) contribution rights. On the other hand, this strategy could also reduce the landowner's negotiation power with the agency, generate substantial litigation costs, expose the landowner to potentially severe penalties for non-compliance with agency directives,[125] and delay site remediation. A landowner could attempt to settle with the EPA or a state agency, giving rise to CERCLA section 113(f)(3) contribution rights;[126] however, neither the EPA nor state regulatory agencies with the power to order cleanups have the resources to negotiate settlements at tens (if not hundreds) of thousands of contaminated sites.[127]

3. *Aviall* and the Role of State Law Cleanup Cost Claims

After *Aviall*, most owners of contaminated property could no longer count on a federal cleanup cost remedy under CERCLA. Because current owners are liable parties under CERCLA, they may not, after *Aviall*, obtain cleanup cost contribution under section 113(f) unless they first have been sued under CERCLA or settled their CERCLA liabilities with the government, and may not have a right to cost recovery under section 107(a).[128] *Aviall*, of course, addressed only federal law claims. In the-

123. Commentators expressed concern that *Aviall* would have a significant chilling effect on PRPs cooperating with regulatory agencies in the absence of a CERCLA contribution remedy. *See, e.g.*, Charles F. Helsten et al., *The Effect of* Aviall *on the Vitality of Brownfields Programs*, A.B.A. ENVTL. TRANSACTIONS & BROWNFIELDS COMM. NEWSL., Mar. 2005, at 4 (*Aviall* decision "has raised many questions that may make developers leery of brownfields projects" in light of uncertain cleanup cost contribution rights); Richard G. Leland & Toni L. Finger, *The Supreme Court's Limitation on Private Cost Recovery Actions Under Superfund: No Good Deed Goes Unpunished - Part II*, METRO. CORP. COUNS., May 2005, at 8 (parties may choose not to enter into voluntary cleanup agreements, particularly in states without state Superfund statute contribution rights); *see also* NAT'L GOVERNORS ASS'N, POLICY POSITION STATEMENT NR-04: SUPERFUND POLICY § 4.4 (2000) (revised Winter Meeting 2005) ("[A] recent U.S. Supreme Court decision [*Aviall*] has the potential to diminish a significant incentive under CERCLA for responsible parties to properly perform voluntary cleanups under state oversight.").

124. *See, e.g.*, Daniel M. Steinway, *The Ramifications of the Aviall Decision: Where Do We Go From Here?*, 20 TOXICS L. REP. (BNA) 190, 194–95 (2005) (potential techniques for a PRP to obtain CERCLA contribution rights include inviting a "friendly" lawsuit from a regulatory agency).

125. *See, e.g.*, 42 U.S.C.A. § 9606(b)(1) (West 2006) (imposing $25,000 per day fines for non-compliance absent "sufficient cause"); 42 U.S.C.A. § 9607(c)(3) (authorizing EPA to seek punitive damages of up to three times the response costs incurred as a result of failure to take proper action); *see also supra* note 27 and accompanying text.

126. *See supra* note 67 and accompanying text.

127. *See* Schnapf, *supra* note 65, at 612 ("[P]arties may have to offer some sort of 'carrots' to state agencies to justify diverting limited resources by performing a more comprehensive cleanup than normally would be required or perhaps implementing a supplemental environmental project."); Albert M. Cohen, *Certainty and Uncertainty in the Post* Cooper v. Aviall *Superfund World*, 20 TOXICS L. REP. (BNA) 73, 77 n.15 (2005) ("EPA could be cooperative and enter into such agreements. On the other hand, it may see the lack of a right to contribution as a means to put additional pressure on parties which refuse to settle.").

128. Cooper Indus., Inc. v. Aviall Servs., Inc., 543 U.S. 157, 167 (2004).

ory, landowners could always look to state law as an alternative to a CERCLA claim. Indeed, PRPs often had joined state law claims to CERCLA cleanup cost claims, for a variety of reasons. First, CERCLA does not provide a petroleum cleanup cost recovery remedy;[129] state law is not so limited. Second, CERCLA does not provide damages for property or personal injury caused by contamination;[130] state law could. Third, CERCLA limits private cleanup cost recovery to "necessary" response costs that are "consistent with the national contingency plan;"[131] state common law theories are free of the complicated, time-consuming and often expensive requirements of the NCP.

So why does *Aviall* matter? It matters because state law often provides no remedy at all, especially for current landowners faced with contamination left by their predecessors or contamination that occurred years ago. The hundreds of thousands of contaminated sites across the country represent a significant national problem requiring a comprehensive solution. Voluntary cleanups are essential to solving this problem, and contribution rights greatly facilitate voluntary cleanups. A federal cleanup cost contribution remedy under CERCLA would ensure the availability of contribution claims for sites in every state. But a federal cleanup cost remedy is not necessarily preferable to a state law remedy. To the contrary, state law remedies, in many ways, could be superior to CERCLA. A federal "safety net" remedy is necessary, however, because current state law fails to provide a coherent nationwide response to a nationwide problem.[132]

II. AN INCOHERENT PATCHWORK QUILT: THE CURRENT STATE OF PRIVATE NUISANCE LAW AS APPLIED TO PRIVATE CLEANUP COST DISPUTES

This section analyzes the unrealized potential of private nuisance as a meaningful rule of decision in private cleanup cost disputes. It first looks at the efficiencies and remedial flexibility that could be realized through application of private nuisance law to contaminated property disputes. This section then analyzes the doctrinal and policy limitations of current private nuisance law that dramatically limit opportunities for it to provide these benefits. Part IIII of the article proposes an environmental nuisance paradigm for soil and groundwater contamination disputes that would eliminate shortcomings in current law and make reliance on state law a meaningful alternative to a uniform private federal remedy.

---

129. 42 U.S.C.A. § 9601(14) (West 2006).

130. *See, e.g.*, Artesian Water Co. v. Gov't of New Castle County, 659 F.Supp. 1269, 1285 (D. Del. 1987) ("Congress in enacting CERCLA clearly manifested an intent not to provide compensation for economic losses or for personal injury resulting from the release of hazardous substances.").

131. 42 U.S.C.A. § 9607(a)(4)(B); *see also supra* note 54 and accompanying text.

132. *See* Aronovsky, *supra* note 5, at 68-79 (arguing that a federal "safety net" remedy is necessary in light of the inter-state inconsistencies and doctrinal limitations of current state law).

A. Flexibility and Efficiency: The Potential Benefits of State Law as the Primary Rule of Decision in Private Cleanup Cost Disputes

1. Efficiency

State law could provide a more attractive source for the rules of decision governing private cleanup cost disputes than federal law.[133] Use of state law to resolve all contamination dispute issues would promote litigation efficiency. First, state law contamination disputes would be heard in state courts[134] before state court judges more likely to be familiar with applicable state law than federal court judges.[135] Second, using state law as the primary rule of decision in private cleanup cost disputes would permit the application of a common body of state law to all contaminated property dispute questions, including claims for petroleum cleanup costs,[136] property damages (beyond cleanup costs), and personal injury damages that are not recoverable under CERCLA. Third, applying a common body of state law to all cleanup cost disputes would avoid conflicts between federal and state law regarding alternative cleanup cost allocations among PRPs, and differing measurements of recoverable damages between costs that were and were not consistent with the NCP.[137] Fourth, a common body of law in multi-party contamination disputes would promote settlements by avoiding potential complications that could arise from inconsistent rules governing settlements not involving all disputants.[138] Fifth, in light of the fact that state or local govern-

---

133. *See id.* at 60-68 (discussing potential flexibility and efficiency benefits of state law as primary rule of decision in private cleanup cost disputes).

134. By contrast, CERCLA section 113(b) provides that federal courts have exclusive subject matter jurisdiction over claims "arising under" CERCLA. 42 U.S.C.A. § 9613(b). State law claims can be joined to CERCLA claims pursuant to a federal court's original (i.e., diversity) or supplemental subject matter jurisdiction. *See, e.g.*, Arawana Mills Co. v. United Techs. Corp., 795 F. Supp. 1238, 1247-49 (D. Conn. 1992) (analyzing under 28 U.S.C.A. § 1367(c) retention of supplemental jurisdiction over state law claims joined with CERCLA claims).

135. A private cleanup cost or other contamination-related dispute based solely on state law always could, of course, be filed in or removed to federal court based on diversity of citizenship jurisdiction. 28 U.S.C.A. §§ 1332, 1441 (West 2006).

136. CERCLA excludes petroleum from the definition of the "hazardous substances" governed by CERCLA. 42 U.S.C.A. § 9601(14) (West 2006).

137. *See* Aronovsky, *supra* note 5, at 89 n.369.

138. Partial settlements (i.e., settlements involving fewer than all defendants) in multi-defendant cleanup cost litigation typically are conditioned on issuance by the court of a "contribution bar order" providing contribution protection for a settling defendant. A contribution bar order prohibits contribution claims against the settling defendant by reducing any judgment against the non-settling defendant in an amount corresponding to the settling defendant's fair share of liability. This reduction in judgment can either be *pro tanto* (reducing the non-settling defendant's liability by the amount paid in settlement), ordered by the court after a "fairness" hearing on the settlement, or *pro rata,* reducing the non-settling defendant's liability by the settling defendant's equitable share of liability as determined at trial. Federal courts are split as to whether to give *pro tanto* or *pro rata* contribution protection in private cleanup cost cases. *See, e.g.*, Eric DeGroff, *Raiders of the Lost ARCO: Resolving the Partial Settlement Credit Issue in Private Cost Recovery and Contribution Claims Under CERCLA*, 8 N.Y.U. ENVTL. L.J. 332, 397 (2000) (arguing for *pro rata* protection in CERCLA contribution claims and *pro tanto* protection in cost recovery claims). Complications can arise when the applicable federal claim contribution protection rule differs from the state law rule. Using a common body of state law in private cleanup cost disputes would eliminate this problem.

ment agencies take the regulatory lead at most sites, the availability of state law claims would facilitate coordination of cost claim elements and regulatory agency remediation requirements. Finally, tort-based state law cleanup cost claims could be asserted against government entity PRPs because federal and state tort claims statutes authorize private tort-based cleanup cost claims.[139]

2. Flexibility

Using state law as the primary rule of decision in private cleanup cost disputes would also promote flexibility in addressing the nation's many contaminated sites. For example, CERCLA limits private cleanup cost recovery to only those response costs consistent with the NCP.[140] Compliance with NCP requirements[141] for conducting an appropriate investigation into the extent of the pollution and evaluating the feasibility of remedial alternatives,[142] as well as affording a meaningful opportunity

139. The federal government has been identified as a PRP at thousands of sites across the country, potentially involving billions of dollars in cleanup costs. Several *amicus curiae* briefs filed with the U.S. Supreme Court in *Aviall* argued that a PRP right of contribution under CERCLA was necessary to promote cost recovery from the federal government at sites where the United States was a PRP. *See* Aronovsky, *supra* note 5, at 44 n.212. State tort law claims may be asserted against the federal government pursuant to the Federal Tort Claims Act (FTCA). 28 U.S.C.A. §§ 2671-2680 (West 2006); *see* Hoery v. United States, 324 F.3d 1220, 1224 (10th Cir. 2003) (holding landowner timely filed FTCA claim for continuing trespass and nuisance under Colorado tort law arising from groundwater contamination); *see also* Elizabeth Ann Coleman, *In Re* Hoery v. United States*: Compensating Homeowners For Loss of Property Value Due to Toxic Pollution Under the Continuing Tort Doctrine*, 16 VILL. ENVTL. L.J. 35, 44-47 (2005) (discussing application of FTCA to continuing tort claim relating to groundwater contamination). Similarly, state law claims may be asserted against state governments pursuant to comparable state tort claim act statutes or other state statutes waiving sovereign immunity. *See, e.g.*, Ayers v. Jackson Tp., 525 A.2d 287, 291-92 (N.J. 1987) (affirming in part nuisance personal injury damage award in action brought by local residents against municipality regarding contaminants leaching from town landfill into aquifer). By contrast, the Eleventh Amendment and state sovereign immunity protect state government entities from CERCLA claims which can only be brought in federal court because CERCLA section 113(b) provides exclusive federal court jurisdiction for claims "arising under" CERCLA. 42 U.S.C.A. § 9613(b) (West 2006); *see, e.g.*, Burnette v. Carothers, 192 F.3d 52, 60 (2d Cir. 1999) (holding private CERCLA action against state barred by the Eleventh Amendment).

140. A private party response action is consistent with the NCP if it is in "substantial" compliance with applicable requirements of the NCP. 40 C.F.R. § 300.700(c)(3)(i) (2005).

141. Private cleanup actions that result in a "CERCLA-quality cleanup" are considered consistent with the NCP. *See* Regional Airport Authority of Louisville v. LFG, LLC, 460 F.3d 697, 707 (6th Cir. 2006) (holding cleanup is consistent with NCP if, taken as a whole, it is in substantial compliance with NCP and results in a CERCLA-quality cleanup). For a private response action to constitute a "CERCLA-quality cleanup," the selected remedy must: (1) protect human health and the environment; (2) employ permanent solutions and alternative treatment technologies to the maximum extent practicable; (3) be cost effective; (4) identify and attain applicable or relevant and appropriate requirements (ARARs, i.e., cleanup standards) for the site; and (5) provide for meaningful public participation in the remedy selection process. *See* National Oil and Hazardous Substances Pollution Contingency Plan, 55 Fed. Reg. 8666-01, 8793 (Mar. 8, 1990); Franklin County Convention Facilities Auth. v. Am. Premier Underwriters, Inc., 240 F.3d 534, 543 (6th Cir. 2001) (identifying "CERCLA-quality cleanup" NCP factors).

142. The NCP requires preparation of a remedial investigation and feasibility study (RI/FS), which involves, *inter alia*, a thorough analysis of contamination conditions and remedial alternatives. 40 C.F.R. § 300.430(e)(9) (2005). *See* Carson Harbor Vill., Ltd. v. County of Los Angeles, 433 F.3d 1260, 1268–69 (9th Cir. 2006) (holding feasibility study fully analyzing only one remedial

for public participation in the remedial process,[143] can be time consuming and expensive.[144] The additional costs and delays required for NCP consistency (and for CERCLA cost recovery from polluters) can discourage prospective purchasers from acquiring and developing brownfield or other properties contaminated by others.[145]

State common law theories, while imposing culpability and causation proof requirements inapplicable to CERCLA claims, do not condition cleanup cost recovery on consistency with the NCP.[146] State law, therefore, could provide an attractive cost claim alternative for current landowners hoping to conduct a technically sound cleanup without the delay and expense often resulting from NCP requirements.[147] Similarly,

---

alternative did not substantially comply with NCP). The final remedy selected by the lead regulatory agency is documented in a Record of Decision (ROD). 40 C.F.R. § 300.430(f)(1).

143. Public participation requirements include creation of a public information and community relations plan and providing sufficient opportunity for public comment on the RI/FS and a proposed remedy. 40 C.F.R. § 300.700(c)(6). Failure to comply with NCP public participation requirements may bar cost recovery. *See, e.g.*, *Regional Airport Authority*, 460 F.3d at 703-08 (affirming summary judgment for prior owner because plaintiff current landowner's claimed costs were not "necessary" within the meaning of section 107(a)(4)(B) and were not consistent with the NCP because, *inter alia*, plaintiff failed to permit public comment on the selected remedy); *Carson Harbor Vill.*, 433 F.3d at 1266–67 (holding plaintiff failed to show compliance with public participation requirement; "minor and ministerial" involvement of public agency did not provide effective substitute for public participation); Union Pac. R.R. Co. v. Reilly Indus., 215 F.3d 830, 835 (8th Cir. 2000) ("Failure to provide a meaningful opportunity for public participation and comment in the selection of a remedial action at a particular cleanup site is inconsistent with the NCP.").

144. *See, e.g.*, Richard G. Opper, *The Brownfield Manifesto*, 37 URB. LAW. 163, 182-83 (2005) (noting that NCP requirements apply on their face to all sites regardless of their complexity, and arguing for amending the NCP to streamline brownfield and other low-risk site regulation); *see also* WILLIAM H. RODGERS, JR., 4 RODGERS' ENVIRONMENTAL LAW § 8:9 (2006 Update) ("It deserves emphasis that the remedial implementation process for NPL sites can be slow (10 to 12 years from initiation of the RI to site cleanup), ponderous (an average of 8 years of study before cleanup begins), expensive (RI/FS cost $750,000 on average with a high of $7 million; remedies average $25 million, several have exceeded $100 million, and a few are approaching $ 1 billion), legally profuse (139 RODs and 4 ROD amendments were signed during Fiscal Year 1989), labor consumptive (some sites will go through 4 or 5 Remedial Project Managers in the years of implementation), and unfailingly complex (single sites can generate multiple RI/FSs and multiple RODs (as many as twenty) for each piece of the cleanup pie." (footnotes omitted)).

145. *See* Opper, *supra* note 144, at 183 ("The NCP needs to be updated to fit the new brownfields paradigm, or else it should adopt language to allow a finding of 'consistency' after little or no elaborate process for certain types of common urban projects."); Schwenke, *supra* note 24, at 297 ("The potential liability for environmental contamination continues to stand as a major impediment to acquisition, financing, and development of these vacant or abandoned sites.").

146. In *PMC, Inc. v. Sherwin-Williams Co.*, 151 F.3d 610, 617–18 (7th Cir. 1998), the Seventh Circuit held that a plaintiff unable to recover response costs under CERCLA section 113(f)(1) because of inconsistency with the NCP could not recover those costs derivatively under the Illinois Contribution Act, reasoning that plaintiff PMC's attempted "invocation of Illinois' contribution statute is an attempt to nullify the sanction that Congress imposed for the kind of CERCLA violation that PMC committed." *See* Aronovsky, *supra* note 5, at 93-99 (discussing *PMC* decision and arguing that the NCP should not provide basis for CERCLA to preempt state statutory or common law direct cleanup cost remedies).

147. *See, e.g.*, Richard G. Opper, *Managing Risks at Brownfield Sites*, 20 NAT. RESOURCES & ENV'T 32, 36 (2006). Mr. Opper observed that:

> There is sometimes less risk, and a greater chance of success, for private cost-recovery plaintiffs in a common law nuisance case than there is in using CERCLA. The use of a continuing nuisance theory, incepting after mitigation measures are complete, can be efficient and effective in state court. It can be a cheaper and faster path through the litiga-

without the specter of NCP consistency hovering over current landowners looking to recover the costs of voluntary cleanup from other PRPs, states could have greater flexibility with which to experiment and promote effective alternative site cleanup policies.[148]

B. Theory Meets Reality: Current Doctrinal Limitations Prevent State Law from Solving the National Problem of Encouraging Voluntary Cleanups

In theory, state law remedies could encourage the voluntary cleanup of contaminated sites as well as or better than a PRP cleanup cost remedy under CERCLA. In fact, current state law is not up to the task.[149] The doctrinal limitations of current state statutory and common law theories, and the significant variations among state law tort regimes, make meaningful state law cleanup cost remedies unavailable at many multi-PRP sites across the country.

Nuisance casts the widest liability and remedial net among state law claims potentially applicable to private cleanup cost disputes[150] because it is not subject to limitations that constrict other state law theories. For example, a trespass claim requires proof of intentional conduct and can be defeated by a possessor consent defense.[151] Moreover, because the tort of trespass is based on the unauthorized invasion of the possessory interest of another in real property, trespass cannot be asserted by a cur-

---

tion than federal court, and it does not require compliance with the National Contingency Plan, providing for cost savings during the cleanup.

Id.

148. *See* Tom Kuhnle, *The Rebirth of Common Law Actions for Addressing Hazardous Waste Contamination*, 15 STAN. ENVTL. L.J. 187, 221 (1996) ("[s]tate and local laws are better than federal laws at reflecting local preferences for environmental quality and at encouraging valuable environmental policy experimentation" (citing Richard B. Stewart, *Pyramids of Sacrifice? Problems of Federalism in Mandating State Implementation of National Environmental Policy*, 86 YALE L. J. 1196, 1210 (1977) ("[s]tate and local governments can better reflect geographical variations in preferences for collective goals like environmental quality."))).

149. *See* Aronovsky, *supra* note 5, at 68-79.

150. *See supra* notes 30-34 and accompanying text; *infra* notes 151-61 and accompanying text; *see also* WARREN FREEDMAN, HAZARDOUS WASTE LIABILITY 121 (1992) ("[N]uisance is a more reliable theory of liability than trespass or negligence for the hazardous waste disposal situation."); Klein, *supra* note 12, at 353-54 (noting that "[i]n light of CERCLA's failures, legal commentators have increasingly suggested that courts supplement or replace the legislative regime through an expanded use of common law tort actions—in particular, nuisance law," and citing the observation of RODGERS, *supra* note 144, at 112–13, that "[t]here is no common law doctrine that approaches nuisance in comprehensiveness or detail as a regulator of land use and of technological abuse."); Roger Meiners & Bruce Yandle, *Common Law and the Conceit of Modern Environmental Policy*, 7 GEO. MASON L. REV. 923, 926 (1999) ("The nuisance cause of action provides the backbone of common law environmental (pollution) litigation."). More than one common law theory, of course, may apply to a private cleanup cost dispute. *See, e.g.*, Joseph F. Falcone III & Daniel Utain, *You Can Teach an Old Dog New Tricks: The Application of Common Law In Present-Day Environmental Disputes*, 11 VILL. ENVTL. L.J. 59, 88 (2000) (noting that "cases have recognized that invasions of property that amount to a trespass may also constitute a nuisance").

151. *See, e.g.*, RESTATEMENT (SECOND) OF TORTS § 158 cmt. e (1965) (intentional conduct is required for trespass, but intrusion with consent of a possessor is privileged); Edgcomb v. Lower Valley Power and Light, Inc., 922 P.2d 850, 859 (Wyo. 1996) (consent of possessor is an absolute defense to a trespass action under Wyoming law).

rent landowner seeking to recover cleanup costs from a prior owner or tenant who disposed of contaminants while possessing the property.[152] Negligence claims require proof that the defendant's conduct did not conform to an applicable standard of care[153] and may not fall within the continuing tort doctrine in order to avoid a statute of limitations defense.[154] Strict liability for an abnormally dangerous activity can be unpredictable because of the multi-factored balancing analysis required to show that an activity is abnormally dangerous,[155] and in some states this doctrine is inapplicable to contaminated property disputes.[156] None of these limitations generally apply to a nuisance claim.

---

152. *See, e.g.*, Wellesley Hills Realty Trust v. Mobil Oil Corp. 747 F. Supp. 93, 99 (D. Mass. 1990) ("Mobil owned and was in possession of the property when it allegedly released the oil causing the contamination. Thus, Mobil's releases of oil were not unprivileged, and Mobil clearly was not intruding on land in the possession of another. Mobil's releases of oil on its own land, therefore, cannot constitute a trespass."); Capogeannis v. Superior Court, 12 Cal. App. 4th 668, 674 (Cal. Ct. Ap. 1993) ("Manifestly one cannot commit an actionable interference with one's own possessory right."); RESTATEMENT (SECOND) OF TORTS § 821D cmt. d (1965) ("A trespass is an invasion of the interest in the exclusive possession of land, as by entry upon it."); *see also* Falcone & Utain, *supra* note 150, at 100 ("While a small minority of jurisdictions are willing to permit current landowners to sue previous owners for contamination under claims of nuisance and strict liability, there are an even smaller number of jurisdictions willing to allow them to sue under a claim for trespass under similar conditions.").

153. *See, e.g.*, RESTATEMENT (SECOND) OF TORTS § 282 (1965) ("[N]egligence is conduct which falls below the standard established by law for the protection of others against unreasonable risk of harm."). Contamination caused by waste disposal practices that at the time reflected state of the art technology may not be actionable years late under a negligence theory. *See* Aronovsky, *supra* note 5, at 70 n.304.

154. *See, e.g.*, Andritz Sprout-Bauer, Inc. v. Beazer E., Inc., 12 F. Supp. 2d 391, 417 (M.D. Pa. 1998) (finding current landowner's negligence claim against former owner and operator of site time-barred; continuing tort doctrine applies only to trespass and nuisance, not negligence); Church v. General Elec. Co. 138 F. Supp. 2d 169, 174 (D. Mass. 2001) (under Massachusetts law, continuing tort doctrine is limited to nuisance and trespass claims and does not apply to negligence or strict liability claims). *But see* Nat'l. Tel. Co-op. Ass'n. v. Exxon Corp., 38 F. Supp. 2d 1, 6-7 (D.D.C. 1998) (holding under District of Columbia law that continuing tort theory applied to toll five year limitations period on commercial property owner's negligence and strict liability for ultra-hazardous activity claims alleging that gasoline leaking from a neighbor's underground storage tanks). Application of the discovery rule, coupled by a tolling of the statute of limitations following discovery of contamination during site evaluation could reduce the risk that a negligence-based cleanup cost claim could become time barred. *See* Melanie R. Kay, *Environmental Negligence: A Proposal for a New Cause of Action for the Forgotten Innocent Owners of Contaminated Land*, 94 CAL. L. REV. 149, 172 (2006) (proposing environmental negligence cause of action that would toll statute of limitations during pendency of site investigation).

155. *See* RESTATEMENT (SECOND) OF TORTS § 520 (1977); *supra* note 37 (identifying the six factors to consider in determining whether an activity is abnormally dangerous); *see, e.g.*, Nnadili v. Chevron U.S.A., Inc., 435 F. Supp. 2d 93, 103 (D.D.C. 2006) (applying District of Columbia law and section 520 balancing factors to grant defendant's motion for summary judgment on the ground that storage of gasoline in defendant's underground storage tanks did not constitute an abnormally dangerous activity); *cf.* Klass, *supra* note 34, at 963 (noting the continued problem of unpredictability in applying section 520 multi-factor analysis).

156. *See, e.g.*, Nat'l Tel. Coop. Ass'n v. Exxon Corp., 38 F. Supp. 2d 1, 8 (D.D.C. 1998) (observing that strict liability for abnormally dangerous activity had not yet been explicitly adopted in the District of Columbia); Jones v. Texaco, Inc., 945 F. Supp. 1037, 1050 (S.D. Tex. 1996) (noting Texas courts have rejected doctrine of abnormally dangerous activities as a basis for strict liability in the context of hazardous wastes); Arawana Mills Co. v. United Techs. Corp., 795 F. Supp. 1238, 1251 (D. Conn. 1992) (concluding Connecticut law would not recognize storage and use of hazardous materials as abnormally dangerous).

Nuisance law similarly avoids many of the limitations of "state Superfund" statutes. For example, some of these statutes limit private cost recovery to costs consistent with the NCP.[157] Some state Superfund statutes apply only to releases occurring after the statute was enacted.[158] Like CERCLA,[159] some state Superfund statutes do not apply to petroleum.[160] Finally, like CERCLA section 113(f), some State Superfund statutes only permit contribution claims during or after judicial or administrative proceedings against a PRP.[161]

As discussed below, nuisance law has great promise as a rule of decision in private cleanup cost disputes because its flexibility corresponds well to the complexity of soil and groundwater contamination problems.[162] Moreover, nuisance law avoids doctrinal limitations presented

---

157. *See, e.g.*, KY. REV. STAT. ANN. § 224.01-400 (West 2006); N.J. STAT. ANN. § 58:10-23.11f3 (West 2006). Some state Superfund statutes do not require NCP consistency for cost recovery. *See, e.g.*, ARIZ. REV. STAT. ANN. § 49-285(B) (West 2006); DEL. CODE ANN. tit. 7, § 9105 (West 2006); FLA. STAT. § 403.727 (West 2006); GA. CODE ANN. § 12-8-96.1 (West 2006); MICH. COMP. LAWS ANN. § 324.20126 (West 2006). Because state Superfund statutes often largely mirror the structure and liability scheme of CERCLA, *see infra* notes 159-61, state Superfund statutes permitting contribution for cleanup costs that are not consistent with the NCP could be subject (properly or otherwise) to a preemption challenge on the ground that using a "state version of CERCLA" to permit recovery of non-NCP costs would undermine the goals of CERCLA. Such an overbroad preemption analysis would be ill-considered. *See* Aronovsky, *supra* note 5, at 90-98.

158. *See, e.g.*, CAL. HEALTH & SAFETY CODE § 25366(a) (Deering 2005) (no liability for actions before January 1, 1982 if the actions were not in violation of then-existing state or federal law); HAW. REV. STAT. § 128D-6 (2005) (no private recovery of costs arising from a release occurring before July 1, 1990).

159. State Superfund statutes often adopt or incorporate provisions from CERCLA. *See, e.g.*, CAL. HEALTH & SAFETY CODE § 25323.5 (Deering 2005) (incorporating CERCLA definitions of "responsible party" and "liable person"); IND. CODE ANN. § 13-25-4-8 (West 2006) (incorporating CERCLA definition of liable parties); UTAH CODE ANN. § 19-6-302(20) (West 2006) ("remedial investigation" means a remedial investigation and feasibility study as defined in the NCP).

160. *See, e.g.*, ALASKA STAT. § 46.09.900 (2005); CAL. HEALTH & SAFETY CODE § 25317 (Deering 2005); KY. REV. STAT. ANN. § 224.01-400(1)(a) (West 2006); VA. CODE ANN. § 10.1-1400 (West 2006).

161. *See, e.g.*, Wacker Chem. Corp. v. Bayer Cropscience, Inc., No. 05-72207, 2006 WL 2404502, at *5 (E.D. Mich. Aug. 18, 2006) (relying on *Aviall* to interpret Mich. Comp. Laws § 324.20129(3) to permit contribution actions only during or after civil actions brought under Michigan state Superfund statute); 35 PA. STAT. ANN. § 6020.705(a) (West 2006). Amending state Superfund statutes to ensure cleanup cost contribution rights for all PRPs, including current landowners, could serve the same "safety net" function as confirming a similar right under CERCLA. Attempting to amend a state Superfund statute, however, could be frustrated by the same risk of "opening Pandora's box" and confronting the often competing demands of contaminated property dispute stakeholders that have frustrated efforts to amend core provisions of CERCLA for the past twenty years. *See* Aronovsky, *supra* note 5, at 81. In any event, state Superfund statute cleanup cost contribution rights (like CERCLA contribution rights) would not provide the litigation efficiency and remedial flexibility benefits available under the proposed private nuisance paradigm, such as the application of a common body of tort law to both cleanup cost and other claims raised in private contaminated property disputes and the absence of the NCP consistency and petroleum exclusion limitations sometimes found in state Superfund statutory schemes.

162. *See infra* note 175 and accompanying text; *see also* Denise E. Antolini, *Modernizing Public Nuisance: Solving the Paradox of the Special Injury Rule*, 28 ECOLOGY L.Q. 755, 771 (2001) ("Although amorphous in definition, all nuisance actions have in common three important doctrinal aspects that provide unique scope to their application by the courts: substantiality of interference, unreasonableness of the defendant's conduct, and equitable flexibility."); Kuhnle, *supra* note 148, at 221-23 (arguing that nuisance liability is more contextual, less predictable than CERCLA, not limited by legislative line-drawing).

by other state law theories.[163] Nevertheless, the current state of nuisance law fails as a meaningful alternative to a uniform federal rule of decision in private cleanup cost disputes for the promotion of voluntary cleanups. Simply put, nuisance law has not evolved to meet the realities of modern contaminated property problems.[164]

C. *The Inadequate State of Current Nuisance Law*

1. Nuisance—A Brief Overview

Nuisance law dates back to the twelfth century;[165] its well-documented history[166] does not require repetition here. William Prosser famously observed that "[t]here is perhaps no more impenetrable jungle in the entire law than that which surrounds the word 'nuisance.'"[167] As one court observed, "in order to alleviate some of the confusion [surrounding the use of the term "nuisance"], it is important to distinguish 'private' and 'public' nuisance, which 'bear little relationship to each other. Although some rules apply to both, other rules apply to one but not the other."[168]

---

163. *See infra* notes 151-61 and accompanying text; *cf.* Kay, *supra* note 154, at 162-68 (arguing for an "environmental negligence" cause of action because, among other things, doctrinal limitations make current nuisance law unavailable or inadequate to meet common contamination problems); Klein, *supra* note 12, at 339 (arguing that "shifting costs through nuisance law is no more efficient than shifting costs through CERCLA-created liability" and proposing instead a broad-based system of revenue collection to fund cleanups).

164. *See, e.g.*, G. Nelson Smith III, *Nuisance and Trespass Claims in Environmental Litigation: Legislative Inaction and Common Law Confusion*, 36 SANTA CLARA L. REV. 39 (1995). Mr. Smith argued that:

> The use of common law nuisance and trespass claims to address environmental problems is an outdated method since much of the case law relied upon precedes the enactment of many of the environmental statutes. Consequently, the cases do not direct the courts on how to confront the complicated problems that are associated with pollution.

Id. at 70.

165. *See, e.g.*, *id.* at 41 (describing twelfth century origins of nuisance); Falcone & Utain, *supra* note 150, at 65 ("The legal theory of nuisance dates back to the twelfth century.").

166. *See, e.g.*, Antolini, *supra* note 162, at 767-68; Robert G. Bone, *Normative Theory and Legal Doctrine in American Nuisance Law: 1850-1920*, 59 S. CAL. L. REV. 1101, 1139-41 (1986); H. Marlow Green, *Common Law, Property Rights and the Environment: A Comparative Analysis of Historical Developments in the United States and England and a Model for the Future*, 30 CORNELL INT'L L.J. 541, 545-46 (1997); Jeff L. Lewin, *Boomer and the American Law of Nuisance: Past, Present and Future*, 54 ALB. L. REV. 189, 192-96 (1990); William A. McRae, Jr., *The Development of Nuisance in the Early Common Law*, 1 U. FLA. L. REV. 27 (1948); William H. Wilson, Comment, *Nuisance as a Modern Mode of Land Use Control*, 46 WASH. L. REV. 47, 54-55 (1970).

167. WILLIAM L. PROSSER, HANDBOOK OF THE LAW OF TORTS 571 (4th ed. 1971).

168. United States v. Hooker Chems. & Plastics Corp., 722 F. Supp. 960, 964 (W.D.N.Y. 1989) (quoting State of New York v. Shore Realty Corp., 759 F.2d 1032, 1050 (2d Cir.1985)).

Public and private nuisance differ in significant respects.[169] A public nuisance is "an unreasonable interference with a right common to the general public."[170] Government entities may bring an action to enjoin and compel the abatement of a public nuisance.[171] Private parties affected by a public nuisance (e.g., the owner of property impacted by a contaminated groundwater plume) must show an injury different than that suffered by other affected property owners to state a *prima facie* public nuisance claim.[172] Major contaminated groundwater problems affecting an entire neighborhood or public water supply would cause or threaten to cause harm to a common interest of the general public and thus qualify as a public nuisance. Most contaminated property disputes, however, do not involve a public nuisance. For example, contaminated soil problems (usually limited to a single parcel of polluted property) and discrete contaminated groundwater problems (i.e., those involving at most only a handful of properties down-gradient from the source property and not threatening a public water supply) likely would not be considered a public nuisance. The environmental nuisance paradigm proposed by this article focuses on private nuisance contaminated property disputes, rather than public nuisance.[173]

---

169. *See, e.g.*, Antolini, *supra* note 162, at 774-75. Professor Antolini observed that:

> Public nuisance offers plaintiffs several important strategic advantages. Its primary advantage is a more direct focus on the merits—the existence of the nuisance, the injury, and the appropriate remedy—than is available in many statutory cases, where the focus is often on procedure or violations of permits or standards. Moreover, public nuisance gives plaintiffs the opportunity to obtain damages and injunctive relief, lacks laches and other common tort defenses, is immune to administrative law defenses such as exhaustion, avoids the private nuisance requirement that the plaintiff be a landowner/occupier of affected land, eliminates a fault requirement, and circumvents any pre-suit notice requirement.

Id. (footnotes omitted).

170. RESTATEMENT (SECOND) OF TORTS § 821B(1) (1979).

171. *Id.* § 821C(2)(b).

172. *See id.* § 821C(1) ("In order to recover damages in an individual action for a public nuisance, one must have suffered harm of a kind different from that suffered by other members of the public exercising the right common to the general public that was the subject of interference."). Distinctions between public and private nuisance vary among the states, including the requirement that a private plaintiff suffer a "special injury" providing standing to bring a public nuisance claim, may further restrict the availability of private cleanup cost remedies at contaminated groundwater sites. *See* Antolini, *supra* note 162, at 761 (discussing the development and various applications of the special injury requirement for private standing to bring a public nuisance claim). For example, some courts have found that cleanup costs incurred by a current landowner constitute such a "special injury," *see, e.g.*, Westwood Pharm., Inc. v. Nat'l Fuel Gas Distrib. Corp., 737 F. Supp. 1272, 1281-82 (W.D.N.Y. 1990) (applying New York law to hold that incurring NCP-consistent response costs constituted special injury providing standing to maintain public nuisance action), while others do not, *see, e.g.*, Philadelphia Elec. Co. v. Hercules, Inc., 762 F.2d 303, 316 (3d Cir. 1985) (applying Pennsylvania law and holding that a current landowner lacked standing to maintain a public nuisance claim because cleanup costs incurred to remediate contamination for which plaintiff is liable as current property owner did not constitute injury suffered exercising right common to general public); Hamlin Group, Inc. v. Int'l Minerals & Chem. Corp., 759 F. Supp. 925, 935 (D. Me. 1990) (applying Maine law and holding that response costs incurred by current landowner to cleanup own property did not constitute special injury suffered exercising right common to general public).

173. Public and private nuisance address different interests (rights common to the general public as compared to the use and enjoyment of private property) and employ a variety of different

Section 822 of the Restatement (Second) of Torts addresses the scope of private nuisance:

> [o]ne 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.[174]

Courts have considerable flexibility in fashioning private nuisance remedies, including money damages or injunctive relief tailored to the nature and scope of the underlying invasion.[175]

Soil or groundwater contamination often results in a substantial interference with the use and enjoyment of property and thus falls squarely within the definition of private nuisance.[176] For example, the current

---

rules (e.g., private actions to enjoin a public nuisance generally are not barred by statutes of limitations and can be brought by a current owner against the former owner of the same property). A detailed analysis of public nuisance is beyond the scope of this article. A large proportion of contaminated property disputes would fall within the scope of private nuisance law; the environmental nuisance paradigm proposed by this article addresses the application of private nuisance to contaminated property disputes.

174. RESTATEMENT (SECOND) OF TORTS § 822 (1979). Private nuisance requires proof of culpability, i.e., intentional and unreasonable, negligent, reckless, or "abnormally dangerous" conduct by the defendant. *Cf.* RODGERS, *supra* note 144, § 2.4 ("Nuisance law is continuing on the road to becoming a strict liability tort although there is more than a little meander in the chosen path."). In some jurisdictions, culpable conduct by the plaintiff may present an obstacle to recovery. *See, e.g.*, Copart Indus., Inc. v. Consol. Edison Co. of N.Y., 362 N.E.2d 968, 970 (N.Y. 1977) ("[A]lthough contributory negligence may be a defense where the basis of the nuisance is merely negligent conduct, it would not be where the wrongdoing is founded on the intentional, deliberate misconduct of defendant."); *see also* RESTATEMENT (SECOND) OF TORTS § 840E cmt. d (explaining that where plaintiff contributes to pollution and, if the harm is capable of apportionment, the apportionment will be made and the defendant will be held liable to the extent of his own contribution, but where apportionment cannot be made the plaintiff's own responsibility for the entire harm will bar recovery). Under the proposed environmental nuisance paradigm, plaintiff's culpability would be relevant to fashioning an appropriate remedy under comparative responsibility principles. *See infra* notes 440-41 and accompanying text.

175. *See,* RODGERS *supra* note 144, § 2.13 (describing available private nuisance remedies and observing that "[t]he balancing associated so prominently with nuisance law comes to the fore in the fashioning of remedies" (footnote omitted)); *see also* James B. Brown & Michael V. Sucaet, *Environmental Cleanup Efficiency: Private Recovery Actions for Environmental Response Costs*, 7 T.M. COOLEY L. REV. 363, 381 (1990) ("The common-law action of nuisance encompasses a wide variety of injuries. This type of action is particularly significant in the environmental context due to the availability of both equitable relief and damages."). When available, the flexibility of private nuisance remedies can tailor the allocation of cleanup cost responsibilities to specific remediation dispute circumstances. The severe doctrinal limitations of current private nuisance law, however, bar its application to many common contaminated property dispute problems. *See infra* notes 177-360 and accompanying text.

176. *See, e.g.*, Williams Pipe Line Co. v. Bayer Corp., 964 F. Supp. 1300, 1330 (S.D. Iowa 1997) (stating that under Iowa law, "[c]hemical contamination of land, such as underground gasoline, can qualify as a nuisance."); MSOF Corp. v. Exxon Corp, 934 So. 2d 708, 721 (La. Ct. App. 2005) (reversing the trial court's granting of summary judgment for the defendant and holding that potential contamination and health advisories arising from defendant's waste disposal support private nuisance claim); Wood v. Picillo, 443 A.2d 1244, 1248 (R.I. 1982) (maintaining hazardous waste dump site constituted public and private nuisance where "defendants' dumping operations have already caused substantial injury to defendants' neighbors and threaten to cause incalculable

owner of a contaminated site may need to incur cleanup costs to comply with regulatory requirements, restore the value and utility of the property, or pursue property development opportunities. Private nuisance, therefore, would appear well-suited to provide a basis for a current landowner to recover from other PRPs who caused or contributed to site contamination their fair share of cleanup costs. An analysis of private nuisance law across the United States, however, reveals four significant doctrinal limitations that severely limit the effectiveness of nuisance as a rule of decision in private cleanup cost disputes: (1) most states limit private nuisance claims to disputes involving neighboring property uses; (2) in many states the doctrine of *caveat emptor* bars private nuisance claims against predecessor owners; (3) many states employ an anachronistic interpretation of the continuing nuisance doctrine to render time-barred private nuisance claims at older contamination sites; and (4) the misplacement of the burden of proof regarding whether a nuisance is permanent or continuing can extinguish claims for unabated contamination and create a series of perverse incentives against proactive site investigation and informal cleanup cost dispute resolution. The following subsections will analyze each of these doctrinal limitations. Part III describes an environmental nuisance paradigm that solves each problem.

2. Geographic Limitations: Same Property Pollution Disputes

Many private cleanup cost disputes concern contamination caused or contributed to by former occupants of the contaminated property, as opposed to contamination that originated on neighboring property. "Same property" pollution arises in many settings, such as former landfills and properties that once housed manufacturing facilities or retail establishments (such as dry cleaners or service stations) that handled hazardous substances. Current landowners at these sites are left with "predecessor pollution" that impairs the value of the property and requires compliance with regulatory cleanup requirements. Most states, however, bar same property private nuisance claims by a current landowner seeking damages or equitable relief for contamination originating on her property.[177]

---

damage to the general public. The Picillos' neighbors have displayed physical symptoms of exposure to toxic chemicals and have been restricted in the reasonable use of their property. Moreover, expert testimony showed that the chemical presence on defendants' property threatens both aquatic wildlife and human beings with possible death, cancer, and liver disease.").

177. *See, e.g.*, Evans v. Lochmere Recreation Club, Inc., 627 S.E.2d 340, 342 (N.C. Ct. App. 2006) ("[A] private nuisance exists in a legal sense when one makes an improper use of his own property and in that way injures the land or some incorporeal right of one's neighbor."); Pestey v. Cushman, 788 A.2d 496, 502 (Conn. 2002) (quoting Nailor v. C.W. Blakeslee & Sons, Inc., 167 A. 548 (Conn. 1933)) ("The law of private nuisance springs from the general principle that '[i]t is the duty of every person to make a reasonable use of his own property so as to occasion no unnecessary damage or annoyance to his neighbor.'"); Demont v. Abbas, 32 N.W.2d 737, 738 (Neb. 1948) ("Generally, an owner of property has a right to make any use of it he sees fit. It is only where his use prevents his neighbors from the enjoyment of their property to their damage that an owner's use may be restricted."); Hydro-Mfg., Inc. v. Kayser-Roth Corp., 640 A.2d 950, 957 (R.I. 1994) ("Under

a. Majority View: No Same Property Private Nuisance Claims

The vast majority of states restrict private nuisance claims to disputes between neighboring landowners.[178] The seminal case barring

---

Rhode Island law it is well settled that a cause of action for a private nuisance 'arises from the unreasonable use of one's property that materially interferes with a neighbor's physical comfort or the neighbor's use of his real estate.' The offensive condition therefore must originate outside the plaintiff's land." (quoting Weida v. Ferry, 493 A.2d 824, 826 (R.I. 1985))); Sans v. Ramsey Golf & Country Club, Inc., 149 A.2d 599, 605 (N.J. 1959) ("The question is not simply whether a person is annoyed or disturbed, but whether the annoyance or disturbance arises from an unreasonable use of the neighbor's land or operation of his business."); Beckman v. Marshall, 85 So. 2d 552, 554 (Fla. 1956) (quoting Antonik v. Chamberlain, 78 N.E.2d 752, 758 (Ohio Ct. App. 1947) for the proposition that "Nuisance, in law, for the most part consists in so using one's property as to injure the land or some incorporeal right of one's neighbor . . . ."); Sewell v. Phillips Petroleum Co., 197 F. Supp. 2d 1160, 1171-72 (W.D. Ark. 2002) (granting summary judgment for defendant on same property private nuisance claim, citing Southeast Arkansas Landfill, Inc. v. State, 858 S.W.2d 665, 667 (Ark. 1993) for the proposition that "the Arkansas Supreme Court defines nuisance to contemplate separate parcels of property"); State v. Jacob Decker & Sons, 196 N.W. 600, 601 (Iowa 1924) ("The essential element in nuisance is the injury to one's neighbors, and involves an invasion of the legal rights of persons sustaining peculiar relations to the property or thing in question, or threatening or impending danger to the public."). Some states that bar "same property" private nuisance claims will nevertheless permit a private plaintiff who has suffered the "special injury" required for standing to seek recovery for "same property" contamination constituting a public nuisance. *See, e.g.*, Westwood Pharm., Inc. v. Nat'l Fuel Gas Distrib. Corp., 737 F. Supp. 1272, 1282 (W.D.N.Y. 1990) (applying New York law); *see also infra* notes 378-79 and accompanying text. In *Moore v. Texaco*, Inc., 244 F.3d 1229 (10th Cir. 2001), the Tenth Circuit affirmed summary judgment under Oklahoma law in favor of defendant former landowner against public and private nuisance claims brought by a current landowner in connection with soil contamination caused by petroleum. The court noted that while a public nuisance claim could be asserted against Texaco, the former landowner, plaintiff Moore had adduced no evidence that Texaco had either caused the pollution on the property or failed to abate prior pollution about which it had knowledge prior to selling the property to Moore. *Moore*, 244 F.3d at 1231-32. The Tenth Circuit further held that Moore could not maintain a private nuisance claim based on the petroleum pollution and the construction of berms on the property, finding it "likely that Oklahoma would reach the same conclusion reached by nearly every other court to consider the issue: that an action for private nuisance is designed to protect neighboring landowners from conflicting uses of property, not successor landowners from conditions on the land they purchased." *Id.* at 1232. Similarly, some courts have recognized an exception to the private nuisance neighboring property limitation for private nuisance claims by a landlord against a tenant. *Compare, e.g.*, Graham Oil Co. v. BP Oil Co., 885 F. Supp. 716, 725 (W.D. Pa. 1994) (denying motion to dismiss private nuisance claim under Pennsylvania law brought by landlord against tenant who operated gas station that allegedly caused contamination on leased property), *with, e.g.*, Patton v. TPI Petroleum, Inc., 356 F. Supp. 2d 921, 931-32 (E.D. Ark. 2005) (dismissing landlord's private nuisance claim against commercial tenant because under Arkansas law nuisance claims limited to neighboring property disputes).

178. *See, e.g.*, Falcone & Utain, *supra* note 150, at 78-84 (describing restriction of nuisance claims disputes between neighboring landowners as the "majority approach"); *see also* 9 RICHARD R. POWELL, POWELL ON REAL PROPERTY § 64.02[3] (Michael Allan Wolf ed., 2000) ("Unreasonableness has a role to play in private nuisance law in that plaintiffs are not expected to tolerate unreasonable interference with use and enjoyment of their real property . . . . The conclusion of 'unreasonableness' depends then upon liability-inviting conduct of the defendant plus a finding that this conduct violates a protected interest of the neighbor-plaintiff."); Edward Rabin, *Nuisance Law: Rethinking Fundamental Assumptions*, 63 VA. L. REV. 1299, 1319 (1977) ("An interference is not a nuisance unless, among other things, it *substantially* interferes with the use and enjoyment of neighboring land."); Andrew Jackson Heimert, *Keeping Pigs Out of Parlors: Using Nuisance Law to Affect the Location of Pollution*, 27 ENVTL. L. 403, 406 (1997) ("Nuisance actions to abate interferences with an owner's interest in land have existed for over eight hundred years. The pre-Revolutionary body of American nuisance law accepted the oft-repeated maxim, sic utere tuo ut alienum non laedes ('one should use his own property in such a manner as not to injure that of another')." (footnotes omitted)). By contrast, CERCLA makes the owner or operator of contaminated

private nuisance claims for same property hazardous substance contamination is the Third Circuit's 1985 decision in *Philadelphia Elec. Co. v. Hercules, Inc.*[179] Philadelphia Electric Company (PECO) acquired property in Chester, Pennsylvania, that had once been used as a hydrocarbon resin manufacturing plant and at which a former owner had disposed of resins and by-products.[180] Following discovery of the contamination, the Pennsylvania Department of Environmental Resources directed PECO to remediate the site.[181] PECO sued Hercules (as successor to the former owner of the PECO property), asserting public and private nuisance claims under Pennsylvania law.[182] PECO sought to recover cleanup costs and lost property rentals caused by the contamination, and obtain an injunction directing Hercules to abate any remaining contamination.[183] A jury awarded cleanup cost and delay damages to PECO and the district court issued the requested injunction against Hercules.[184] On appeal, the Third Circuit reversed, holding that Pennsylvania law did not permit a same property private or public nuisance claim.[185]

The Third Circuit started its private nuisance analysis by turning to the Restatement (Second) of Torts. The court assumed that the prior owner had created a nuisance within the meaning of section 821D of the Restatement (Second) of Torts, which "defines a 'private nuisance' as 'a nontrespassory invasion of another's interest in the private use and enjoyment of land.'"[186] The court found that "[t]he crucial and difficult question for us is *to whom* Hercules may be liable."[187] The court concluded that while Hercules might be liable in nuisance for a claim brought by a neighboring property owner, Hercules owed no duty under nuisance law to a successor owner of the Chester site, such as PECO.[188] The court reasoned that barring same property nuisance claims "is consonant with the historical role of private nuisance law as a means of efficiently resolving conflicts between *neighboring*, contemporaneous land uses."[189] The court noted that:

---

property at the time of disposal of a hazardous substance liable for response costs without limiting direct or derivative private response cost recovery rights to neighboring property owners. 42 U.S.C.A. § 9607(a)(2), (4) (West 2006).

179. Phila. Elec. Co. v. Hercules, Inc., 762 F.2d 303 (3d Cir. 1985).

180. *Phila. Elec. Co.*, 762 F.2d at 306.

181. *Id.* at 307.

182. *Id.* at 307-08.

183. *Id.* at 308.

184. *Id.*; *see also* Phila. Elec. Co. v. Hercules, Inc., 587 F. Supp. 144, 147 (E.D. Pa. 1984).

185. The Third Circuit affirmed that Hercules was the successor to the prior owners of the PECO property under a *de facto* merger theory, but held that PECO could not state a same property nuisance claim under Pennsylvania law. *Phila. Elec. Co.*, 762 F.2d at 312-14.

186. *Phila. Elec. Co.*, 762 F. 2d at 313.

187. *Id.*

188. *Id.* at 314. The court noted that while "PECO also owned an adjoining piece of land, and thus was a neighbor of Hercules . . . . [PECO did] not, however, allege that pollution of the Chester site interfered with its use and enjoyment of this adjoining site." *Id.* at 314 n.8.

189. *Id.* at 314.

> [A]ll of the very useful and sophisticated economic analyses of private nuisance remedies published in recent years proceed on the basis that the goal of nuisance law is to achieve efficient and equitable solutions to problems created by *discordant* land uses. In this light nuisance law can be seen as a complement to zoning regulations, . . . and not as an additional type of consumer protection for the purchasers of realty. Neighbors, unlike the purchasers of land upon which a nuisance exists, have no opportunity to protect themselves through inspection and negotiation. The record shows that PECO acted as a sophisticated and responsible purchaser—inquiring into the past use of the Chester site, and inspecting it carefully. We find it inconceivable that the price it offered Gould did not reflect the possibility of environmental risks, even if the exact condition giving rise to this suit was not discovered.[190]

The court concluded that to "extend private nuisance beyond its historical role would render it little more than an epithet, 'and an epithet does not make out a cause of action.'"[191] Finally, the court underscored its discomfort with extending private nuisance to same property disputes by referencing the nascent state of federal and state environmental statutes, observing that "[s]uch an extension of common law doctrine is particularly hazardous in an area, such as environmental pollution, where Congress and the state legislatures are actively seeking to achieve a socially acceptable definition of rights and liabilities."[192]

The court also found PECO's public nuisance claim unavailing. On the one hand, the court rejected Hercules' contention that the neighboring property dispute limitation on private nuisance claims applied equally to public nuisance claims,[193] noting that the "doctrine of public nuisance protects interests quite different from those implicated in actions for private nuisance, . . . [and that] an action for public nuisance may lie even though neither the plaintiff nor the defendant acts in the exercise of private property rights."[194] The court nevertheless concluded that under Pennsylvania law spending money to remediate predecessor pollution did not constitute the "special injury" or "particular damage"

---

190. *Id.* (citation omitted). The court further held that the doctrine of *caveat emptor* barred PECO's claim. *Id.* at 312-13. *See infra* notes 221-43 and accompanying text. The court also found analogous a series of Pennsylvania cases (later overruled) that did not permit "*tenants* to circumvent traditional limitations on the liability of *lessors* by the expedient of casting their cause of action for defective conditions existing on premises (over which they have assumed control) as one for private nuisance." *Phila. Elec. Co.*, 762 F.2d at 313. The court noted that although these tenant-landlord cases had been overruled by the time of the *Philadelphia Electric* decision, they nevertheless reflected "sound tort theory." *Id.* at 314 n.8.

191. *Id.* at 315 (citation omitted).

192. *Id.* at 315 n.13; *see also infra* notes 379-82 and accompanying text.

193. *Phila. Elec. Co.*, 762 F.2d at 315 n.13.

194. *Id.* at 315.

different in kind from that suffered by the general public required for PECO to assert a public nuisance claim.[195]

Over the next two decades, courts around the country embraced the reasoning of *Philadelphia Electric* to reject same property pollution private nuisance claims under the laws of other states.[196] These decisions tended to rely on the traditional role of private nuisance as a vehicle for resolving disputes between neighboring land uses; they typically neither identified nor rested on positive law that barred same property private nuisance claims.

---

195. *Id.* at 316.

196. *See, e.g.*, Lilly Indus., Inc. v. Health-Chem Corp., 974 F. Supp. 702, 706-08 (S.D. Ind. 1993) (applying Indiana law); Truck Components, Inc. v. K&H Corporation, No. 94 C 50250, 1995 WL 692541, at *12 (N.D. Ill., Nov. 22, 1995) (applying Illinois law); Metro. Water Reclamation Dist. of Greater Chicago v. Lake River Corp., 365 F. Supp. 2d 913, 918–19 (N.D. Ill. 2005) (stating that under Illinois law the purpose of private nuisance action is to resolve disputes between neighboring, contemporaneous landowners); Amland Props. Corp. v. Aluminum Co. of Am., 711 F. Supp. 784, 808 (D.N.J. 1989) ("Because New Jersey courts have read private nuisance to encompass only instances of danger to the public or interference with the use of adjoining land, Amland's claim here must fall."); Wellesley Hills Realty Trust v. Mobil Oil Corp., 747 F. Supp. 93, 98-99 (D. Mass 1990) ("Because the law of private nuisance requires that the interference be to persons outside the land upon which the condition is maintained, the Massachusetts Court of Appeals has held that a vendee of land does not have a private nuisance action against a vendor for its contamination prior to the sale."); Hanlin Group, Inc. v. Int'l Minerals & Chem. Corp., 759 F. Supp. 925, 935 (D. Me. 1990) (applying Maine law); Berry v. Armstrong Rubber Co., 780 F. Supp. 1097, 1103 (S.D. Miss. 1991) (applying Mississippi law to grant summary judgment on private nuisance claim, stating that "[t]he common law of nuisance does not protect a landowner from interference or harm resulting from a previous use of his property by a prior landowner."); Rosenblatt v. Exxon Co., U.S.A., 642 A.2d 180, 190 (Md. 1994) (affirming summary judgment on ground that Maryland law was consistent with holding in *Philadelphia Electric* and other courts that "a cause of action for private nuisance requires an interference with a *neighbor's* use and enjoyment of the land."); Hydro-Mfg., Inc. v. Kayser-Roth Corp., 640 A.2d 950, 957 (D.R.I. 1994) (applying Rhode Island law); Dartron Corp. v. Uniroyal Chem. Comp., Inc., 893 F. Supp. 730, 741 (N.D. Ohio 1995) (applying Ohio law); Cross Oil Co. v. Phillips Petroleum Co., 944 F. Supp. 787, 792-93 (E.D. Mo. 1996) (stating that "[t]he law of nuisance evolved as a means of resolving conflicts between neighboring contemporaneous land users" and holding that under Missouri law a current property owner may not recover against a prior owner under a private nuisance theory for a condition created on the property by the prior owner); Reg'l Airport Auth. of Louisville & Jefferson County v. LFG, LLC, 255 F. Supp. 2d 688, 691 (W.D. Ky. 2003) ("We believe that were a Kentucky court to consider this issue, it would follow the lead of the majority of the courts, which have consistently rejected allowing a subsequent landowner to recover in private nuisance from a prior land owner [sic]."); Rudd v. Electrolux Corp., 982 F. Supp. 355, 368 (M.D.N.C. 1997) (noting that under North Carolina law, "a nuisance is the improper use of one's own property in a way which injures the land or other right of one's neighbor"); Middlebury Office Park Ltd. P'ship. v. Timex Corp., No. 3:95-CV-2160 (EBB), 1998 WL 351583, at *4 & n. 2 (D. Conn. June 16, 1998) (dismissing private nuisance claim because "[i]n Connecticut, the tort of nuisance is only applicable where a person is 'making use of his own property so as to occasion unnecessary damage or annoyance to his neighbor'" and because "claims asserted by a landowner against a former owner of property are barred by the doctrine of caveat emptor unless they fall within two limited exceptions: first, the defects existing in the property were not discoverable upon reasonable inspection by the purchaser, and, second, the seller was chargeable with knowledge of the defects").

b. Minority View: Same Property Nuisance Claims Permitted

A small minority of states recognize same property private nuisance pollution remediation claims.[197] Indeed, reported decisions applying the law of just two states—California and Minnesota—have recognized such claims.[198] In 1991, the California Court of Appeal in *Mangini v. Aerojet-General Corp.* (*Mangini I*)[199] held that a current owner could sue a prior lessee for contamination originating on the current owner's property.[200] In *Mangini I*, the owners of 2,400 acres of land near Sacramento, California sued Aerojet-General, which had leased the property years before from the prior owner, for allegedly contaminating the property with waste rocket fuel materials.[201] Plaintiffs asserted claims under nine legal theories, including private nuisance.[202] The trial court sustained defendant's general demurrer without leave to amend.[203] The California Court of Appeal reversed.[204]

Aerojet-General, relying on general tort law principles, tort law treatises, and *Philadelphia Electric*,[205] argued that private nuisance law did not permit a claim by a current owner for injuries resulting from acts by a defendant committed on the same property.[206] The Court of Appeal rejected these arguments as not reflective of California law, noting that

197. *See* Falcone & Utain, *supra* note 150, at 84-91 (describing as the minority view that a property owner may sue predecessors-in-interest of the same property for private nuisance).

198. Indiana law regarding same property private nuisance claims is unclear. In *Gray v. Westinghouse Electric Corp.*, 624 N.E.2d 49 (Ind. Ct. App. 1993), plaintiffs lived next to a former landfill site at which they alleged that defendant Westinghouse disposed of PCBs. Plaintiffs sued Westinghouse, asserting, among other things, a claim for private nuisance. *Id.* at 51-52. Westinghouse argued that it could not be liable to abate the nuisance because it did not own the landfill. *Id.* at 52-53. The Indiana Court of Appeals reversed the trial court's order granting a motion to dismiss, holding that "the creator of a nuisance of land which it does not own can be required to abate the nuisance." *Id.* at 53. The court also stated that "[w]e hold that the party which causes a nuisance can be held liable, regardless of whether the party owns or possesses the property on which the nuisance originates." *Id.* Subsequently, the federal district court in *Lilly Industries, Inc. v. Health-Chem Corp.*, 974 F. Supp. 702 (S.D. Ind. 1993) held that Indiana law did not recognize same property private nuisance claims, citing to and relying on the reasoning in *Philadelphia Electric*. *Id.* at 706-08. The court found *Gray* distinguishable because it did not involve a same property nuisance claim. *Id.* at 706. *See also* Wickens v. Shell Oil Co., No. 1:05-CV-645-SEB-JPG, 2006 WL 3254544, at *6 (S.D. Ind. Nov. 9, 2006) (citing *Lilly Industries* and granting summary judgment for defendant on a same property private nuisance claim and observing that "[r]ecently, heretofore undiscovered environmental pollution on real property seems to have fostered imaginative attempts to construct or expand on a common law tort theory of recovery. But these efforts have not found support under Indiana law since the state courts have declined to utilize trespass or nuisance doctrines to resolve environmental clean-up disputes.").

199. Mangini v. Aerojet-Gen. Corp. (*Mangini I*), 230 Cal. App. 3d 1125 (Cal. Ct. App. 1991).

200. *Mangini I*, 230 Cal. App. 3d at 1133-37.

201. *Id.* at 1131-32.

202. *Id.* at 1132-33.

203. *Id.* at 1133.

204. *Id.* at 1156.

205. *Id.* at 1133-34.

206. *Id.*

"California nuisance law is a creature of statute."[207] California Civil Code section 3479 defines a nuisance as:

> [a]nything which is injurious to health, or is indecent or offensive to the senses, or an obstruction to the free use of property, so as to interfere with the comfortable enjoyment of life or property, or unlawfully obstructs the free passage or use, in the customary manner, of any navigable lake, or river, bay, stream, canal, or basin, or any public park, square, street, or highway, is a nuisance.[208]

The court traced the origins of section 3479 to nineteenth century statutes enacted during California's Gold Rush, which "recognized that a property owner could sue for nuisance where miners entered the owner's property without consent and dug ditches that interfered with the free use of the property."[209] The court acknowledged that California's statutory definition of nuisance might be broader than the traditional scope of common law nuisance, noting that "[t]he statutory definition of nuisance appears to be broad enough to encompass almost every conceivable type or interference with the enjoyment or use of land or property."[210] The court concluded that "[t]he California nuisance statutes have been construed, according to their broad terms, to allow an owner of property to sue for damages caused by a nuisance created *on* the owner's property. Under California law it is not necessary that a nuisance have its origin in neighboring property."[211] Subsequent California cases extended *Mangini I* to private nuisance claims against former property owners as well as former tenants.[212]

Minnesota law similarly has been interpreted to permit same property private nuisance claims. In *Union Pacific Railroad Co. v. Reilly Industries, Inc.*,[213] Union Pacific, the owner of real property in Minnesota, sued Reilly Industries, the successor of Republic Creosoting Company, to recover the past and future costs of cleaning up contamination caused by Republic's former creosoting operations on the property.[214] Union Pacific included a private nuisance claim in its complaint. Reilly moved for summary judgment on that claim, arguing that private nui-

---

207. *Id.* at 1134.

208. CAL. CIV. CODE § 3479 (West 2006). California Civil Code section 3480 provides that "[a] public nuisance is one which affects at the same time an entire community or neighborhood, or any considerable number of persons, although the extent of the annoyance or damage inflicted upon individuals may be unequal." *Id.* § 3480. Civil Code section 3481 provides that "[e]very nuisance not included in the definition of the last section is private." *Id.* § 3481.

209. *Mangini I*, 230 Cal. App. 3d at 1135.

210. *Id.* at 1136 (citations omitted).

211. *Id.* at 1134.

212. *See, e.g.*, Capogeannis v. Superior Court, 12 Cal. App. 4th 668, 672-73 (Cal. Ct. App. 1993) (allowing private nuisance claim by current owner against former owner and tenant); Newhall Land & Farming Co. v. Superior Court, 19 Cal. App. 4th 334, 340-43 (Cal. Ct. App. 1993) (allowing private nuisance claim by current owner against former owner).

213. Union Pac. R.R. Co. v. Reilly Indus., Inc., 4 F. Supp. 2d 860 (D. Minn. 1998).

214. *Id.* at 862-63.

sance applies only to neighboring property use disputes.[215] The court denied Reilly's motion.[216]

The court acknowledged that, "as a general matter," a claim for "private nuisance resolves conflicts between neighboring or adjacent landowners," but observed that Minnesota law imposed no such limitation.[217] The court noted that under Minnesota's nuisance statute:

> Anything which is injurious to health, or indecent or offensive to the senses, or an obstruction to the free use of property, so as to interfere with the comfortable enjoyment of life or property, is a nuisance. An action may be brought by any person whose property is injuriously affected or whose personal enjoyment is lessened by the nuisance, and by the judgment the nuisance may be enjoined or abated, as well as damages recovered.[218]

Based on the broadly-worded Minnesota statute, the court concluded that "Union Pacific may maintain a nuisance claim even though the claim does not involve adjacent property of neighboring property holders."[219]

No other state has recognized a same property private nuisance claim brought by the current owner of property against a former owner or tenant.[220] Accordingly, in the vast majority of states, the owner of contaminated property cannot use nuisance law to recover a PRP's fair share of cleanup costs regarding contamination that originated on the property rather than on a neighboring property. Under this rule, the current owner of most brownfield sites, former industrial facilities, and former landfills would be barred from recovering cleanup costs under a private nuisance theory and, after *Aviall*, could have no right to cleanup cost contribution at all.

3. Market-Based Limitations: The Doctrine of *Caveat Emptor*

The neighboring property requirement bars landowners in most states from recovering cleanup costs from predecessor PRPs under a private nuisance theory. A closely related barrier to recovery also faced by landowners is the ancient doctrine of *caveat emptor*.[221] The commentary

---

215. *Id.* at 863, 867.
216. *Id.* at 867.
217. *Id.*
218. MINN. STAT. ANN. § 561.01 (West 2006).
219. *Union Pac. R.R.*, 4 F. Supp. 2d at 867.
220. Some courts have permitted a landlord to recover cleanup costs from a tenant under a private nuisance theory. *See, e.g.*, Arawana Mills Co. v. United Techs. Corp., 795 F. Supp. 1238, 1250 (D. Conn. 1992) (applying Connecticut law). Other relationship-based theories may also be available to a landlord for recovery of cleanup costs from a tenant, including waste (failure to restore property to pre-leasehold condition) or breach of contract (e.g., breach of warranty regarding property usage or breach of lease provision requiring restoration of property to pre-leasehold condition).
221. *See, e.g.*, Klein, *supra* note 12, at 356 (observing that historically "private nuisance has had virtually no application in cases where current waste site owners and operators seek contribution

to the Restatement (Second) of Torts, section 352, summarizes the application of *caveat emptor* ("let the buyer beware") to the vendor of land:

> Under the ancient doctrine of caveat emptor, the original rule was that, in the absence of express agreement, the vendor of the land was not liable to his vendee, or a fortiori to any other person, for the condition of the land existing at the time of transfer. As to sales of land this rule has retained much of its original force, and the implied warranties which have grown up around the sale of chattels never have developed. This is perhaps because great importance always has been attached to the deed of conveyance, which is taken to represent the full agreement of the parties, and to exclude all other terms and liabilities. The vendee is required to make his own inspection of the premises, and the vendor is not responsible to him for their defective condition, existing at the time of transfer.[222]

Caveat emptor proponents embrace economic arguments to support its application to cleanup cost claims by a vendee against a vendor of real property. They argue that the doctrine encourages: (a) thorough inspection of the property by the vendee; (b) full disclosure of property defects by the vendor to avoid misrepresentation liability; and (c) the clear, efficient shifting of risk regarding property conditions without wasteful litigation transaction costs.[223]

Indeed, many states recognize *caveat emptor* as a defense to claims by the purchaser of real property against the seller with regard to property conditions existing on the transferred land, including private nuisance claims relating to hazardous substance contamination.[224] For ex-

---

from previous landowners for the costs of cleaning waste. The reason behind this lack of application is the adherence of the courts to the doctrine of caveat emptor.").

222. RESTATEMENT (SECOND) OF TORTS, § 352, cmt. a (1965).

223. *See, e.g.*, Klein, *supra* note 12, at 365-67 (arguing in favor of applying *caveat emptor* to private party nuisance claims); Michael Andrew O'Hara, *The Utilization of Caveat Emptor in CERCLA Private Party Cleanups*, 56 LAW & CONTEMP. PROBS. 149, 160 (1993) ("As a defense, caveat emptor completely disallows any recovery to a plaintiff who has had ample opportunity to examine the item under consideration. Allowing the defense of caveat emptor creates an incentive for the purchaser to inspect closely the good being purchased. This incentive is desirable from an equity standpoint because it promotes a true 'meeting of the minds' in transactions, placing the purchaser and seller on equal terms."); *see also* Truck Components, Inc. v. K-H Corp., No. 94-C 50250, 1995 WL 692541, at *12 (N.D. Ill. Nov. 22, 1995) (rejecting same property private nuisance claim and observing without expressly referencing *caveat emptor* that "[a] suit for private nuisance between successive owners of the same property would effectively displace the market's allocation of risks and subject sellers to unbargained for future liability to remote buyers.").

224. *See* Klein, *supra* note 12, at 356 ("Historically, however, private nuisance has had virtually no application in cases where current waste site owners and operators seek contribution from previous landowners for the costs of cleaning waste. The reason behind this lack of application is the adherence of the courts to the doctrine of *caveat emptor*."); *cf., e.g.*, James B. Brown & Glen C. Hansen, *Nuisance Law and Petroleum Underground Storage Tank Contamination: Plugging the Hole in the Statutes*, 21 ECOLOGY L. Q. 643, 683-94 (1994) (arguing in favor of predecessor owner liability for leaking underground storage tank contamination). *See also, e.g.*, Gordon v. Nat'l R.R. Passenger Corp., No. CIV. A. 10753, 1997 WL 298320, at *10 (Del. Ch. Mar. 19, 1997) (stating that under Delaware law, *caveat emptor* applied to private nuisance claim by vendee of contaminated property against vendor); Wilson Auto Enters., Inc. v. Mobil Oil Corp., 778 F. Supp. 101, 107 (D.R.I. 1991) (granting motion for judgment on the pleadings, because "[w]hen Wilson purchased

ample, the Third Circuit embraced *caveat emptor* as an alternative ground to bar PECO's private nuisance claim under Pennsylvania law in *Philadelphia Electric Co. v. Hercules, Inc.*[225] In *Philadelphia Electric*, the court recognized that Pennsylvania law had abolished the rule of *caveat emptor* as to the sale of new homes by a builder-vendor.[226] The court nevertheless concluded that where "corporations of roughly equal resources contract for the sale of an industrial property, and especially where the dispute is over a condition on the land rather than a structure, caveat emptor remains the rule."[227] Similarly, in *Amland Properties Corp. v. Aluminum Co. of America*,[228] a federal district court in New Jersey held that *caveat emptor* barred a property owner from asserting a private nuisance claim against the vendor / prior owner for the cost of remediating polychlorinated biphenyl (PCB) contamination.[229] Likewise, a New York federal district court in *Westwood Pharmaceuticals, Inc. v. National Fuel Gas Distribution Corp.*,[230] granted a motion to dismiss on the ground that New York courts would apply the doctrine of *caveat emptor* to preclude a private nuisance claim by a current property owner against a former owner for the costs of remediating soil and groundwater contamination.[231] The *Westwood* court, however, would not apply *caveat emptor* to dismiss the current owner's damages claim on a *public* nuisance theory. The court attempted to distinguish private nuisance *caveat emptor* cases by concluding that

---

the property from LRRC in an arms-length transaction, he lost any possible standing to sue the previous owners and lessees in negligence, nuisance, trespass, or strict liability"); Rosenblatt v. Exxon Co., U.S.A., 642 A.2d 180, 188 n.7 (Md. 1994) (rejecting same property private nuisance claim, and noting that "the common law rule of caveat emptor, although legislatively abrogated in the context of residential property, is still applicable in Maryland with regard to the sale of commercial property"); Middlebury Office Park Ltd. P'ship. v. Timex Corp., No. 3:95-CV-2160 (EBB), 1998 WL 351583, at *4 n.2 (D. Conn. 1998) (dismissing private nuisance claim because "[i]n Connecticut, the tort of nuisance is only applicable where a person is 'making use of his own property so as to occasion unnecessary damage or annoyance to his neighbor'" and because "claims asserted by a landowner against a former owner of property are barred by the doctrine of caveat emptor unless they fall within two limited exceptions: first, the defects existing in the property were not discoverable upon reasonable inspection by the purchaser, and, second, the seller was chargeable with knowledge of the defects").

225. Phila. Elec. Co. v. Hercules, Inc., 762 F.2d 303, 312-13 (3d Cir. 1985) (holding that *caveat emptor* barred current landowner's claim against successor to former owner).

226. *Phila. Elec. Co.*, 762 F. 2d at 312.

227. *Id.* at 313.

228. Amland Props. Corp. v. Aluminum Co. of Am., 711 F. Supp. 784 (D.N.J. 1989).

229. *Amland Props. Corp.*, 711 F. Supp. at 808.

230. Westwood Pharm., Inc. v. Nat'l Fuel Gas Distrib. Corp., 737 F. Supp. 1272 (W.D.N.Y. 1990).

231. *Westwood Pharm. Inc.*, 737 F. Supp. at 1282-84. The *Westwood* court stated:

> While the New York Court of Appeals has abolished the defense of caveat emptor in cases involving contracts for the construction and sale of homes, . . . Westwood has provided no convincing reason for this court to assume that the Court of Appeals would likewise refuse to apply the doctrine where, as in the present case, the vendor and the vendee of the property at issue are both sophisticated commercial enterprises who agreed to a purchase price based, apparently in large part, on the condition of the property at the time of conveyance.

Id. (citation omitted).

> [i]n light of the different interests and public-policy concerns involved in public-nuisance actions, the court has not been convinced that New York courts would dismiss Westwood's claim. This conclusion is bolstered by the nature of the alleged public nuisance involved here, contamination of the environment by hazardous substances. Knowledge about the hazards to public health and to the environment posed by hazardous wastes is increasing constantly, and this court is not willing to assume that the New York law of public nuisance is too inflexible to meet the growing public need for avenues to address these hazards, including lawsuits where public interests are being protected through a cause of action brought by a private party.[232]

Some states have refused to apply *caveat emptor* to private nuisance claims relating to hazardous substance contamination.[233]

For example, in *Hanlin Group, Inc. v. International Minerals & Chemical Corp.*,[234] the purchaser of a chemical manufacturing plant sued the seller in a Maine federal court to recover cleanup costs under a variety of theories, including private nuisance.[235] The seller moved to dismiss the buyer's tort claims, arguing that sections 352[236] and 353[237] of the Restatement (Second) of Torts barred vendor liability to the

---

232. *Id.* at 1282; *cf.* Klein, *supra* note 12, at 361-62 (noting that, while courts that apply *caveat emptor* to private nuisance claims will not do so to private claims for public nuisance, application of the special injury rule may have same limiting effect on actions against prior owner as *caveat emptor* does in private nuisance cases). Although public and private nuisance address different interests, a private party bringing a public nuisance claim nevertheless seeks a remedy (damages or injunctive relief) relating to a private injury (e.g., the "special injury" that gave the plaintiff standing to bring the claim). *Id.*

233. *See, e.g.*, Sealy Conn., Inc. v. Litton Indus., Inc., 989 F. Supp. 120, 125 (D. Conn. 1997) (holding that *caveat emptor* did not bar nuisance claim by lessor against lessee, but did bar nuisance claim by vendee against vendor); Hanlin Group, Inc. v. Int'l Minerals & Chem. Corp., 759 F. Supp. 925, 932-33 (D. Me. 1990) (applying Maine law).

234. 759 F. Supp. 925 (D. Me. 1990).

235. *Hanlin Group, Inc.*, 759 F. Supp at 927.

236. RESTATEMENT (SECOND) OF TORTS § 352 (1965) (providing that: "[e]xcept as stated in § 353, a vendor of land is not subject to liability for physical harm caused to his vendee or others while upon the land after the vendee has taken possession by any dangerous condition, whether natural or artificial, which existed at the time that the vendee took possession").

237. RESTATEMENT (SECOND) OF TORTS § 353 (1965) provides:

> (1) A vendor of land who conceals or fails to disclose to his vendee any condition, whether natural or artificial, which involves unreasonable risk to persons on the land, is subject to liability to the vendee and others upon the land with the consent of the vendee or his subvendee for physical harm caused by the condition after the vendee has taken possession, if
>
> (a) the vendee does not know or have reason to know of the condition or the risk involved, and
>
> (b) the vendor knows or has reason to know of the condition, and realizes or should realize the risk involved, and has reason to believe that the vendee will not discover the condition or realize the risk.
>
> (2) If the vendor actively conceals the condition, the liability stated in Subsection (1) continues until the vendee discovers it and has reasonable opportunity to take effective precautions against it. Otherwise the liability continues only until the vendee has had reasonable opportunity to discover the condition and to take such precautions.

Id.

vendee.[238] The court denied the motion to dismiss on the basis of *caveat emptor*.[239] The court noted that the Maine Law Court previously had declined to apply *caveat emptor* to the sale of new houses by a builder vendor and had not yet addressed a private nuisance claim for chemical contamination of land caused by a vendor.[240] The court concluded that it could not assume that Maine state courts would apply *caveat emptor* to contaminated property tort disputes.[241]

Buyers and sellers of real property should be able to expressly and knowingly allocate responsibility between them for addressing contamination found on transferred property.[242] Applying *caveat emptor* to same property contamination private nuisance cases accepts the primacy of the economic theory underlying the doctrine by automatically effecting a default allocation of cleanup cost risks to the buyer as a matter of law. There remains, however, an underlying tension between, on the one hand, economic theory assuming that a real property transfer resulting from arm's length, misrepresentation-free bargaining reflects an efficient allocation of risk that the law should respect, and on the other hand, environmental policy encouraging proactive remediation through the equitable allocation of cleanup cost responsibilities. The issue is not whether transactional circumstances and market forces are relevant to private cleanup cost disputes. They are. Instead, the question with which courts must wrestle is how these factors should affect the allocation of cleanup cost responsibilities, i.e., should *caveat emptor* (or related issues such as purchaser knowledge of site conditions reflected, perhaps, in a discounted purchase price) provide a complete defense to vendor liability or instead affect the nature and scope of any remedy.[243]

---

238. *Hanlin Group, Inc.*, 759 F. Supp. at 932.

239. *Id.* at 933. The court nevertheless granted the seller's motion to dismiss on the ground that Maine law only permitted private nuisance claims in disputes involving neighboring properties. *Id.* at 935.

240. *Id.* at 932-33.

241. *Id.* at 933.

242. *See* 42 U.S.C.A. § 9607(e)(1) (West 2006) (nothing in subsection barring agreements to transfer liability under CERCLA "shall bar any agreement to insure, hold harmless, or indemnify a party to such agreement for any liability under this section"); SmithKline Beecham Corp. v. Rohm & Haas Co., 89 F.3d 154, 158 (3d Cir. 1996) (business purchase agreement indemnification agreement enforceable against seller as "responsible parties can lawfully allocate CERCLA response costs among themselves while remaining jointly and severally liable to the government for the entire clean-up"); *see also infra* notes 404-07 and accompanying text.

243. *See, e.g.*, Seneca Meadows, Inc. v. ECI Liquidating, Inc., 983 F. Supp. 360, 364 (W.D.N.Y. 1997) (denying motion to dismiss private and public nuisance and trespass claims for property damage regarding contamination that migrated from neighboring property, rejecting defendant's argument that pre-purchase knowledge of contamination barred liability but noting that proof of such knowledge may be considered by the jury in determining damages); W. Props. Serv. Corp. v. Shell Oil Co., 358 F.3d 678, 691 (9th Cir. 2004) (buyer aware of environmental hazards as evidenced by reduced purchase price precluded from recovering entire cost of cleanup); *see also* Joseph Y. Ybarra, *Refining California's "Consent" Defense in Environmental Nuisance Cases: Determining the Proper Scope of Liability for Responsible Former Owners*, 74 S. CAL. L. REV. 1191, 1215-21 (2001) (analyzing role of "consent" defense to private nuisance claims in context of cases involving (a) release agreements in recent transfers of property, (b) older transfers of property; (c) "as is"

4. Temporal Limitations: The Doctrine of Continuing Nuisance

Many polluted sites were contaminated decades ago. Under CERCLA, the fact that a PRP contaminated a site long ago would not present a statute of limitations problem because (1) CERCLA imposes retroactive liability and (2) the limitations periods for recovering costs under CERCLA sections 107(a) and 113(f) are measured from the completion of remedial activity[244] or the date of a CERCLA settlement or judgment creating rights to contribution,[245] respectively. After *Aviall*, however, a federal law cleanup cost remedy may no longer be available to most PRPs.

Under state law, the statute of limitations can present insurmountable problems at older contamination sites. The limitations period for a nuisance claim generally begins to run when the nuisance is created[246] or, under the discovery rule, when the potential plaintiff knew or reasonably should have known of the claim.[247]

Because tortious injury to property is considered an injury to the property itself rather than to the property owner, the limitations period does not begin to run anew every time ownership of the property changes hands.[248] As a result, a current landowner who only recently discovered

---

agreements; and (d) purchaser awareness of contamination at the time of sale); *cf., e.g.*, N.Y. Tel. Co. v. Mobil Oil Corp. 473 N.Y.S.2d 172, 174 (N.Y. App. Div. 1984) (under New York law, prior owner nuisance liability was terminated if there was reasonable opportunity for the new owner to discover harmful conditions on the property).

244. Section 113(g)(2) provides that a section 107(a) claim to recover the cost of a removal action must be brought within three years of the completion of a removal action and a section 107(a) claim to recover the costs of a remedial action must be brought within six years "after initiation of physical on-site construction of remedial action." 42 U.S.C.A. § 9613(g)(2) (West 2006); *see also* 42 U.S.C.A. § 9607(23) (defining "removal"); 42 U.S.C.A. § 9607(24) (defining "remedial").

245. Section 113(g)(3) provides that no action for response cost contribution may be commenced more than three years after the date of: (a) entry of judgment under CERCLA for recovery of response costs; (b) a *de minimis* settlement under CERCLA section 122(g); (c) a cost recovery settlement under CERCLA section 122(h); or (d) entry of a judicially approved response cost settlement. 42 U.S.C.A. § 9613(g)(3).

246. *See, e.g.*, *Mangini I*, 281 Cal. Rptr. 827, 838 (Cal. Ct. App. 1991) ("[W]here a private citizen sues for damage from a permanent nuisance, the statute of limitations begins to run upon creation of the nuisance.").

247. *See, e.g.*, McKelvey v. Boeing N. Am., Inc., 86 Cal. Rptr. 2d 645, 651 (Cal. Ct. App. 1999) (under California law, "the period of limitations will begin to run without regard to whether the plaintiff is aware of the specific facts necessary to establish his claim, provided that he has . . . 'notice or information of circumstances to put a reasonable person on inquiry.'" (citations omitted)). CERCLA section 309(a)(1) applies a "federally required commencement date" to state statutes of limitation when the state law commencement date would provide a shorter limitations period. 42 U.S.C.A. § 9658(a)(1) (West 2006). The "federally required commencement date" is defined in CERCLA section 309(b)(4) as "the date the plaintiff knew (or reasonably should have known) that the personal injury or property damages [allegedly caused by a CERCLA hazardous substance, pollutant, or contaminant] were caused or contributed to by the hazardous substance or pollutant or contaminant concerned." 42 U.S.C.A. § 9658(b)(4); *see also* Freier v. Westinghouse Elec. Corp., 303 F.3d 176, 196-97 (2d Cir. 2002) (federally required commencement date preempts state statute of limitations if state law otherwise would provide for earlier commencement date).

248. *See, e.g.*, Beck Dev. Co. v. S. Pac. Transp. Co., 52 Cal. Rptr. 2d 518, 556 (Cal. Ct. App. 1996) ("In an action involving tortious injury to property, the injury is considered to be to the property itself rather than to the property owner, and thus the running of the statute of limitations against

contamination or learned of its significance may find a nuisance claim time-barred if a prior landowner had discovered the problem years before.

Otherwise time-barred nuisance claims for older contamination problems may survive depending on whether the nuisance is characterized as "continuing" rather than "permanent."[249] Broadly speaking, if an interference with the use or enjoyment of property "will not terminate or be abated, the nuisance is said to be permanent."[250] If a nuisance is characterized as permanent, the plaintiff must recover all past, present, and future damages from the activity in a single action.[251] Once the limitations period has expired, however, all claims for permanent nuisance damages are time-barred.

---

a claim bars the owner and all subsequent owners of the property."); *see also* Bauer v. Weeks, 600 S.E.2d 700, 702 (Ga. Ct. App. 2004) (affirming summary judgment for the defendant on a property damage tort claim regarding defective synthetic stucco exterior cladding even though limitations period expired before the plaintiff purchased property because, under Georgia law, "if a homeowner sells a house after the statute of limitations has run, the conveyance does not revive the cause of action" for defects in the house); *cf.* Pinole Point Props., Inc. v. Bethlehem Steel Corp., 596 F. Supp. 283, 292-93 (N.D. Cal. 1984) (applying California law to dismiss private nuisance claim as time barred where the plaintiff, who acquired property beyond the three year California limitations period and four years after the release of pollutants ceased, failed to allege lack of knowledge about the state of property or justifiable delay in discovering contamination). For same property private nuisance claims, the statute of limitations presumably could not begin to run until the owner responsible for contaminating the site sold the property. The limitations period then would begin to run when the successor owner discovered or should have discovered the contamination under the applicable state or federal discovery rule. *See supra* note 247. Statutes of limitations applicable to nuisance claims traditionally look to historic conduct (when the last act constituting the cause of action occurred) or the consequence of historic conduct (when the contamination caused by defendant was abated). Alternatively, states could take a prospective approach to environmental nuisance statutes of repose, with the trigger for the limitations period based not on historic acts but on remedial conduct, e.g., a designated period of time after plaintiff first incurs cleanup costs. *Cf., e.g.*, 42 U.S.C.A. §9613(g) (West 2006) (CERCLA section 107(a) cost recovery statutes of limitations triggered by dates of removal or remedial expenditures); *see also* Kay, *supra* note 154, at 172 (proposing environmental negligence cause of action that would toll statute of limitations during pendency of site investigation). Under such an approach, subsequent property owners would still be able to assert a nuisance-based claim for cleanup costs if the statute of limitations began to run afresh for each landowner who incurred cleanup costs. On the other hand, this variation on the discovery rule would conflict with policies of finality underlying statutes of limitations and present complications arising from different statutes of limitations applying to nuisance claims depending on the nature of the damages sought.

249. *See infra* notes 272-349 and accompanying text. Some jurisdictions refer to a "continuing" nuisance as a "temporary" nuisance.

250. *See* DAN B. DOBBS, THE LAW OF TORTS § 468 (2000). The distinction between permanent and continuing nuisance can vary dramatically by state. *See supra* notes 259-334 and accompanying text.

251. JAMES M. FISCHER, UNDERSTANDING REMEDIES § 84(b)(ii) (permanent nuisance damages include diminution in property value and lost opportunity to sell the property); *see, e.g.*, Baker v. Burbank-Glendale-Pasadena Airport Auth., 218 Cal. Rptr. 293, 294 (Cal. 1985) (all past, present and future permanent nuisance damages to be recovered in single action); Mitchell Energy Corp. v. Bartlett, 958 S.W.2d 430, 444 (Tex. Ct. App. 1997) (submitting claim for future damages to jury constituted election to proceed on permanent nuisance theory); City of Clanton v. Johnson, 17 So. 2d 669, 672 (Ala. 1944) (all permanent nuisance damages, including diminution in property value, must be sought in single suit).

In contrast, a "continuing nuisance" broadly speaking is an interference with the use and enjoyment of property that can be terminated.[252] A plaintiff may bring successive continuing nuisance actions so long as the harmful inference continues; new causes of action effectively continue to accrue until the inference comes to an end.[253] Because a continuing nuisance can stop or be abated at some point in the future, continuing nuisance remedies typically are limited to injunctive relief to abate the nuisance, an action for damages that accrued during the nuisance limitations period, or both.[254] As discussed below, characterization of a nuisance as

252. *See, e.g.*, Capogeannis v. Super. Ct., 15 Cal. Rptr. 2d 796, 801 (Cal. Ct. App. 1993) (under California law, a nuisance is continuing if offensive condition can be discontinued or reasonably abated at any time); N. States Power Co. v. Franklin, 122 N.W.2d 26, 30-31 (Minn. 1963) (under Minnesota law, "[w]here a structure is erected or junk is stored and the harmful effect is 'one that may be abated or discontinued at any time,' there is 'a continuing wrong so long as the offending object remains,' and the courts regard such as a continuing trespass." (footnote omitted)).

253. *See, e.g.*, Wilshire Westwood Ass'n v. Atl. Richfield Co., 24 Cal. Rptr. 2d 563, 569 (Cal. Ct. App. 1993) ("[W]here the nuisance involves a use which may be discontinued at any time, it is characterized as a continuing nuisance, and persons harmed by it may bring successive actions for damages until the nuisance is abated."); Tri-County Inv. Group, Ltd. v. So. States, Inc., 500 S.E.2d 22, 25 (Ga. Ct. App. 1998) (quoting City Council of Augusta v. Lombard, 28 S.E. 994, 998 (Ga. 1897), to show that under Georgia law, "where a nuisance is not permanent in its character, but is one which can and should be abated by the person erecting or maintaining it, every continuance of the nuisance is a fresh nuisance for which a fresh action will lie").

254. *See, e.g.*, FISCHER, *supra* note 251, § 84(b)(i) ("For a temporary nuisance, i.e., a nuisance that a plaintiff may seek to abate or, at its election, seek damages, when the plaintiff elects damages, the award will cover losses only up to the date of the commencement of the action or the date of trial."). Continuing nuisance damages typically are limited to damages consistent with the non-permanent nature of the nuisance (e.g., abatement costs, lost rental, lost use) suffered during the limitations period. For example, in a state with a three year statute of limitations for nuisance, a continuing nuisance plaintiff could only recover damages suffered during the three years before the suit was filed. *See, e.g.*, Lyons v. Twp. of Wayne, 888 A.2d 426, 434 (N.J. 2005) ("One who suffers a continuing nuisance, therefore, is able to collect damages for each injury suffered within the limitations period." (citation omitted)); Silvester v. Spring Valley Country Club, 543 S.E.2d 563, 567 (S.C. Ct. App. 2001) (continuing nuisance damages limited to previously unrecovered damages suffered during limitations period). These courts typically hold that prospective damages (e.g., diminution in property value, stigma to property value) are unavailable to a continuing nuisance plaintiff. *See, e.g.*, Spaulding v. Cameron, 239 P.2d 625, 629 (Cal. 1952) (directing trial court to determine whether landslide onto plaintiff's property constituted a permanent nuisance; if so, the court should award diminution in property damages and if not, the court should award injunctive relief to abate the nuisance); Baker v. Burbank-Glendale-Pasadena Airport Auth., 705 P.2d 866, 869 (Cal. 1985) ("[I]f a nuisance is a use which may be discontinued at any time, it is considered continuing in character and persons harmed by it may bring successive actions for damages until the nuisance is abated. Recovery is limited, however, to actual injury suffered prior to commencement of each action. Prospective damages are unavailable." (citation omitted)). Some courts, however, have treated a nuisance as continuing only for statute of limitations purposes, choosing to award a "permanent" (i.e., prospective) nuisance damage remedy. *See, e.g.*, Cook v. Rockwell Int'l Corp., 358 F. Supp. 2d 1003, 1012 (D. Colo. 2004) (presence on plaintiffs' property of plutonium released from nearby nuclear weapons plant constituted continuing nuisance and trespass for statute of limitations purposes, and could provide basis for prospective, diminution in value damages effectively purchasing an easement for invasion to continue); Beatty v. Wash. Metro Area Transit Auth., 860 F.2d 1117, 1125 (D.C. Cir. 1988) ("[N]uisances may be classified for two distinct purposes, one for the assessment of damages, and the other for the application of the statute of limitations."); *cf.* Mangini v. Aerojet-Gen. Corp. (*Mangini III*), 912 P.2d 1220, 1230 (Cal. 1996) (holding that plaintiff failed to prove nuisance was continuing and, because permanent nuisance claim was time-barred, declining to decide whether the same characterization of a nuisance as permanent or continuing should apply for both limitations and damages purposes). *See infra* note 439 and accompanying text.

permanent or continuing thus has three distinct consequences: (a) when the claim accrues (i.e., is claim barred by the applicable statute of limitations); (b) whether a series of successive nuisance suits is permitted (i.e., whether the doctrine of *res judicata* bars a subsequent nuisance suit for later accruing damages caused by an unabated nuisance); and (c) whether damages are available for future or only past injuries.[255]

At first blush, the "continuing nuisance" doctrine would seem to eliminate statute of limitations problems for private nuisance claims arising from the remediation of older contamination. Closer examination, however, demonstrates dramatic variation regarding how courts define a continuing nuisance[256] and, indeed, whether the continuing nuisance doctrine is available at all.[257] These variations, in turn, can determine whether any state common law private cleanup cost remedy exists for older contamination conditions.[258]

A review of continuing nuisance law across the country reveals a range of different approaches to distinguishing between a "continuing" and a "permanent" nuisance. These standards vary depending on whether the focus of the inquiry is on: (a) the failure of the defendant to abate the offensive condition; (b) the nature of the conduct creating the

---

255. *See, e.g.*, DAN B. DOBBS, DOBBS LAW OF REMEDIES: DAMAGES-EQUITY-RESTITUTION § 5.11(1) (2d ed. 1993).

256. *See, e.g.*, Schneider Nat'l Carriers, Inc. v. Bates, 147 S.W.3d 264, 272-73 (Tex. 2004) ("in other jurisdictions there is no consensus as to where the line between permanent and temporary nuisances should be . . . or how it should be applied . . . ."); Kay, *supra* note 154, at 166-67 (discussing different definitions of continuing nuisance); *see also* Robert E. King, *Chemical Contamination in California: A Continuing Nuisance*, 1997 U. CHI. LEGAL F. 483, 489-90 (1997) (arguing that failure of California legislature or courts to define standards distinguishing between permanent and continuing nuisance in contaminated property cases has led to conflicting judicial outcomes); G. Nelson Smith, *Nuisance and Trespass Claims in Environmental Litigation: Legislative Inaction and Common Law Confusion*, 36 SANTA CLARA L. REV. 39, 48-49 (1995) (noting that plaintiffs turn to common law nuisance and trespass actions to address petroleum contamination because of CERCLA and California state Superfund statute petroleum exclusion but that "what constitutes a nuisance or trespass is not defined in the California Code, particularly as applied to environmental matters. As a result, courts are compelled to hypothesize as to the true meaning of an environmental nuisance or trespass, and as to when a cause of action under such theories accrues. Such lack of statutory guidance has led to inconsistent interpretations, many times within the same jurisdiction, causing both plaintiffs and defendants in environmental nuisance and trespass litigation to question the actual availability of such remedies.").

257. *See infra* notes 330-34 and accompanying text.

258. Most states recognize both "continuing nuisance" and "continuing trespass" theories and apply a common distinction between "permanent" and "continuing" torts to each cause of action. *See, e.g.*, Mangini v. Aerojet-Gen. Corp. (*Mangini I*), 281 Cal. Rptr. 827, 841-42 (Cal. Ct. App. 1991) (applying common standard for distinguishing between permanent and continuing trespass and nuisance). Some states do not extend the "continuing tort" doctrine to negligence or strict liability for ultra-hazardous activity claims, thereby limiting the availability of these theories for disputes involving older contamination sites. *Compare, e.g.*, Church v. Gen. Elec. Co., 138 F. Supp. 2d 169, 174 (D. Mass. 2001) (under Massachusetts law, continuing tort doctrine is limited to nuisance and trespass claims and does not apply to negligence or strict liability claims), *with* Nat'l Tel. Co-op. Ass'n v. Exxon Corp., 38 F. Supp. 2d 1, 6-7 (D.D.C. 1998) (under District of Columbia law, continuing tort theory applied to toll five year limitations period on commercial property owner's negligence and strict liability for ultra-hazardous activity claims alleging that gasoline leaking from a neighbor's underground storage tanks).

offensive condition; (c) any social benefit derived from the condition, or (d) the prospective impact or form of the condition.

a. Continued Abatable Harm

In some states, the "continuing" nature of a nuisance turns on whether the offensive condition is "abatable."[259] In these jurisdictions, a new cause of action accrues every day, in effect, because of the defendant's continued failure to abate a nuisance for which the defendant is liable, resulting in the continued presence of an offensive condition. The *Mangini* decisions illustrate the abatability standard.

As described above, in 1991, the California Court of Appeal in *Mangini I* held that California's broad statutory definition of private nuisance allows a current landowner to maintain a private nuisance action against the former tenant of a prior landowner.[260] That holding, however, did not clear the way for the plaintiff's nuisance claim to proceed because the defendant also argued that the plaintiff's claim should be dismissed as a time-barred permanent nuisance action.[261] The court agreed that a permanent nuisance claim was time-barred but allowed plaintiff to file an amended complaint alleging continuing nuisance.[262] The court stressed that under California law "the crucial distinction between a permanent and continuing nuisance is whether the nuisance may be discontinued or abated,"[263] and observed that:

> [P]laintiff's land may be subject to a continuing nuisance even though defendant's offensive conduct ended years ago. That is be-

---

259. *See, e.g.*, Valdez v. Mtn. Bell Tel. Co., 755 P.2d 80, 84 (N.M. Ct. App. 1988) (presence of utility pole that could be "easily removed at a reasonable expense" constitutes a temporary nuisance); Fletcher v. City of Indep., 708 S.W.2d 158, 178 (Mo. Ct. App. 1986) (nuisance should be characterized as "temporary" if it is "scientifically possible and reasonably practicable" to abate); Knight v. City of Missoula, 827 P.2d 1270, 1277 (Mont. 1992) ("We have held that a nuisance is a continuing nuisance when: . . . at all times, the [defendant] could have abated the nuisance by taking curative action. Since the nuisance was so terminable, it cannot be deemed to be a permanent nuisance as of the creation date . . . ."); N. States Power Co. v. Franklin, 122 N.W.2d 26, 30 (Minn. 1963) ("[T]he harmful effect is 'one that may be abated or discontinued at any time.'"); Russo Farms, Inc. v. Vineland Bd. of Educ., 675 A.2d 1077, 1086 (N.J. 1996) ("[N]uisance is continuing when it is the result of a condition that can be physically removed or legally abated. In such a case, it is realistic to impute a continuing duty to the defendant to remove the nuisance, and to conclude that each new injury includes all elements of a nuisance, including a new breach of duty."); Caron v. Margolin, 147 A. 419, 420 (Me. 1929) (describing continuing nuisance as a nuisance that "is not of such a permanent nature that it can not readily be removed and thus abated"); City of Phoenix v. Johnson, 75 P.2d 30, 34-35 (Ariz. 1938) ("If it is within the power of the person by the exercise of skill and labor to abate the nuisance, he must do so. If he fails in his duty, but allows the same to continue, he is responsible for maintaining a continuing nuisance . . . . If one responsible for maintaining a nuisance is unable by the exercise of skill and labor to abate it, then it is to be regarded as permanent, because it will continue indefinitely without change . . . .").

260. *See supra* notes 199-211 and accompanying text.

261. *Mangini I*, 281 Cal. Rptr. at 837.

262. *Id.* at 841. The court also held that the continuing/permanent distinction applies equally to private nuisance claims and to private suits for damages based on public nuisances. *Id.* at 839, 842 n.15.

263. *Id.* at 840.

> cause the "continuing" nature of the nuisance refers to the continuing damage caused by the offensive condition, not to the acts causing the offensive condition to occur.[264]

Mangini I held that a nuisance was continuing if it was "abatable."[265] The court, however, did not define what it meant for a nuisance to be "abatable." This issue was not clarified until after the remanded case proceeded to trial and a jury awarded plaintiffs $13.2 million in damages.[266] The California Court of Appeal overturned the jury's verdict in *Mangini II*, ruling that the trial court should have granted defendant's motion for judgment notwithstanding the verdict because the plaintiffs had failed to prove that the contamination was a continuing nuisance, i.e., that it was "abatable."[267] In 1996, the California Supreme Court affirmed the Court of Appeal (*Mangini III*).[268] The court concluded that evidence of "mere technological feasibility" alone could not prove abatability.[269] Instead, the court held that a nuisance was "abatable"—and thus continuing—if it could be "remedied at a reasonable cost by reasonable means."[270] No such evidence was offered at trial; indeed, there was no evidence of how the site actually would be remediated.[271] The court concluded that defendant was entitled to judgment notwithstanding the verdict because:

> [W]e do not know how much land or water has to be decontaminated. We do not know how deep the decontamination would have to go. We have no idea how much it would cost but know only that it would cost unascertainable millions of dollars. [¶] On this record, there is no substantial evidence that the nuisance is abatable.[272]

---

264. *Id.* at 841. The court applied the same "abatability" standard to conclude that plaintiff could file an amended complaint alleging a continuing trespass claim. *Id.*; *see also supra* note 258.

265. *Id.* at 840.

266. Mangini v. Aerojet-Gen. Corp. (*Mangini III*), 912 P.2d 1220, 1221 (Cal. 1996).

267. *Id.* at 1224; *see also* Mangini v. Aerojet-Gen. Corp. (*Mangini II*), 31 Cal. Rptr. 2d 696, 708 (Cal. Ct. App. 2004). Pursuant to California Rule of Court 976(d)(1), the Court of Appeal decision in *Mangini II* could not be considered published once the California Supreme Court granted Mangini's petition for review. *See* CAL. RULES OF COURT R. 976(d)(1).

268. *Mangini III*, 912 P.2d at 1221.

269. *Id.* at 1227.

270. *Id.* at 1229.

271. The Plaintiffs' own experts testified that there was not enough known yet to assess what remedial measures would be necessary or effective, and the plaintiffs' counsel in closing argument acknowledged a lack of evidence about the extent of site contamination or what it would take to decontaminate plaintiff's property. *Id.* at 1225, 1227. The court also noted that there was no evidence that a government regulatory agency had set cleanup levels for the site. *Id.* at 1227. Some courts have looked to site cleanup levels set by a regulatory agency as evidence that site contamination presumptively was abatable and as evidence of the degree of cleanup required to effect reasonable abatement. *See, e.g.*, Wilshire Westwood Assocs. v. Atl. Richfield Co., 24 Cal. Rptr. 2d 562, 569 (Cal. Ct. App. 1993) (agency letter advising that cleanup satisfied agency concerns demonstrated that nuisance was abatable and thus continuing); Capogeannis v. Super. Ct., 15 Cal. Rptr. 2d 796, 805 (Cal. Ct. App. 1993) ("We are satisfied to presume that cleanup standards set by responsible public agencies sufficiently reflect expert appraisal of the best that can be done to abate contamination in particular cases.").

272. *Mangini III*, 912 P.2d at 1230.

Absent evidence that the contamination could be abated at a reasonable cost and by reasonable means, the court deemed the plaintiffs' claim to be a claim for permanent nuisance and thus barred by the statute of limitations.[273]

b. Continued Presence of Harmful Condition

Some courts have held that contamination constitutes a continuing nuisance as long as it remains on the plaintiff's property, without regard to whether the contamination is reasonably abatable. For example, in *Hoery v. United States*,[274] the plaintiff sued the United States in a Colorado federal court under the Federal Tort Claims Act alleging that toxic chemicals negligently released into the ground at an Air Force base had migrated onto the plaintiff's residential property.[275] The government moved to dismiss the plaintiff's nuisance and trespass claims as time-barred under Colorado's two year statute of limitations, arguing that all government operations causing the release of the chemicals had stopped by 1994 and that while the plaintiff knew or should have known about

273. *Id.* at 1221. Other states have similarly embraced reasonable abatability as the standard for continuing nuisance. *See, e.g.*, Traver Lakes Cmty. Maint. Ass'n v. Douglas Co., 568 N.W.2d 847, 853 (Mich. Ct. App. 1997) (continuing nuisance is "abatable by reasonable curative or remedial action"); Hudson v. Peavey Oil Co., 566 P.2d 175, 179 (Or. 1977) ("Temporary injury, or injury which is reasonably susceptible of repair, justifies damages measured by the loss of use or rental value during the period of the injury, or the cost of restoration, or both, depending on the circumstances."); City of Sioux Falls v. Miller, 492 N.W.2d 116, 118 (S.D. 1992) ("[I]f a structure, even though permanent, can be changed, repaired, or remedied at a reasonable expense to abate a nuisance, the condition is temporary." (quoting 58 AM. JUR. 2D § 30)); Pate v. City of Martin, 614 S.W.2d 46, 48 (Tenn. 1981) ("A nuisance which can be corrected by the expenditure of labor or money is a temporary nuisance. Where the nuisance is temporary, damages to property affected by the nuisance are recurrent and may be recovered from time to time until the nuisance is abated." (citation omitted)); Hedgepath v. Am. Tel. & Tel. Co., 559 S.E.2d 327, 337 (S.C. Ct. App. 2001) ("A continuing nuisance is defined as a nuisance that is intermittent or periodical. It is described as one which occurs so often that it is said to be continuing although it is not necessarily constant or unceasing. A nuisance is continuing if abatement is reasonably and practically possible."); Valdez v. Mtn. Bell Tel. Co., 755 P.2d 80, 83 (N.M. Ct. App. 1988) ("Courts have also distinguished between a permanent or temporary structure or nuisance. A permanent structure or nuisance is one that may not be readily remedied, removed or abated at a reasonable expense, or one of a durable character evidently intended to last indefinitely, costing as much to alter as to build."); Taylor v. Culloden Pub. Serv. Dist., 591 S.E.2d 197, 203 (W. Va. 2003) ("A nuisance is temporary or continuing where it is remediable, removable, or abatable, or if abatement is reasonably and practicably possible, or, according to some cases, where it is abatable at a reasonable cost, or by the expenditure of labor or money, by the defendant, or by legal process at the instance of the injured party, against the will of the person creating it."); Campbell v. Anderson, 866 S.W.2d 139, 143 (Mo. Ct. App. 1993) ("When a nuisance can be reasonably and practicably abated, it is a temporary nuisance."); Idaho Gold Dredging Corp. v. Boise Payette Lumber Co., 22 P.2d 147, 150-51 (Idaho 1933) (discharge of oil and grease into waterway constituted permanent nuisance where jury could conclude that abatement would be impractical); *cf.* Isnard v. City of Coffeeyville, 917 P.2d 882, 889 (Kan. 1996) (nuisance was not abatable and thus permanent); *accord* KY. REV. STAT. ANN. § 411.530 (West 2006) (defining permanent nuisance as "any private nuisance that: (a) cannot be corrected or abated at reasonable expense to the owner; and (b) is relatively enduring and not likely to be abated voluntarily or by court order"); KY REV. STAT. ANN. § 411.540 (West 2006) ("Any private nuisance that is not a permanent nuisance shall be a temporary nuisance.").

274. 64 P.3d 214 (Colo. 2004).

275. *Hoery*, 64 P.3d at 215-16.

the release by 1995, the suit was not filed until 1998.[276] The court granted the government's motion, rejecting the plaintiff's argument that the continued presence and migration of chemicals constituted a continuing nuisance.[277] The plaintiff appealed, and the Colorado Supreme Court accepted a request by the Tenth Circuit Court of Appeals to answer two questions pertaining to Colorado state law: "(1) Does the continued migration from defendant's property to plaintiff's property, allegedly caused by chemical releases by the defendant, constitute continuing trespass and/or nuisance under Colorado law? (2) Does the ongoing presence of those toxic chemicals on plaintiff's property constitute continuing trespass and/or nuisance under Colorado law?"[278] The court answered both questions in the affirmative.[279]

The Colorado Supreme Court concluded that:

> [i]f the defendant causes the creation of a physical condition that is of itself harmful, even after the activity that created it has ceased, a person who carried on the activity that created the condition is subject to continuing liability for the physical condition. [¶] For continuing intrusions – either by way of trespass or nuisance – each repetition or continuance amounts to another wrong, giving rise to a new cause of action. The practical significance of the continuing tort concept is that for statue of limitation purposes, the claim does not begin to accrue until the tortuous conduct has ceased.[280]

In contrast, the court distinguished cases characterizing irrigation ditches and the location of railway lines as permanent nuisances[281] because those structures were designed to promote the state's economic develop-

---

276. *Id.* at 216-17.
277. *Id.* at 217.
278. *Id.* at 215.
279. *Id.* at 218.
280. *Id.* (citations omitted). The court relied on comment (e) to section 834 of the Restatement (Second) of Torts, which provides, in pertinent part, that:

> Activities that create a physical condition differ from other activities in that they may cause an invasion of another's interest in the use and enjoyment of land after the activity itself ceases. When the invasion continues only so long as the activity is carried on, a person who ceases to have any part in the activity is not liable for the continuance of the invasion by others. But if the activity has resulted in the creation of a physical condition that is of itself harmful after the activity that created it has ceased, a person who carried on the activity that created the condition or who participated to a substantial extent in the activity is subject to the liability for a nuisance, for the continuing harm. His active conduct has been a substantial factor in creating the harmful condition and so long as his condition continues the harm is traceable to him. This is true even though he is no longer in a position to abate the condition and to stop the harm. If he creates the condition upon land in his possession and thereafter sells or leases it to another, he is subject to liability for invasions caused by the condition after the sale or lease as well as for those occurring before. When the vendor or lessor has created the condition his liability continues until the vendee or lessee discovers it and has reasonable opportunity to take effective precautions against it.

RESTATEMENT (SECOND) OF TORTS § 834 cmt. e (1965).
281. *Hoery*, 64 P.3d at 219-20.

ment.[282] Accordingly, the court concluded that the only exception under Colorado law to a continuing nuisance where a defendant fails to eliminate an ongoing harmful physical condition wrongfully placed on plaintiff's land is where the invasion serves a "socially beneficial" purpose intended to be permanent.[283] The court found the exception inapplicable in *Hoery*. The court reasoned that:

> public policy favors the discontinuance of both the continuing migration and the ongoing presence of toxic chemicals into Hoery's property and irrigation well. Under Colorado law, a tortfeasor's liability for continuing trespass and nuisance creates a new cause of action each day the property invasion continues. Hence, the alleged tortfeasor has an incentive to stop the property invasion and remove the cause of damage.[284]

Unlike the *Mangini III* decision, the *Hoery* analysis did not further require a showing that contamination could be abated by reasonable means at a reasonable cost as a condition to characterizing it as a continuing nuisance.[285]

---

282. *Id.* at 220.

283. *Id.*; *accord* Cook v. Rockwell Int'l Corp., 358 F. Supp. 2d 1003, 1007 (D. Colo. 2004) (applying *Hoery* and Colorado law and holding that a tort based on property invasion that continues in fact is a continuing tort unless the invasion will continue indefinitely and the invasion "is integral to an enterprise vital to the development of the state."); *see also infra* note 354. The *Hoery* court also recognized as a separate basis for finding a continuing nuisance the fact that the toxic pollution continued to migrate unabated onto plaintiff's land. *Hoery*, 64 P.3d at 222. *See infra* notes 315-24 and accompanying text. Other courts have focused on the continued presence of a harmful condition to characterize a nuisance as continuing. *See, e.g.*, Bradley v. Am. Smelting and Ref. Co., 709 P.2d 782, 791-92 (Wash. 1985) (applying Washington law); Kulpa v. Stewart's Ice Cream, 534 N.Y.S.2d 518, 520 (N.Y. App. Div. 1988) (continuing nuisance claim stated where contamination that migrated from underground storage tanks removed from neighboring property remained in plaintiff's well). Applying Ohio law, the Sixth Circuit in *Nieman v. NLO, Inc.*, 108 F.3d 1546 (6th Cir. 1997), also employed a continued presence approach in the context of a continuing trespass case. The court reversed a district court order granting a motion to dismiss a continuing trespass claim based on continued presence of contaminants on plaintiff's property because "under Ohio law, a claim for continuing trespass may be supported by proof of continuing damages and need not be based on allegations of continuing conduct" and noted that defendants "may be ordered to remove the uranium waste if the trespass is determined to be continuing and abatable." *Nieman*, 108 F. 3d at 1559. In dissent, Judge Krupansky argued that under Ohio law the dumping of waste constitutes a permanent trespass and nuisance, contending that the majority opinion "unjustifiably licenses the plaintiff to assert a stale permanent tort claim which could and should have been litigated within four years of Nieman's constructive discovery of the alleged radiation emissions caused by the defendants." *Id.* at 1568 (Krupansky, J., dissenting).

284. *Hoery*, 64 P.3d at 223. In a dissenting opinion, Justice Kourlis argued that a continuing tort "theory that concentrates on the nature of the conduct of the tortfeasor is the one that comports best with general tort law and the concepts of predictability and deterrence." *Id.* at 224.

285. The court did observe that "the record at this stage of the litigation indicates that the contamination is not permanent - that is, it is remediable or abatable. Although the United States did not address the factual issue, Hoery's expert opined under oath that Hoery's property could be remediated." *Id.* at 222-23 (footnote omitted). *See* Coleman, *supra* note 139, at 53-54 ("In summary, the [*Hoery*] court noted five reasons for reaching its conclusion. First, the TCE pollution is an ongoing presence and migrates continuously onto Hoery's property. Second, the daily migration and presence of TCE on Hoery's property constitute the continuing tort. Third, the undisputed record shows that the contamination is not permanent because it is abatable and remediable. Fourth, the pollution is neither socially nor economically beneficial. Finally, public policy favors termination of the

c. Continued Harmful Conduct

Other courts find that a continuing nuisance standard based on the continued presence of contamination "would clearly undermine the purposes behind statutes of limitations."[286] Instead, these courts focus on the defendant's continued polluting conduct or instrumentalities (e.g., a currently leaking underground storage tank) as the basis for characterizing a nuisance as continuing rather than permanent.[287] For example, in *Union Pacific Railroad Company v. Reilly*,[288] the current owner of Minnesota property brought a nuisance action in a Minnesota federal court against the successor to the former property owner, alleging creosote that had leached from storage tanks previously located on the property constituted a nuisance.[289] The defendant moved for summary judgment on the ground that the nuisance claim was time-barred.[290] The court found a genuine issue of material fact regarding whether the plaintiff discovered the contamination within the limitations period,[291] but ruled that the continued presence of the creosote contamination did not constitute a continuing nuisance.[292] The court concluded that under Minnesota law:

---

continued migration and ongoing presence of TCE in Hoery's property." (footnote omitted)); Dana L. Eismeier, *Continuing Trespass and Nuisance for Toxic Chemicals*, 32 COLO. LAW. 107, 110 (2003) (observing that the Colorado Supreme Court left unanswered in *Hoery* whether it "will adopt a 'reasonableness' legal standard regarding what nuisances or trespasses can be abated"); *cf.* Tatum v. Basin Res., Inc., 141 P.3d 863, 867 (Colo. Ct. App. 2005) (stating, without supporting citation or elaboration, that *Hoery* identified reasonable abatability as a permanent injury claim factor).

286. Breiggar Props. L.C. v. H. E. Davis & Sons, Inc., 52 P.3d 1133, 1136 (Utah 2002) (granting summary judgment for defendant contractor on ground that continued presence of debris dumped on plaintiff landowner's property beyond the limitations period constituted a permanent rather than continuing trespass). *See infra* notes 287-300 and accompanying text.

287. *See, e.g.*, Hicks Family Ltd. Partnership v. 1st Nat. Bank of Howell, No 268400, 2006 WL 2818514, at *9 (Mich. App. Oct 3, 2006) (stating that under Michigan law a plaintiff alleging continuing nuisance or trespass must show "continuing tortious *acts*, not merely continual harmful effects from a completed act"); Wilson v. McLeod Oil Co., Inc., 398 S.E.2d 586, 594-95 (N.C. 1990) (under North Carolina law, continued harm from migration of gasoline from present or former USTs is a "renewing" nuisance rather than a "continuing," i.e., permanent, nuisance.); Soo Line R.R. Co. v. Tang Indus., Inc., 998 F. Supp. 889, 896-97 (N.D. Ill. 1998) (private nuisance claim against a tenant for contamination arising from tenant waste disposal time barred because under Illinois law continuing tort is occasioned by continuing unlawful acts or conduct, not continued harm, so nuisance became permanent when lease expired and tenant vacated site); Blake v. Gilbert, 702 P.2d 631, 639-40 (Alaska 1985) *overruled on other grounds by* Bibo v. Jeffrey's Rest., 770 P.2d 290, 292 n.9 (Alaska 1989) (continuing nuisance (or trespass) only where "repeated and continued tortious acts are committed."); Anderson v. State, 965 P.2d 783, 789 (Haw. Ct. App. 1998) ("[A] continuing tort is defined as one inflicted over a period of time; it involves a wrongful conduct that is repeated until desisted, and each day creates a separate cause of action. A continuing tort sufficient to toll a statute of limitations is occasioned by continual unlawful acts, not by continual ill effects from an original violation, and for there to be a continuing tort there must be a continuing duty."); Carpenter v. Texaco, Inc., 646 N.E.2d 398, 399 (Mass. 1995) ("[A] continuing trespass or nuisance must be based on recurring tortious or unlawful conduct and is not established by the continuation of harm caused by previous but terminated tortious or unlawful conduct.").

288. 4 F. Supp. 2d 860 (D. Minn. 1998).

289. *Union Pac. R.R. Co.*, 4. F. Supp. 2d. at 862-63.

290. *Id.* at 863.

291. *Id.* at 865.

292. *Id.* at 867.

> [t]he continuous presence of the contaminants is insufficient to constitute a recurring damage. The disposal of creosote was in the nature of a permanent tort, rather than a continuing tort. To the extent that leakage from storage tanks or basins could constitute a continuing wrong, such wrong ceased when the storage tanks and settling basins no longer existed. There is no evidence of contamination or damage as a consequence of conduct or events which recurred on or after December 22, 1988. Thus, the court agrees with defendant that the "continuing wrong" doctrine is inapplicable to plaintiff's claims.[293]

Similarly, in *Carpenter v. Texaco, Inc.*,[294] the Supreme Judicial Council of Massachusetts held that the continued presence of gasoline contamination on the plaintiffs' property did not constitute a continuing nuisance. In *Carpenter*, the plaintiffs' property was contaminated by gasoline that had leaked from an underground storage tank formerly located on a nearby property.[295] Defendant Texaco had owned a gas station on that property but sold the station in 1980; the tank was removed in 1981 and no continuing release of gasoline from the gas station property had occurred after 1984.[296] The plaintiffs discovered the contamination in 1982 but did not bring their consolidated actions until 1991 and 1992.[297] The defendant moved for summary judgment on the ground that the plaintiffs' claims were time-barred under Massachusetts' three-year statute of limitations for permanent nuisance; plaintiffs contended that the ongoing presence of the gasoline on their property constituted a continuing nuisance.[298] The trial court granted the defendant's motion for summary judgment. The trial court's ruling was affirmed by the Supreme Judicial Council, which held that "a continuing trespass or nuisance must be based on recurring tortious or unlawful conduct and is not established by the continuation of harm caused by previous but terminated tortious or unlawful conduct."[299] The court viewed the gasoline seepage as a single encroachment that had stopped by 1985 and thus became a permanent nuisance.[300]

---

293. *Id.*
294. 646 N.E.2d 398 (Mass. 1995).
295. *Carpenter*, 646 N.E.2d at 399.
296. *Id.*
297. *Id.*
298. *Id.*
299. *Id.*
300. *Id.* at 400 & n.4. The *Carpenter* court distinguished *Sixty-Eight Devonshire, Inc. v. Shapiro*, 202 N.E.2d 811, 815 (Mass. 1964) (gutter repeatedly poured water onto the plaintiff's property); and *Asiala v. Fitchburg*, 505 N.E.2d 575, 577 (Mass. App. Ct. 1987) (continuing damage from ongoing encroachment by defective retaining wall) as cases where "rights were being invaded from time to time, and thus there were continuing trespasses or continuing nuisances" Carpenter, 646 N.E.2d at 400 n.4. The *Carpenter* court also distinguished *Wishnewsky v. Saugus*, 89 N.E.2d 783, 786 (Mass. 1950) (recurrent drainage system flooding harmed plaintiff's land); *Wells v. New Haven & Northampton Co.*, 23 N.E. 724 (Mass. 1890) (culvert repeatedly channeled water onto plaintiff's property); and *Prentiss v. Wood*, 132 Mass. 486, 487 (1882) (dam repeatedly set water back on

d. Predictability of Future Damage

Texas takes a "fairly unique"[301] approach to distinguishing between permanent and "temporary" nuisances by looking at the predictability of future harm arising from the nuisance rather than abatability or continued harmful conduct. In 2004, the Supreme Court of Texas discussed this standard at length in *Schneider National Carriers, Inc. v. Bates*.[302] In *Schneider National Carriers*, nearby residents sued industrial plant operators, alleging that air contaminants, odors, lights, and noise emitted from the plants for many years interfered with the use and enjoyment of the plaintiffs' property.[303] The trial court granted the defendants' motions for summary judgment on the ground that the plaintiffs' nuisance claims were time-barred because these long-standing conditions constituted a permanent nuisance.[304] The court of appeals reversed, finding issues of fact regarding the frequency of nuisance conditions and the feasibility of injunctive relief.[305] The Texas Supreme Court reversed the court of appeals after granting a petition for review in order "to try to clarify the distinction [between permanent and temporary nuisance], one of the oldest and most complex in Texas law."[306]

The court found that the distinction between a temporary and a permanent nuisance "must take into account the [three] reasons for which that distinction is drawn,"[307] i.e., when a nuisance claim accrues, whether a series of nuisance suits should be permitted, and whether damages for future as well as past harm should be allowed.[308] The court held that:

> [I]f a nuisance occurs several times in the years leading up to a trial and is likely to continue, jurors will generally have enough evidence of frequency and duration to reasonably evaluate its impact on neighboring property values. In such cases, the nuisance should be treated as permanent, even if the exact dates, frequency, or extent of future damage remain unknown. Conversely, a nuisance as to which any future impact remains speculative at the time of trial must be deemed "temporary."[309]

Based on this distinction, the court found that a nuisance should be deemed permanent "if it is sufficiently constant or regular (no matter how long between occurrences) that future impact can be reasonably

---

plaintiff's property) as cases involving "continuing nuisances that were not barred by the statute of limitations because of the recurring nature of the harm." *Id.* at 400 n.4.

301. Schneider Nat'l Carriers, Inc. v. Bates, 147 S.W.3d 264, 271 (Tex. 2004).
302. 147 S.W.3d at 264.
303. *Schneider Nat'l Carriers, Inc.*, 147 S.W.3d at 268.
304. *Id.* at 269.
305. *Id.*
306. *Id.* at 268.
307. *Id.*
308. *Id.* at 275.
309. *Id.* at 280 (footnote omitted).

evaluated."[310] Under this standard, there would be no need for a series of successive suits if future damages could be reasonably ascertained in one action.[311] The court also held that "the characterization of a nuisance as temporary or permanent should not depend on whether it can be abated."[312] The court reasoned that a nuisance should be enjoined only if the plaintiff did not have an adequate legal remedy (i.e., damages); to award both future damages and enjoin the nuisance would amount to a double recovery by the plaintiff.[313] It further reasoned that abating an otherwise permanent nuisance would not necessarily render its effects temporary, observing that harm such as stigma arising from the nuisance might impact the value of plaintiff's property for the foreseeable future.[314]

e. Impact of Harmful Condition Varies Over Time

Some courts have based the distinction between a continuing and permanent tort on whether the form or effect of the offensive condition changes over time so that, in effect, a new cause of action accrues with every change in condition.[315] For example, the California Court of Appeal in *Field-Escandon v. DeMann*[316] observed that "[t]he salient feature of a continuing trespass or nuisance is that its impact may vary over time."[317] In *Field-Escandon*, the plaintiff argued that because a sewer pipe buried on the plaintiff's property years before plaintiff acquired the land could be removed at any time, it constituted a continuing trespass.[318] The court affirmed summary judgment in favor of defendants on the ground that the trespass was permanent and the plaintiff's claim was

310. *Id.* at 281.
311. *Id.* at 283.
312. *Id.* at 284.
313. *Id.*
314. *Id.* at 285-86.
315. *See, e.g.*, Field-Escandon v. DeMann, 251 Cal. Rptr. 49, 53 (Cal. Ct. App. 1988) ("The salient feature of a continuing trespass or nuisance is that its impact may vary over time."); *see also* Kuhnle, *supra* note 148, at 197 n.61 (changing over time standard "appears to be controlling in hazardous waste cases"); *cf.* Briggs & Stratton Corp. v. Concrete Sales & Servs., 29 F. Supp. 2d 1372, 1378 (D. Ga. 1998) (granting summary judgment for counterclaim defendant on ground that presence of 300 barrels of chemical waste on property no longer owned by counterclaimant and without evidence of continued leaching from barrels did not support claim for continued nuisance, observing that "[a] cause of action for continuing nuisance is limited to situations where the contamination has continued to spread: the fact that Briggs & Stratton's barrels remained on the PMI property is insufficient to constitute a continuing nuisance."); City of Phoenix v. Johnson, 75 P.2d 30, 35 (Ariz. 1938) ("Permanent nuisance expresses the idea that a nuisance may continue in the same state, unless the person obligated to abate it performs his duty and changes its form so as to destroy its character as a nuisance.").
316. 251 Cal. Rptr. 49 (Cal. Ct. App. 1988).
317. *Field-Escandon*, 251 Cal. Rptr. at 53. The Supreme Court of California in *Mangini III* quoted this language from *Field-Escandon* but did not question or overrule its continuing nuisance standard in the course of affirming entry of judgment notwithstanding the verdict for defendant on the ground that plaintiff failed to meet its burden of offering substantial evidence that the nuisance was reasonably abatable. *See* Mangini v. Aerojet-Gen. Corp. (*Mangini III*), 912 P.2d 1220, 1223 (Cal. 1996); *see also supra* notes 337-45 and accompanying text.
318. *Field-Escandon*, 251 Cal. Rptr. at 52.

time-barred because the pipe's impact on the property did not vary or increase over time.[319] In *Spar v. Pacific Bell*,[320] the court took the reasoning of *Field-Escandon* one step further to hold that utility lines that the defendant already had voluntarily removed from plaintiff's property nevertheless constituted a permanent nuisance and trespass.[321] The court noted that "because defendant is a public utility, it might have been able to keep the facilities on plaintiff's property by paying just compensation to plaintiff."[322]

Under the "varying impact over time" standard, whether a private nuisance (or trespass) claim for older contamination is timely as a continuing tort or time-barred as a permanent tort could depend on the media contaminated by the defendant. Groundwater contamination conditions vary over time, as a plume of contaminated groundwater migrates and spreads through plaintiff's property and, perhaps, beyond.[323] Unlike groundwater contamination, soil contamination from a now-terminated source may not change materially over time. For example, some hazardous substances, such as DDT or PCBs, are relatively insoluble in water, readily sorb to soil and, absent the introduction of a chemical solvent or other catalyst, rarely will leach or migrate in the subsurface.[324] Even though abatement of soil contamination typically is much faster and less expensive than groundwater remediation, under a "varying over time" standard an older groundwater contamination nuisance claim could survive a statute of limitations challenge as a continuing tort while an older soil contamination claim would likely be dismissed as time-barred.

---

319. *Id.* at 53.

320. 1 Cal. Rptr. 2d 480 (Cal. Ct. App. 1991).

321. *Spar*, 1 Cal. Rptr. 2d at 484.

322. *Id.* at 483.

323. In *Hoery v. United States*, 64 P.3d 214 (Colo. 2004), the Colorado Supreme Court ruled that the continued migration of contaminated groundwater onto plaintiff's property constituted a continuing nuisance. *Hoery*, 64 P.3d at 222. The court also found that the continued presence of wrongfully placed contaminants on plaintiff's property provided an alternative basis for characterizing the nuisance as continuing for statute of limitations purposes. *Id.*; *see also supra* notes 272-85 and accompanying text. Similarly, in *Arcade Water Dist. v. United States*, 940 F.2d 1265 (9th Cir. 1991), the Ninth Circuit reversed a district court order granting a motion to dismiss as time barred a private nuisance claim based on California law brought under the Federal Tort Claims Act, alleging that contamination of a water district well from the operations of a former military laundry constituted a nuisance. *Arcade Water Dist.*, 940 F.2d at 1267. The court held that continuing damage rather than continuing wrongful activity was required to establish a continuing nuisance but, without expressly stating that changing nuisance conditions were relevant to determining whether a nuisance was permanent or continuing, stressed that "the most salient allegation is that contamination continues to leach into Arcade's Well 31." *Id.* at 1268; *see also* Stanley Works v. Snydergeneral Corp., 781 F. Supp. 659, 665-67 (E.D. Cal. 1990) (denying defendant's statute of limitations summary judgment motion on plaintiff's groundwater contamination continuing nuisance claim based on disposals that occurred years before, observing "[t]he fact that the area of contamination is getting bigger as the years go by does not convert the nuisance into a permanent nuisance. As noted above, 'the salient feature of a continuing trespass or nuisance is that its impact may vary over time.'" (quoting *Field-Escandon*, 251 Cal. Rptr. at 53)).

324. *See, e.g.*, ROBERT D. MORRISON, ENVIRONMENTAL FORENSICS: PRINCIPLES AND APPLICATIONS § 1.5 (2000) (describing immobility of hydrophobic compounds like DDT and PCBs and their remobilization through cosolvency by the introduction of a solvent such as gasoline or oil).

f. Multi-Factor Balancing

Given the wide range of tests applied by courts across the country, it is not surprising that some courts have concluded that no one factor should be determinative of whether a nuisance is continuing or permanent. Instead, a number of courts have employed a multi-factored balancing test to arrive at a case-by-case determination of whether a nuisance is continuing or permanent.[325] The California Court of Appeal's decision in *Beck Development Co. v. Southern Pacific Transportation Co.*,[326] illustrates such a multi-factored approach. In *Beck*, the owner of contaminated property sued the former owner to recover cleanup costs and obtain an injunction compelling remediation of oil contamination placed on the property decades before.[327] In determining that the nuisance was permanent rather than continuing, the California Court of Appeal considered a variety of different factors: (a) whether the defendant's offending activities had been discontinued (it had); (b) whether the nuisance would vary or change over time (tar-like petroleum contamination in the soil would not); (c) whether the contaminants would continue to migrate and cause new damage (it would not); (d) whether the nuisance could be abated at any time (it could) and, if so; (e) whether abatement was the best alternative in light of its cost (more than the value of the property), technical feasibility (excavation was feasible), legitimate competing interests (local agencies opposed excavation in light of minimal risk to groundwater or human receptors), benefits of abatement (remediated property), and risk of abatement (transporting contaminants to another location for disposal).[328] After balancing each of these factors, the court concluded that the contamination constituted a permanent nuisance and that Beck's claim was time-barred.[329] Multi-factored balancing on a case-by-case basis permits the court substantial flexibility in determining whether a claim should be time-barred as a permanent nuisance; it also is an unpredictable standard that does not promote consistent adjudication.

---

325. *See, e.g.*, Beck Dev. Co. v. S. Pac. Transp. Co., 52 Cal. Rptr. 2d 518 (Cal. Ct. App. 1996) (multi-factor balancing test under California law); City of Sioux Falls v. Miller, 492 N.W.2d 116, 118 (S.D. 1992) (applying South Dakota law and defining continuing nuisance by whether nuisance can be discontinued at any time, is intermittent or periodical in character, or involves a solid structure or a structure that can be modified); L'Enfant Plaza E., Inc. v. John McShain, Inc., 359 A.2d 5, 6 (D.C. 1976). ("Three factors for determining permanency are articulated in *D. Dobbs, Remedies* § 5.4 (1973): '(1) is the source of the invasion physically permanent, i.e., is it likely, in the nature of things, to remain indefinitely? (2) is the source of the invasion the kind of thing an equity court would refuse to abate by injunction because of its value to the community or because of relations between the parties? (3) which party seeks the permanent or prospective measure of damages?'").

326. 52 Cal. Rptr. 2d 518 (Cal. Ct. App. 1996).

327. *Beck Dev. Co.*, 52 Cal. Rptr. 2d at 526.

328. *Id.* at 557-60.

329. *Id.* at 560.

g. Continuing Nuisance Unavailable

Finally, in some states the continuing tort doctrine is simply unavailable to address property contamination problems. For example, in *Citizens & Southern Trust Co. v. Phillips Petroleum Co.*,[330] the Georgia Court of Appeals affirmed summary judgment for the defendants in an action alleging that gasoline leaking from defendants' underground storage tanks had contaminated plaintiff's commercial property, ruling that the continuing tort doctrine only applied to cases involving personal injury and was inapplicable to cases involving only property damage.[331] In New York, the state legislature in Civil Practice and Rules section 214-c effectively abolished the doctrine of continuing torts for any "action to recover damages for . . . injury to property caused by the latent effects of exposure to any substance or combination of substances" by imposing an absolute discovery rule-based three-year statute of limitations for such claims.[332] In *Jensen v. General Electric Co.*,[333] the New York Court of Appeals interpreted section 214-c to bar a continuing nuisance money damages claim brought by property owners against an electric utility that allegedly released hazardous substances that contaminated the plaintiffs' property, but not to bar a continuing nuisance claim for injunctive relief.[334]

The absence of a continuing tort doctrine dramatically limits the availability of a tort-based cleanup cost claim for sites at which the contamination was created years ago. The discovery rule could preserve a nuisance-based cleanup cost claim for landowners at sites recently sold by parties who caused the contamination years before. Otherwise, without a continuing tort doctrine, a current property owner who is voluntarily cleaning up an older contamination site could not turn to state tort law to obtain from other PRPs their fair share of cleanup costs.

5. Procedural Limitations: Burden of Proving Continuing Nuisance

The continuing nuisance doctrine theoretically can preserve private cleanup cost claims for older contamination otherwise barred by the stat-

330. 385 S.E.2d 426 (Ga. Ct. App. 1989).

331. *Citizens & S. Trust Co.*, 385 S.E.2d at 427-28.

332. N.Y. C.P.L.R. § 214-c (McKinney 2006). Section 214-c states, in pertinent part:

> Notwithstanding the provisions of section 214, the three year period within which an action to recover damages for personal injury or injury to property caused by the latent effects of exposure to any substance or combination of substances, in any form, upon or within the body or upon or within property must be commenced shall be computed from the date of discovery of the injury by the plaintiff or from the date when through the exercise of reasonable diligence such injury should have been discovered by the plaintiff, whichever is earlier.

Id.

333. 623 N.E.2d 547 (N.Y. 1993).

334. *Jensen*, 623 N.E.2d at 548-49; *see also* Syms v. Olin Corp., 408 F.3d 95, 110 (2d Cir. 2005) (under *Jensen* and section 214-c, continuing tort doctrine is unavailable for Federal Tort Claims Act nuisance claim for cleanup cost and diminution in value damages).

ute of limitations. In application, however, the ability of a current landowner to recover cleanup costs from other PRPs may turn on which party has the burden of proof (i.e., the burden of production and burden of persuasion) about whether a nuisance is permanent or continuing. The significance of the burden of proof, in turn, depends on the continuing nuisance standard employed by the court.

Few reported cases address the burden of proving whether a nuisance is permanent or continuing. Perhaps this is because the burden of proof on the issue of continuing nuisance is less likely to present a challenge to a plaintiff in jurisdictions that employ a continued harmful activity standard. Absent a dispute regarding the source of contamination, in these jurisdictions once defendant's conduct or structure causing the contamination on plaintiff's property has stopped or been removed, the continuing nuisance doctrine would no longer apply. Similarly, in courts that focus on whether the nuisance has changed over time, evidence of the basic characteristics of the contamination (e.g., immobile soil contamination as compared to ongoing leaching or contaminated groundwater migration) should resolve whether the nuisance is continuing.

The burden of proof, however, can be determinative in continuing nuisance cases in courts that employ an abatability of harm standard (or include abatability as part of a multi-factored balancing analysis), particularly those that require proof that a nuisance can be abated by reasonable cost and by reasonable means. To address the abatability of soil or groundwater contamination, a court may well require at least three types of evidence.

First, the court may expect evidence of what a cleanup that is minimally acceptable to relevant government regulatory agencies would look like. This would require evidence of an adequate site characterization, i.e., identifying the extent and nature of the contamination. It would also require evidence of applicable cleanup standards, i.e., at what point would regulatory agencies consider the site remediated.[335] In the absence of applicable government cleanup standards, the court may turn to the testimony of expert witnesses regarding appropriate cleanup levels.

Second, the court may expect evidence of whether it is technologically feasible to remediate the site to the court's satisfaction. This would require evidence of a detailed plan for accomplishing the cleanup and the

335. In theory, abatement of a chemical contamination nuisance could mean complete removal of all detectable contaminants. It is unlikely that a regulatory agency would require such a cleanup in most instances; moreover, complete removal of all contaminants at sites with groundwater or extensive soil contamination could prove technically impossible as well as unnecessary to protect public health and the environment. *See* Capogeannis v. Super. Ct., 15 Cal. Rptr. 2d 796, 805 (Cal. Ct. App. 1993) ("[W]e are satisfied to presume that cleanup standards set by responsible public agencies sufficiently reflect expert appraisal of the best that can be done to abate contamination in particular cases.").

likelihood that the contemplated technique would effect the desired remediation.

Third, the court may expect evidence of the cost of remediation. To the extent that the court will be evaluating the "reasonableness" of the proposed abatement, the court may also require evidence about the current value of the contaminated property as compared to the property's value after remediation.

A party cannot satisfy this burden of production unless (a) environmental professionals have adequately characterized the site; and (b) a regulatory agency has participated in developing a cleanup plan. At a small site with discrete soil contamination that will require only excavation and proper disposal, this burden of production could easily be met. At a larger site with complicated contamination problems, however, the burden of production could present a significant challenge. For example, development of cleanup standards could be delayed because of regulatory agency inattention or understaffing. At sites involving contaminated groundwater, fully characterizing the nature, extent, and source(s) of groundwater contamination and developing a remediation plan could take years.[336] The party who must produce evidence that the contamination constitutes a continuing nuisance runs the risk of not having necessary evidence to present at trial or to defeat summary judgment, or to persuade the court for a trial date continuance (or perhaps multiple continuances) in order to gather the necessary abatability evidence.

Courts have commonly placed the burden of proving a continuing nuisance on plaintiffs.[337] Two California appellate decisions illustrate the consequences of plaintiffs' failure to meet this burden. In *Mangini v. Aerojet-General Corp.* (*Mangini III*),[338] the Supreme Court of California affirmed a Court of Appeal decision overturning a $13.2 million jury verdict[339] for the plaintiffs-owners of contaminated property on the ground that plaintiffs failed to meet their burden of producing evidence that the contamination constituted a continuing nuisance, i.e., that it was reasonably abatable.[340]

---

336. *See, e.g.*, EPA, CLEANING UP THE NATION'S WASTE SITES, supra note 21, at 49-57 (describing remediation techniques used at NPL sites); *id.* at 269-83 (describing source location and remediation problems at contaminated groundwater sites where dense non-aqueous phase liquid contamination is present).

337. *See* Coleman, *supra* note 139, at 46 ("To establish a continuing tort, the plaintiff must show a 'substantial nexus' between the violations outside and within the limitations period. If the violations are 'sufficiently related' they are treated as one continuous violation and the statute of limitations will not be tolled because the tortious act has not ceased." (footnote omitted)).

338. 912 P.2d 1220 (Cal. 1996); *see also supra* notes 266-73 and accompanying text.

339. *Mangini III*, 912 P.2d at 1230; Mangini v. Aerojet-Gen. Corp. (*Mangini II*), 31 Cal. Rptr. 2d 696, 700 (Cal. Ct. App. 1994) (damage award was not for cleanup costs but for "a reduction in the value of the use of the property within the limitations period").

340. Plaintiffs had won the right to proceed with a same property continuing nuisance claim in Mangini v. Aerojet-General Corp. (*Mangini I*), 281 Cal. Rptr. 827 (Cal. Ct. App. 1991). *See supra* notes 199-211 and accompanying text.

The court rejected the plaintiffs' argument that they had met their burden by offering evidence that the property was contaminated and that the technology existed to decontaminate the property.[341] Instead, the court required the plaintiffs to have offered evidence that the property could be remediated by reasonable means and without unreasonable expense.[342] The court agreed that something less than total decontamination could suffice to show abatability but found that plaintiffs had failed to submit evidence of cleanup levels acceptable to or ordered by agencies for the property,[343] or expert evidence about the means and cost of site remediation.[344] Because the plaintiffs failed to prove that the nuisance was reasonably abatable, it was deemed permanent and their nuisance claim was time-barred.[345]

Less than one month after the Supreme Court of California decided *Mangini III*, the California Court of Appeal in *Beck Development Co. v. Southern Pacific Transportation Co.*[346] held that the plaintiff had failed to carry its burden of proving that a nuisance was continuing.[347] Beck had attempted to develop a large tract of land in Tracy, California, on which Southern Pacific, a prior owner, had built a reservoir to store up to 3,000,000 barrels of oil.[348] Beck sued Southern Pacific for an injunction and damages relating to oil contamination on the property. The trial court found that the oil contamination constituted a continuing nuisance, issued an injunction requiring Southern Pacific to determine the extent of contamination and remediate the contamination identified to standards provided by regulatory agencies, and ordered Southern Pacific to pay Beck $1,205,613.18 in damages.[349] The California Court of Appeal reversed Beck's continuing nuisance judgment against Southern Pacific.[350]

---

341. *Mangini III*, 912 P.2d at 1227.

342. *Id.*

343. *Id.* at 1226-27. Indeed, no such evidence was available. Although the EPA and state agencies had entered into a consent decree with Aerojet-General in 1988 to complete a remedial investigation and feasibility study to identify and evaluate remedial alternatives for the site, the study had not been prepared by the time of trial, and apparently was not expected to be ready until 1998. *Id.* at 1231 (Mosk, J. dissenting).

344. *Id.* at 1224. To the contrary, plaintiffs' expert testified that the site had not been sufficiently characterized to know what remedial measures would be necessary or whether they could be effectively accomplished, describing the potential range of cleanup costs as over $20 million and perhaps as high as $75 million. *Id.* The court also noted that plaintiffs' counsel had acknowledged to the jury during closing argument that "'[N]obody really knows how much is there, where it is, what the chemicals are, or how much it's going to cost to abate the chemicals . . . . So I guess the bottom line, if you ask yourself the question, how bad really is this site, the answer's got to be you just don't know.'" *Id.* at 1224-25.

345. *Mangini III*, 912 P.2d at 1230. Plaintiffs apparently did not object to bearing the burden that the nuisance was continuing as an element of their cause of action, but did object to proving that the contamination could be remediated by reasonable means and at a reasonable cost. *Id.* at 1226.

346. 52 Cal. Rptr. 2d 518 (Cal. Ct. App. 1996).

347. *Beck Dev. Co.*, 52 Cal. Rptr. 2d at 560.

348. *Id.* at 526-27.

349. *Id.* at 533.

350. *Id.* at 560. Beck also sued both the California Department of Toxic Substances Control (DTSC), seeking a writ of mandate compelling the DTSC to make a determination regarding

The court employed a multi-factor balancing analysis to determine whether the oil contamination constituted a continuing nuisance.[351] Relying on *Mangini III*, the court included in that analysis whether the contamination could be abated at "reasonable cost by reasonable means."[352] The court emphasized the testimony of Beck's experts that the site had been insufficiently characterized to draw firm conclusions about either the extent of contamination or the cost of remediation, and that the range of remediation costs (between $6,500,000 and $16,200,000) estimated by Beck's consultant exceeded the value of the land after remediation.[353] The court also noted that a regulatory agency had concluded that nothing further needed to be done with the contamination and instead had cautioned that excavating, treating, and disposing of the contaminated soil at an off-site location "would be significantly burdensome and from a public and regulatory point of view may not be the most advisable option."[354] The court found that "the result of the uncertainty in the record is that there is no substantial evidence of abatability,"[355] and that because Beck had the burden to prove reasonable abatability the oil contamination was deemed to constitute a permanent nuisance and Beck's private nuisance claim was time-barred.[356]

Few authorities place the burden of proving whether a nuisance is permanent or continuing on the defendant.[357] An Iowa Supreme Court

---

whether the property constituted a hazardous waste site or a site presenting no known environmental hazard, and the City of Tracy (the "City"), seeking an order that the City accept and process Beck's proposed development plan. *Id.* at 526. For several years, Beck had submitted technical reports to the DTSC at its request only to have the DTSC respond with requests for still more reports, while the City refused to process Beck's development plan until the DTSC had made a determination about the environmental condition of the property. *Id.* at 528-32. The trial court issued the requested writ to the DTSC but refused to order the City to process Beck's development plan. *Id.* at 526-27. The California Court of Appeal affirmed the judgment against the DTSC and reversed the judgment in favor of the City. *Id.* at 540, 547.

351. *Id.* at 560; *see also supra* notes 325-29 and accompanying text.

352. *Id.* at 559 (quoting *Mangini III*, 51 Cal. Rptr. at 281).

353. *Id.* at 560. The court noted that the record lacked an estimate of the actual detriment that Beck would suffer if the property was not remediated. *Id.* For example, the court's opinion does not address the potential profitability that Beck would have derived from development of remediated property.

354. *Id.* at 559-60. In *Hoery v. United States*, 64 P.3d 214, 220 (Colo. 2004), the Colorado Supreme Court held that the presence of contaminants on plaintiff's property constituted a continuing nuisance unless the contamination could and should be removed. In the abstract, the presence of contaminants in soil or groundwater does not serve any socially beneficial purpose. On the other hand, the court's concerns in *Beck* might reflect an instance where, under the *Hoery* standard, leaving the soil contamination in place would serve such a purpose, i.e., when removing the contamination would create a greater risk of harm than leaving it alone.

355. *Beck Dev. Co.*, 52 Cal. Rptr. 2d at 560.

356. *Id.* at 560. The court also concluded that the oil contamination was not a nuisance *per se* under California statutes, *id.* at 551, and that Beck failed to prove that the contamination posed a risk to drinking water supplies or other pathways to the public in order to constitute a public nuisance. *Id.* at 554.

357. *See* Brown & Hansen, *supra* note 224, at 698-99 (arguing that defendants should bear burden of proving that underground storage tank contamination constitutes a permanent nuisance rather than plaintiffs bearing the burden that the site constitutes a continuing nuisance, likening the issue to the defendant's burden of proving an affirmative defense that a claim is barred by the statute of limitations); *see also* Smith, *supra* note 164, at 60 (noting that, while plaintiffs typically bear

decision in 1903 regarding whether an improperly placed building constituted a permanent or continuing nuisance turned on which party bore the burden of proof. In *Pettit v. Incorporated Town of Grand Junction, Greene County*,[358] the court held that a private landowner could maintain a public continuing nuisance action for an injunction ordering the removal of public buildings improperly located in a street:

> As the statute did not authorize the construction of the buildings in the streets, the intention to permanently locate them there is not, in the absence of evidence to the contrary, to be inferred; that is, the burden is upon the party asserting that an obstruction in the highway is a permanent nuisance, instead of a continuing one, to establish the fact by proof. This was not done. These buildings were frame, and such as could readily have been removed by the use of modern machinery, practically without injury. The character of the foundation was not shown, and no evidence concerning the feasibility of removal was introduced. Moreover, as already observed, they were placed in a situation where the municipality had no right to locate them; and in this respect the case is distinguishable from most of those cited, where the nuisance consisted in negligently making an improvement where the party at fault had the right to construct it.[359]

Pettit, however, appears to be the exception. Plaintiffs typically bear the burden of production that a nuisance is continuing—a burden that can be determinative in cases involving a current landowner who voluntarily cleans up older contamination and wants to obtain from other PRPs their fair share of cleanup costs.[360]

---

burden of proving damages within continuing tort limitations period, statute of limitations usually is an affirmative defense).

358. 93 N.W. 381 (Iowa 1903).

359. *Pettit*, 93 N.W. at 383.

360. *See e.g.*, Morsey v. Chevron, USA, Inc., 94 F.3d 1470, 1476-77 (10th Cir. 1996) (applying Kansas law to hold that the plaintiff failed to carry burden of proof that harm to leasehold was remediable, removable, or abatable and thus constituted temporary rather than time-barred permanent nuisance damages). Some courts will presume in the face of ambiguous pleading that an alleged nuisance is continuing (or temporary). For example, in *King v. City of Independence*, 64 S.W.3d 335, 339-40, 343 (Mo. Ct. App. 2002), *overruled on other grounds*, George Ward Builders, Inc. v. City of Lee's Summit, 157 S.W.3d 644, 650 (Mo. Ct. App. 2004), the court directed a temporary nuisance judgment for plaintiff notwithstanding plaintiff's failure to specify in the complaint whether the nuisance was permanent or temporary. The court concluded that:

> If allegations are doubtful as to whether the pleaded cause of action is for a permanent nuisance or a temporary nuisance, courts should treat the nuisance as temporary. This is so "because adjudication of a permanent nuisance amounts to a grant of an easement to the wrongdoer to continue to interfere with the plaintiff's land.

King, 64 S.W.3d at 339-40 (quoting Scantlin v. City of Pevely, 741 S.W.2d 48, 50 (Mo. Ct. App. 1987)).

D. *Uncertainty and Opportunity after Aviall: The Unrealized Potential of Private Nuisance Law in Cleanup Cost Allocation Disputes*

Private nuisance has the potential to provide a flexible legal framework well-suited to resolving cleanup cost allocation disputes between the current owners of contaminated property and others who contributed to site contamination. Providing current landowners with the means to require other PRPs to pay or contribute to cleanup costs is essential to helping solve the national problem of encouraging voluntary cleanups at the nation's hundreds of thousands of contaminated sites. The current state of private nuisance law, however, is not up to the task. The dramatic variations in nuisance law among the states and significant doctrinal limitations affecting the application of nuisance law to many common contamination problems make current nuisance law an unsatisfactory response to this national problem. Moreover, the incoherent patchwork quilt of private nuisance law across the country undermines the ability of state law to restore the reliance interest equilibrium among contaminated property dispute stakeholders upended by *Aviall*, and, of perhaps greater importance, the capacity for state law to offer flexible, efficient rules of decision in contaminated property disputes that federal law cannot provide.

States can respond in one of two ways to the current deficiencies in private nuisance law. The first is to do nothing. State courts and legislatures can simply stand by and let the uncertainties created by *Aviall* about the role of federal law in private cleanup cost disputes sort themselves out. Perhaps Congress will amend CERCLA to ensure cleanup cost contribution rights for all PRPs, not just PRPs who have already been sued under CERCLA sections 106 or 107(a) or settled with the government. Or perhaps the United States Supreme Court will definitively resolve whether a PRP may recover from other PRPs their fair shares of cleanup costs under CERCLA section 107(a)(4)(B).

If all PRPs ultimately are found to lack a federal right to recover cleanup costs from other PRPs, then private cleanup cost disputes will descend into a Balkanized state of confusion, with the cleanup cost recovery rights of a PRP who voluntarily cleans up property dependant on which of 50 differing rules of decision happens to apply to a particular dispute by happenstance of geography. In the vast majority of states, no private nuisance action will be available for same property cleanup cost disputes, either because of traditional interpretations of ancient nuisance law or because of the operation of the similarly ancient doctrine of *caveat emptor*. In many states, the availability of cleanup cost claims relating to older contamination sites will turn on which of the many competing "continuing nuisance" tests a state may happen to employ. Current landowners in states with nuisance doctrine inhospitable to cleanup cost claims will be discouraged from voluntarily taking the lead in site cleanup. Instead, cleanup cost allocation disputes will increasingly shift

from direct to derivative claims, as current property owners wait for—or seek out—litigation against them as a trigger for cleanup cost contribution rights under federal or state law.

On the other hand, the chilling effect on voluntary cleanups created by *Aviall* will largely disappear if federal legislation or court decisions provide all PRPs with CERCLA cleanup cost contribution rights. Confirmation of CERCLA cleanup cost rights, however, would do nothing to effect the efficiency and flexibility benefits that potentially could be available under nuisance law, including: (a) application of a common body of state law to the cleanup cost and other issues raised in contaminated property disputes; (b) encouragement of technically sound, efficient cleanups; and (c) avoidance of the remediation and litigation transaction costs associated with NCP consistency requirements under CERCLA.[361]

The second option is for states to seize the initiative and re-examine the current state of nuisance law in their jurisdictions. Limiting private nuisance claims to disputes between contemporaneous neighboring property owners, employing overly-restrictive continuing nuisance standards, and applying outmoded interpretations of nuisance that undermine prompt site remediation ignore the realities of environmental contamination problems and private cleanup cost disputes. States should adopt a new environmental nuisance paradigm that applies to all private cleanup cost disputes faced by the current owners of contaminated property. The next section of this article proposes such a paradigm.

III. AN ENVIRONMENTAL NUISANCE PARADIGM FOR PRIVATE CLEANUP COST DISPUTES

Before *Aviall*, states allowed federal law to serve as the primary rule of decision in private cleanup cost disputes in large part by clinging to narrow, anachronistic interpretations of private nuisance law. The uncertain status of current landowner federal law cleanup cost contribution rights after *Aviall* has provided an opportunity for states to make private nuisance law relevant to a wide range of contaminated property disputes. State courts and legislatures can revitalize private nuisance law by revisiting unnecessary doctrinal limitations barring its application to same property and older contamination problems through adoption of the environmental nuisance paradigm described below. Combining such a modernized liability framework with its traditional remedial flexibility would transform private nuisance into a valuable tool for allocating remediation responsibilities at sites across the country. Moreover, it can assure cleanup cost contribution rights to CERCLA-liable but non-culpable

361. *See supra* notes 140-45 and accompanying text.

current owners of contaminated property.[362] This section identifies and analyzes the environmental nuisance paradigm components required to bring the efficiency and flexibility of private nuisance law to virtually all cleanup cost disputes faced by the current owners of contaminated property.

A. The Proposal

Every state, whether by statute or case law, should modernize its law of private nuisance as applied to contaminated property disputes in accordance with the following environmental nuisance paradigm:

> (1) A claim for private nuisance should be available in same property as well as neighboring property contamination disputes.
>
> (2) The doctrine of *caveat emptor* should be abolished as a defense to private nuisance liability in environmental nuisance cases, but the underlying circumstances surrounding a plaintiff's acquisition of contaminated property should be relevant to awarding and fashioning a private nuisance remedy.
>
> (3) Soil and groundwater contamination should presumptively constitute a continuing nuisance capable of abatement.
>
> (4) The burden of proof as to both liability and damages regarding whether a nuisance is permanent or continuing should be on the party (plaintiff or defendant) contending that a nuisance is permanent.
>
> (5) A party seeking to meet this burden of proving that an environmental nuisance is permanent would be required to show that the contamination cannot be abated by reasonable means and at a reasonable cost as measured against the highest and best potential use of the property as remediated.

This environmental nuisance paradigm serves several goals. First, it creates a template that allows every state to adopt a nuisance-based cleanup cost remedy, providing current landowners with an incentive to initiate prompt characterization and remediation efforts.[363] Second, it

362. Because the current owner of contaminated property is a liable party under CERCLA section 107(a)(1) regardless of whether she contributed to site contamination, she may not after *Aviall* obtain cleanup cost contribution from other PRPs under section 113(f) unless she first has been sued under CERCLA or settled with the government, and may not have a right to cost recovery at all under section 107(a). *See supra* notes 72-121 and accompanying text. Current private nuisance law would bar this non-culpable but CERCLA-liable current owner from recovering cleanup costs in same property or many older contamination disputes. The paradigm proposed by this article would provide cleanup cost recovery rights for this non-culpable current owner.

363. A nuisance-based paradigm requires that a plaintiff seeking to recover cleanup costs must have a property interest in the contaminated site. It thus would provide a remedy for current property owners at multi-PRP sites but would not benefit PRPs who incur cleanup costs but who are not current landowners, e.g., former owners, arrangers, transporters, neighbors. This paradigm nevertheless should address most private cleanup cost disputes. Regulatory agencies often look to the current owner of property to take the lead in site cleanup. Agencies would look to PRPs other than the current site owner to take the lead in cleaning up a site in a limited set of circumstances, e.g., the

encompasses same property contamination problems, not just neighboring property disputes. Third, it encourages site cleanup rather than payment of prospective damage awards that effectively constitute the purchase of permanent contamination easements. Fourth, it encourages informal resolution of cleanup cost disputes by allowing equitable cost allocation and presumptively making available successive continuing nuisance actions if abatable contamination is not remediated. Finally, it is consistent with the basic structure and tradition of nuisance law. Although it could most efficiently be implemented by statute, courts in many states could adopt the paradigm by interpreting the existing common law of private nuisance law to reflect the realities of soil and groundwater contamination.[364] The following discussion looks in more detail at each component of the proposed paradigm.

B. Same Property Environmental Nuisance Disputes

A private right of action for environmental nuisance should include same property contamination disputes. Properties often have been contaminated in whole or in part as a result of hazardous substance handling practices by prior site occupants, such as at former landfills, industrial facilities, gas stations, and abandoned brownfield sites. The contamination caused by these former occupants may significantly impair the current landowner's ability to develop or otherwise use the property and

---

current landowner is insolvent or otherwise incapable of proceeding with the cleanup, or the current owner meets the requirements for a defense to remediation liability, such as the third party defense found in CERCLA section 107(b)(3). Non-current landowner PRPs who incur cleanup costs because they are named on a cleanup order or otherwise take the lead in site cleanup would need to rely on legal theories other than a direct private nuisance action (e.g., contract rights, state or federal cleanup cost statutes, derivative claims) to recover cleanup costs from other PRPs.

364. Some state appellate courts may not encounter significant *stare decisis* problems in order to embrace the paradigm because they have not specifically applied the common law of private nuisance to contaminated property disputes. Many decisions applying state nuisance law to private contaminated property disputes are from federal courts estimating how a state court of last resort would apply nuisance law to environmental contamination problems. *See, e.g.*, Phila. Elec. Co. v. Hercules, Inc., 762 F.2d 303, 312 (3d Cir. 1985) (applying Pennsylvania law); Lilly Indus., Inc. v. Health-Chem Corp., 974 F. Supp. 702, 708 (S.D. Ind. 1997) (applying Indiana law); Wellesley Hills Realty Trust v. Mobil Oil Corp., 747 F. Supp. 93, 98-99 (D. Mass. 1990) (applying Massachusetts law); Amland Props. Corp. v. Aluminum Co. of Am., 711 F. Supp. 784, 808 (D.N.J. 1989) (applying New Jersey law); Westwood Pharm., Inc. v. Nat'l Fuel Gas Distrib. Corp., 737 F. Supp. 1272, 1280 (W.D.N.Y. 1990) (applying New York law). State appellate courts which have not specifically addressed the application of private nuisance law to contaminated property disputes are not bound by federal court interpretations of state law. *See, e.g.*, Howard Contracting, Inc. v. G.A. MacDonald Constr. Co., Inc., 83 Cal. Rptr. 2d 590, 597 (Cal. Ct. App. 1998) ("[F]ederal decisional authority is neither binding nor controlling in matters involving state law"); SI Sec. v. Bank of Edwardsville, 841 N.E.2d 995, 1001 (Ill. App. Ct. 2005) (Illinois courts not bound by federal court interpretation of Illinois statutes not involving federal questions). Moreover, rather than assume that a state court will apply an anachronistic interpretation of nuisance law to a contaminated property dispute, federal courts may wish to follow the lead of the Tenth Circuit in *Hoery v. United States*, 324 F.3d 1220, 1222-23 (10th Cir. 2003), and certify nuisance law questions applicable to soil or groundwater contamination cases to state courts of last resort for resolution.

subject her to costly environmental regulatory obligations.[365] Nevertheless, the vast majority of states continue to ignore the realities of environmental contamination problems by restricting private nuisance claims to disputes involving neighboring property uses.[366] This majority rule should be abandoned for several reasons.

First, the neighboring use dispute limitation is not mandated by the fundamental structure of private nuisance law. On the contrary, private nuisance provides a remedy for significant interference with the use and enjoyment of another's property. It is well-recognized that "[p]rivate nuisance is traditionally a claim based upon activities outside the land by a stranger to the title, for instance, based upon a neighbor's noise or pollution."[367] Courts have bootstrapped this traditional role of private nuisance law into a doctrinal requirement barring same property claims. For example, in *Philadelphia Electric Company v. Hercules, Inc.*,[368] the Third Circuit in 1985 focused on "the historical role of private nuisance law as a means of efficiently resolving conflicts between *neighboring* contemporaneous land uses" to hold that the current owner of contaminated property could not assert a same property private nuisance claim.[369] Courts throughout the country, over the following two decades, have cited *Philadelphia Electric* and embraced the "traditional role" of private nuisance law—often without further analysis—to bar same property private nuisance claims involving hazardous substance contamination.[370]

The fact that private nuisance claims historically served as a common law zoning tool does not exclude other purposes that could be served well by private nuisance law. The hallmark of nuisance is its flexibility; the doctrine can and should adapt as a solution to modern environmental problems.[371] The current owner of property required to

365. *See* RESTATEMENT (SECOND) OF TORTS, § 821D & cmt. e (1975) (private nuisance is an invasion of another's use and enjoyment of property, including interference with possessory interest); *id.* § 821F (claim for private nuisance requires significant harm).

366. *See supra* notes 178-96 and accompanying text.

367. DAN B. DOBBS, THE LAW OF TORTS § 463 (2000) (noting also that a few recent cases had imposed same property liability for serious contamination or ultra-hazardous conduct).

368. 762 F.2d 303 (3d Cir. 1985); *see also supra* notes 179-96 and accompanying text.

369. *Phila. Elec. Co.*, 762 F.2d at 307, 314.

370. *See supra* note 196 and accompanying text.

371. *See, e.g.*, Vogel v. Grant-Lafayette Elec. Coop., 548 N.W.2d 829, 834 (Wis. 1996) (holding that Wisconsin private nuisance law supported damages judgment for dairy farmers against electric cooperative for damage to their dairy herd allegedly caused by stray electrical voltage, noting that "[The Wisconsin Supreme Court] has previously characterized the common law doctrine of private nuisance as being both 'broad' to meet the wide variety of possible invasions, and 'flexible' to adapt to changing social values and conditions." (citation omitted)); Cook v. Rockwell Int'l. Corp., 358 F. Supp. 2d 1003, 1012-14 (D. Colo. 2004) (presence on plaintiffs' property of plutonium released from nearby nuclear weapons plant constituted a continuing nuisance and trespass for statute of limitations purposes, and could provide basis for prospective diminution in value damages, effectively purchasing an easement for invasion to continue given Colorado's "adoption of a flexible approach to determining the appropriate measure of damages for injury to real property"); Carter v. Monsanto Co., 575 S.E.2d 342, 348 (W. Va. 2002) (Starcher, J., concurring) ("The cause of action

remediate the contamination caused by others experiences the same interference with the use and enjoyment of the property whether the contamination came from a prior owner or a predecessor. Indeed,

> [a]s Oliver Wendell Holmes commented on the development of the common law: "It is revolting to have no better reason for a rule of law than that so it was laid down in the time of Henry IV. It is still more revolting if the grounds upon which it was laid down have vanished long since, and the rule simply persists from blind imitation of the past."[372]

Simply put, the historic pedigree of private nuisance law in neighboring property owner disputes should not pose an insurmountable barrier to its adoption as a tool for resolving same property contamination disputes.

Second, current common law or statutory definitions of nuisance do not compel courts to bar same property private nuisance claims. Section 821D of the Restatement (Second) of Torts defines a private nuisance as an "invasion of another's interest in the private use and enjoyment of land."[373] The *Philadelphia Electric* court, among many others, cited section 821D to define a private nuisance for purposes of conducting its Pennsylvania law analysis.[374] The section 821D definition is broad enough to encompass same property disputes where the continued presence of contamination caused by the prior owner invades and substantially interferes with the use and enjoyment of land by another, i.e., the current owner.

Some courts adopting the *Philadelphia Electric* analysis have distinguished California cases[375] recognizing same property private nuisance claims on the ground that (a) the law of nuisance is codified in California and (b) California's broad statutory definition of nuisance has

---

for private nuisance has been for centuries a highly flexible one, giving courts substantial latitude to fashion appropriate and reasonable remedies, depending on the harm to be avoided or remedied."); Antolini, *supra* note 162, at 771 ("Although amorphous in definition, all nuisance actions have in common three important doctrinal aspects that provide unique scope to their application by the courts: substantiality of interference, unreasonableness of the defendant's conduct, and equitable flexibility.").

372. Antolini, *supra* note 162, at 790 (quoting Oliver W. Holmes, *The Path of the Law*, 10 HARV. L. REV. 457, 469 (1897)).

373. RESTATEMENT (SECOND) OF TORTS § 821D (1975); *see also id.*, cmt. e ("If the interference with the use and enjoyment of the land is a significant one, sufficient in itself to amount to a private nuisance, the fact that it arises out of or is accompanied by a trespass will not prevent recovery for the nuisance, and the action may be maintained upon either basis as the plaintiff elects or both.").

374. *Phila. Elec. Co.*, 762 F.2d 303, 313 (3d Cir. 1985).

375. The California state courts are the only state courts as of this writing to have recognized in a published opinion private nuisance same property claims. *See supra* notes 199-212 and accompanying text. A Minnesota federal district court in *Union Pacific Railroad Co. v. Reilly Industries, Inc.*, 4 F. Supp. 2d 860, 867 (D. Minn. 1998) interpreted Minnesota's nuisance statute to permit private nuisance same property claims. *See supra* notes 213-19 and accompanying text. Research conducted for this article revealed no reported Minnesota state court decision or state court decision from any jurisdiction other than California recognizing private nuisance same property claims.

been expansively interpreted by the California courts.[376] The breadth of California's statutory definition of nuisance, however, is immaterial because the definition of nuisance in most states (including those that adopt the section 821D definition from the Restatement (Second) of Torts is sufficiently broad to encompass same property private nuisance claims as well.[377]

Third, some courts that bar same property private nuisance claims nevertheless permit same property public nuisance claims brought by private parties with standing to sue because in exercising a common public right they have suffered a "special injury" different in kind or significant degree from that suffered by the general public.[378] However, "it is difficult to see any material difference between a public and a private nuisance in the context of a subsequent private landowner seeking to sue the previous owner for contamination of the property."[379] Permitting a current property owner to sue a predecessor for *damages* (e.g., cleanup costs) under a public nuisance theory while barring such a claim under a private nuisance theory cannot meaningfully be explained by the different interests protected by public and private nuisance and further demonstrates the need for courts to think afresh about the role of nuisance in contaminated property disputes.

Fourth, the *Philadelphia Electric* court was reluctant in its 1985 decision to "extend private nuisance beyond its historical role"[380] because "[s]uch an extension is particularly hazardous in an area, such as environmental pollution, where Congress and state legislatures are actively seeking to achieve a socially acceptable definition of rights and liabilities."[381] Two decades later, such concerns should no longer inhibit a court from recognizing same property private nuisance claims. Statutory

376. *See, e.g.*, Truck Components, Inc. v. K&H Corp, No. 94 C 50250, 1995 WL 692541, at *12 n.9 (N.D. Ill. Nov. 22, 1995) (California same property nuisance case law based on unique California statutory scheme).

377. *See* Lilly Indus., Inc. v. Health-Chem Corp., 974 F. Supp. 702, 707 (S.D. Ind. 1997) (noting that the breadth of California's statutory definition of private nuisance is not unique to California).

378. *See supra* notes 170-72 and accompanying text. *See, e.g.*, Westwood Pharm., Inc. v. Nat'l Fuel Gas Distrib. Corp., 737 F. Supp. 1272, 1281-84 (W.D.N.Y. 1990) (applying New York law to bar private nuisance claim under doctrine of *caveat emptor* but permitting public nuisance same property claim because of unspecified "different interests and public-policy concerns involved in public nuisance actions"); *cf., e.g.*, Hydro-Mfg., Inc. v. Kayser-Roth Corp., 640 A.2d 950, 957-58 (R.I. 1994) (citing *Philadelphia Electric* to bar same property private nuisance claim while recognizing same property public nuisance claim, but granting summary judgment for defendant because plaintiff failed to suffer "special damage").

379. Truck Components, Inc. v. K&H Corp., No. 94 C 50250, 1995 WL 692541, at *12 (N.D. Ill. Nov. 22, 1995) (applying Illinois law to bar both private nuisance and public nuisance same property claims and also observing that "[w]hile a public nuisance might be actionable against the prior owner if brought on behalf of the public, the present property owner cannot avoid the limitation against bringing a private nuisance action merely by converting his claim to one of public nuisance.").

380. *Phila. Elec. Co.*, 762 F.2d at 315.

381. *Id.*

environmental law has matured from its nascent stages in the mid-1980s. Legislatures have had an opportunity to develop statutory liability schemes addressing environmental regulatory liability and in the course of doing so have largely chosen not to preempt state common law tort theories of liability.[382] Moreover, same property private nuisance claims are consistent in structure with provisions of CERCLA and state Superfund laws that impose strict liability on prior property occupants responsible for site contamination.[383] Recognizing same property private nuisance claims in the early twenty-first century will not interfere with legislative environmental policy making; on the contrary, it would promote legislative goals of encouraging prompt, voluntary cleanups and promoting informal resolution of cleanup cost allocation disputes.

Fifth, courts should re-examine the propriety of same property private nuisance claims in light of the fact that the development of state nuisance law governing private cleanup cost disputes has disproportionately occurred in federal court.[384] Under the *Erie*[385] doctrine, "a federal court sitting in diversity must apply the state substantive law as pronounced by the state's highest court or, if there has been no such decision, must predict how the state's highest court would decide were it confronted with the problem."[386] A federal court presented with a question of first impression regarding whether to permit a same property private nuisance claim involving hazardous substance contamination damages might be reluctant to re-examine the "historical role" of private nuisance as a tool to resolve neighboring property use disputes. For exam-

382. *See, e.g.*, 42 U.S.C.A. §§ 9614(a), 9652(d), 9659(h) (West 2006) (CERCLA "savings clauses" regarding liability under state law theories); Edwards v. First Nat. Bank of Ne., 712 A.2d 33 (Md. Ct. Spec. App. 1998) (trial court erred by granting motion to dismiss groundwater contamination common law claims against lender on ground that Maryland environmental code lender liability exemption preempted nuisance, trespass, negligence, and strict liability claims); City of Lodi v. Randtron, 13 Cal. Rptr. 3d 107, 119-22 (Cal. Ct. App. 2004) (holding that local ordinance modeled after CERCLA preempted as conflicting with and unauthorized by California Hazardous Substance Account Act (HSAA) but noting that HSAA preserved common law liabilities for parties responsible for hazardous substance contamination). *But see* City of Modesto Redevelopment Agency v. Dow Chem. Co., No. 999345, 999643, 2005 WL 1171998, at *18-19 (Cal. Super. Ct., April 11, 2005) (barring common law damage claims that conflicted with HSAA); *see also* Aronovsky, *supra* note 5, at 85-86.

383. *See, e.g.*, 42 U.S.C.A. § 9607(a)(2) (definition of "covered persons" includes persons who owned or operated a facility at the time of disposal of a hazardous substance); CAL. HEALTH & SAFETY CODE § 25323.5 (West 2005) (incorporating CERCLA definitions of "responsible party" and "liable person"); IND. CODE ANN. § 13-25-4-8 (West 2006) (incorporating CERCLA definition of liable parties).

384. Federal courts have exclusive subject matter jurisdiction over claims "arising under" CERCLA. 42 U.S.C.A. § 9613(b). CERCLA plaintiffs have often joined state law cleanup cost and other contamination damage claims, such as private nuisance claims, to CERCLA claims pursuant to a federal court's supplemental jurisdiction to hear state law claims for which the court otherwise would lack subject matter jurisdiction. 28 U.S.C.A. § 1367(a) (West 2006). Private nuisance claims also could be brought in federal court (either joined to a CERCLA claim or as an independent claim) pursuant to diversity of citizenship federal court subject matter jurisdiction. 28 U.S.C.A. §§ 1331-1332.

385. Erie R.R. Co. v. Tompkins, 304 U.S. 64 (1938).

386. Craig v. Lake Asbestos of Quebec, Ltd., 843 F.2d 145, 149 (3d Cir. 1988) (citation omitted).

ple, in *Westwood Pharmaceuticals, Inc. v. National Fuel Gas Distribution Corp.*,[387] the plaintiff argued to a New York federal district court that the doctrine of *caveat emptor* did not bar a same property private nuisance claim brought under New York law, pointing to a series of New York state court decisions recognizing various exceptions to the doctrine.[388] The court declined to accept plaintiff's argument, stating that:

> [w]ithout a more definitive indication from New York's courts that the state's common law doctrine of caveat emptor does not apply to cases such as the one at bar, well-established principles of federalism dictate that this court refrain from extending the scope of private nuisance liability beyond its traditional bounds.[389]

These federal court decisions[390]—particularly *Philadelphia Electric*—came to form many of the building blocks for the majority rule that private nuisance contaminated property claims must involve a neighboring property dispute. State (as well as federal) courts need to re-examine the majority rule rather than further enable the development of a body of same property contamination private nuisance jurisprudence created largely by happenstance because of the *Erie* doctrine.

Finally, as a matter of policy, "expanding" private nuisance law beyond its traditional role in neighboring property use disputes also is consistent with the traditional flexibility of nuisance law to address and adapt to changing societal and economic conditions.[391] The "traditional" role of private nuisance law—as a vehicle for resolving neighboring property disputes—was established long before the development of modern environmental law, the creation of contaminated property regulatory obligations, and growth of scientific knowledge and invention of technical tools necessary to discover (or begin to discover) the presence of health risks and environmental problems presented by hazardous substance contamination. Private nuisance law is broad and flexible enough to meet today's modern contamination problems and can apply with equal force to interference with the use and enjoyment of property caused by same property as well as neighboring property contamination.

---

387. Westwood Pharm., Inc. v. Nat'l Fuel Gas Distrib. Corp., 737 F. Supp. 1272 (W.D.N.Y. 1990); *see also supra* notes 231-32 and accompanying text.

388. *Westwood*, 737 F. Supp. at 1283-84.

389. *Id.* at 1284. The court also cited to the Third Circuit's decision in *Philadelphia Electric*, barring a same property private nuisance claim under Pennsylvania law. *Id.* at 1284 n.12.

390. *See supra* note 195.

391. For example, economic disruption occasioned by the issuance of injunctive relief in private nuisance cases has evolved from an irrelevant consideration to a component in a balancing analysis (comparing the social utility of the actor's conduct and the total amount of economic and other harm that would result from enjoining an offending use) employed to determine whether an injunction should issue at all and, if so, under what conditions. *See, e.g.*, Jeff L. Levin, *Compensated Injunctions and the Evolution of Nuisance Law*, 71 IOWA L. REV. 775, 793-803 (1986) (describing development of modern nuisance law and economic theory). There is a wide range of scholarship discussing in detail the evolution of nuisance law from its twelfth century origins to the present. *See, e.g.*, *supra* note 166.

C. *Abolish the* Caveat Emptor *Defense to Environmental Nuisance Liability*

Along with adherence to the "traditional role" of private nuisance as a rule of decision only for neighboring property use disputes, the doctrine of *caveat emptor* has presented a second barrier to same property private nuisance cleanup cost dispute claims.[392] *Caveat emptor* should not provide a defense to private environmental nuisance liability; instead, the circumstances surrounding the current landowner's acquisition of the property should remain relevant to the nature and scope of any remedy ordered by the court.

Several arguments could be advanced in support of *caveat emptor* as a defense in contaminated property disputes: (a) it provides a bright line for risk allocation in real property transactions;[393] (b) it encourages full disclosure by sellers and diligent property investigations by buyers, promoting a true "meeting of the minds" regarding risk allocation in real property transactions;[394] and (c) it prevents wasteful transaction costs associated with vendor/vendee litigation about property conditions that are better addressed by the allocation of risk by market forces.[395] These arguments, however, fail to justify allocating exclusively to current landowners all of the costs of remediating contamination to which they may have made little or no contribution.

First, any benefits of a *caveat emptor* bright line for allocating risk[396] are outweighed by the inequity of allowing sellers of contami-

---

392. *See supra* notes 221-43 and accompanying text. *See also* Kay, *supra* note 154, at 170-71 (proposing that *caveat emptor* should not provide a defense to a same property contamination damages claim brought under a proposed theory of environmental negligence, arguing that Restatement (Second) of Torts sections 352 (*caveat emptor* exception when seller knows of harmful condition but conceals or fails to disclose it to seller) and 840A (seller remains liable as if still in position for physical harm caused by nuisance condition) provide a "strong foundation on which to rest liability for land sellers who polluted the property before sale.").

393. *See* Albert G. Besser, *Caveat Emptor - Where Have You Gone?*, 4 HOFSTRA PROP. L.J. 203, 226-27 (1992) (noting the unfairness of allowing a buyer of contaminated property to sue a seller notwithstanding "as is" clause but barring a seller from shifting risk to a buyer or enforcing an indemnity clause if contamination is unknown at time of purchase).

394. *See, e.g.*, Klein, *supra* note 12, at 366 (by encouraging seller disclosure and buyer investigative diligence, "[t]he rule of caveat emptor, therefore, encourages parties to reach a true meeting of the minds in real estate transactions."); Besser, *supra* note 393, at 210-11 (purchaser of contaminated property not a "mere bystander" injured by circumstances outside his control; rather, "[t]he purchaser can and, in this environmentally conscious era, should inspect the property or subject it to environmental investigations.").

395. *See, e.g.*, Besser, *supra* note 393, at 215-16 (changing common law allocation of risk between successive owners is better left to legislatures); Klein, *supra* note 12, at 365 ("using nuisance law to shift cleanup costs to intermediate landowners will cause such entities to expend resources in unproductive litigation-related activities.").

396. The *caveat emptor* risk allocation line is not always as bright as its proponents might suggest. For example, in some states risks associated with undisclosed site conditions do not shift until the buyer has had a "reasonable opportunity" to discover undisclosed conditions. *See, e.g.*, New York Tel. Co. v. Mobil Oil Corp., 473 N.Y.S.2d 172, 174 (N.Y. App. Div. 1984) (under New York law, the owner of land ceases to be liable in tort after the conveyance at such time as the new owner has had a reasonable opportunity to discover the condition by making prompt inspection and necessary repairs).

nated property to avoid common law remediation liability as a matter of law. The doctrine's underlying economic theory—that the market should allocate risks associated with real property conditions—assumes that the buyer (and, for that matter, the seller) has the technical and financial capacity to discover site contamination and understand its regulatory, economic and health significance during the narrow window of opportunity available for inspection of commercial or residential real estate. This assumption is unrealistic. Until relatively recently, few buyers fully understood the consequences of soil or groundwater contamination. A properly conducted invasive (e.g., Phase II) site inspection can be expensive and, without information about past waste disposal practices to guide investigators, potentially fruitless. Moreover, discovery of contamination does not necessarily mean that all potential environmental problems at the site have been identified much less understood. A site assumed to present one set of risks based on data collected years ago may prove to present additional concerns in light of newly developed contaminant detection techniques.[397] It is not realistic to assume that the market can efficiently or fairly allocate responsibility for such unknown and unknowable risks.

Second, applying the doctrine is likely to have little or no material impact on expanding the scope of seller disclosures. In many jurisdictions, the scope and content of real estate disclosure obligations, particularly those involving the use or disposal of hazardous substance on the property, are now often mandated by statute or regulation.[398] In addition, under the proposed paradigm a seller would not be tempted by a prospective *caveat emptor* liability defense to minimize disclosures about past property use and avoid scaring away potential buyers. Similarly, *caveat emptor* is likely to have little effect on whether a buyer diligently investigates the property in question. To the contrary, the scope of a potential buyer's site investigation likely would be determined by the buyer's interest in identifying any potential risk of assuming the environmental liabilities and regulatory requirements that arise from owning contaminated property.

---

397. *See, e.g.*, RANDALL L. ERICKSON & ROBERT D. MORRISON, ENVIRONMENTAL REPORTS AND REMEDIATION PLANS: FORENSIC AND LEGAL REVIEW § 8.3 (1995) (discussing risks of misidentification of compounds as a result of analytical testing method selected and how high detection limits can mask the presence of a compound in a soil or groundwater sample).

398. *See, e.g.*, MD. CODE ANN., REAL PROP. § 10-702(e)(2)(vii) (West 2006) (mandatory disclosure form shall include a list of defects in relation to hazardous or regulated materials); CAL. HEALTH & SAFETY CODE § 25359.7 (West 2006) (mandatory disclosure by owners to buyers, lessees or renters of real property in writing of release of any hazardous substance that has "come to be located on or beneath that real property"); NEB. REV. STAT. ANN. § 76-2, 120(4)(g) (West 2006) (mandatory disclosure of "[a]ny hazardous conditions, including substances, materials, and products on the real property which may be an environmental hazard"); *see also* N.J. STAT. ANN. 13:1K-9 (West 2006) (absent alternative contractual arrangements between buyer and seller of industrial property, seller must provide buyer with a no further action letter from state environmental regulatory agency or an approved remediation plan to be funded by seller).

Third, litigation expenses are not "wasteful transaction costs" when they serve to equitably reallocate the costs of cleaning up contamination to the party creating it. The potential for same property private nuisance litigation could further encourage pre-dispute agreements expressly allocating cleanup cost risks. Indeed, the threat or prosecution of litigation also can encourage post-dispute resolution that efficiently allocates remediation and cleanup cost responsibilities between vendor and vendee.

Finally, market forces can still efficiently allocate risk through contract. A seller can (and should) be able to avoid liability to a buyer for site contamination cleanup costs if the buyer contractually agrees to accept such responsibility.[399] A contractual allocation of responsibility by the buyer, however, must reflect a clear assumption of a known or knowable risk, in contrast (particularly in connection with agreements that predate CERCLA) to a generic "as is" clause unaccompanied by disclosures about site hazardous substance use or waste disposal history.[400] Accordingly, without *caveat emptor*, a seller seeking to avoid site contamination liability would have the additional incentive to fully disclose the site's hazardous substance history in order to obtain an enforceable contractual release of liability from the buyer.[401]

Just as many courts have now rejected "coming to the nuisance" as a defense to adjoining property dispute nuisance liability,[402] they should

399. *See, e.g.*, Mardan Corp. v. C.G.C. Music, Ltd., 804 F.2d 1454, 1461 (9th Cir. 1986) (buyer barred from asserting CERCLA claims against seller pursuant to prior agreement resolving and releasing claims relating to sale of property and business).

400. *See, e.g.*, Ybarra, *supra* note 243, at 1216-19 (arguing for a "standard of clear awareness of the specific contamination by the purchaser before rights against the responsible party (seller) are relinquished" and distinguishing among cases involving release agreements in recent transfers of property, older transfers of property, "as is" agreements, and purchaser knowledge of site contamination).

401. Similarly, because a seller's agreement to indemnify a buyer regarding hazardous substance liabilities can be enforceable, *see, e.g.*, 42 U.S.C.A. §9607(e)(3) (West 2006) ("[n]othing in this subsection [of CERCLA] shall bar any agreement to insure, hold harmless, or indemnify a party to such agreement for any liability under this section"), buyers have additional incentive to conduct thorough pre-acquisition site investigations to determine whether contamination issues may be present that warrant negotiation of an indemnification agreement.

402. *See* 1 WILLIAM H. RODGERS, JR., ENVIRONMENTAL LAW: AIR AND WATER § 2:9 (1986 & Supp. 2006) ("The priority in time of conflicting uses is pertinent to the question of whether a nuisance is proven. The issue arises typically as a defense to a nuisance claim on the theory that plaintiff has 'come to the nuisance.'"); RESTATEMENT (SECOND) OF TORTS § 840D, cmt. b (1979) ("The rule generally accepted by the courts is that in itself and without other factors, the "coming to the nuisance" will not bar the plaintiff's recovery. Otherwise the defendant by setting up an activity or a condition that results in the nuisance could condemn all the land in his vicinity to a servitude without paying any compensation, and so could arrogate to himself a good deal of the value of the adjoining land. The defendant is required to contemplate and expect the possibility that the adjoining land may be settled, sold or otherwise transferred and that a condition originally harmless may result in an actionable nuisance when there is later development."); *cf., e.g.*, Spur Indus., Inc. v. Del E. Webb Dev. Co., 494 P.2d 700, 706, 708 (Ariz. 1972) (granting private nuisance injunction requested by developer who purchased property located near cattle feeding operation on condition that developer indemnify cattle feeder for the reasonable cost of moving or shutting down); Jerry Harmon Motors, Inc. v. Farmers Union Grain Terminal Ass'n, 337 N.W.2d 427, 432 (N.D. 1983) (plaintiff that comes to an alleged nuisance has heavy burden to establish liability); Green, *supra* note 166, at 583-84

also turn away from *caveat emptor* as a liability defense to private nuisance claims.[403] Nevertheless, the circumstances surrounding a plaintiff/current landowner's acquisition of contaminated property should remain relevant to the nature and scope of any nuisance remedy. A wide range of transactional factors could impact the shaping of private nuisance relief, including: (a) a current landowner's knowledge about past on-site waste disposal practices before acquiring the property; (b) contractual language relating to but not directly addressing remediation cost risk allocation; (c) purchase price reduction or other consideration adjustment related to site contamination; and (d) post-transactional behavior by the current owner (e.g., contributing to site contamination, failing to promptly and effectively address discovered contamination).

These factors should not bar liability for a defendant who created contamination conditions constituting a nuisance. They should, however, be taken into account at the remedy stage of private nuisance litigation. For example, a court may consider such factors in deciding whether to grant an injunction,[404] fashioning the terms under which injunctive relief should be granted,[405] calculating damages to avoid any "double recovery" by a current owner who received a purchase price discount because of known or suspected contamination problems,[406] or evaluating the extent to which cleanup cost or other contamination damages should be borne by plaintiff rather than defendant under a "comparative respon-

---

(arguing that coming to the nuisance should be an affirmative defense if pollution costs have become capitalized into surrounding land values).

403. *See* T & E Indus., Inc. v. Safety Light Corp., 587 A.2d 1249, 1258 (N.J. 1991) (rejecting application of *caveat emptor* to same property ultra-hazardous activity contamination claim, observing that "the rule of *caveat emptor* has not retained its original vitality. With time, and in differing contexts, we have on many occasions questioned the justification for the rule."). Courts have turned away from *caveat emptor* as a matter of public policy in a variety of contexts, such as claims by the buyer of a new home against the developer-seller. *See, e.g.*, Hanlin Group, Inc. v. Int'l Minerals & Chem. Corp., 759 F. Supp. 925, 932-33 (D. Me. 1990) (applying Maine law and holding that *caveat emptor* did not bar private nuisance claim and noting that Maine courts previously had declined to apply the defense in cases involving the sale of new homes); 17 SAMUEL WILLISTON & RICHARD A. LORD, WILLISTON ON CONTRACTS § 50:26, at 352-53 (4th ed. 2000) ("Over the years, however, the number of cases which have strictly applied the rule of caveat emptor appears to be diminishing, and there is a distinct tendency to depart from the rule, either by way of interpretation, or exception, or by simply refusing to adhere to the rule where it would work injustice."). Sound public policy similarly requires that courts reject *caveat emptor* as a liability defense in same property private nuisance cases involving soil or groundwater contamination.

404. *Cf.* Kellogg v. Village of Viola, 227 N.W.2d 55, 58 (Wis. 1975) ("While coming to the nuisance may properly be considered while weighing the equities in an abatement action, it is irrelevant in a damage suit.").

405. *See, e.g.*, Spur Indus., Inc. v. Del E. Webb Dev., Co., 494 P.2d 700, 706, 708 (Ariz. 1972) (granting private nuisance injunction requested by developer who purchased property located near cattle feeding operation on condition that developer indemnify cattle feeder for the reasonable cost of moving or shutting down).

406. *See, e.g.*, W. Props. Serv. Corp. v. Shell Oil Co., 358 F.3d 678, 691 (9th Cir. 2004) (buyer aware of environmental hazards as evidenced by reduced purchase price precluded from recovering entire cost of cleanup).

sibility" approach to environmental nuisance damages.[407] Economics and market forces thus should continue play a role in private nuisance disputes. That role, however, should be at the remedy rather than the liability phase of the case.

D. A Different Approach to Continuing Environmental Nuisance

Whether a nuisance is characterized as continuing or permanent implicates several critical issues in private environmental nuisance cases: (a) identifying when a cause of action accrues and the applicable statute of limitations begins to run; (b) determining the appropriate nature and scope of available legal and equitable relief; and (c) resolving whether a plaintiff may bring successive actions to address future harm and conduct.[408] The current status of private nuisance law is a hopeless jumble largely because of the widespread inconsistency among (and sometimes within) states regarding the definition of a continuing nuisance and its application to soil and groundwater contamination cases. This inconsistency must be eliminated so that private nuisance law can become a meaningful alternative to a federal rule of decision in private cleanup cost disputes across the country.

To begin with, under the proposed paradigm every state should recognize a tort of continuing environmental nuisance. The concept of continuing tort may seem inconsistent with the policies of finality, predictability and encouraging prompt prosecution of claims that underlie statutes of repose. Closer examination, however, reveals that a continuing nuisance is not an open-ended exception to the statute of limitations; it should be viewed as a tort of wrongful inaction—a defendant's failure to abate a condition that continues to harm plaintiff and the environment. Viewed this way, defendant's ongoing refusal to abate continues to accrue a new cause of action based on defendant's ongoing inaction. Moreover, soil and groundwater contamination presents statute of limitations challenges significantly different in degree (if not in kind) from those presented by other torts. For example, a property owner may have noticed discolored soil decades ago but not understood its health, environmental or regulatory significance. Similarly, recent changes in technology (e.g., development of laboratory equipment capable of identifying increasingly smaller concentrations of contaminants)[409] may provide a far more complete understanding of the harm and regulatory obligations triggered by contamination caused by the defendant. Without a continu-

407. For a thorough discussion of "comparative nuisance" and the equitable allocation of cleanup costs based on the comparative responsibility of the parties under a nuisance cause of action, see generally Jeff L. Lewin, *Comparative Nuisance*, 50 U. PITT. L. REV. 1009 (1989).

408. *See, e.g.*, Schneider Nat'l Carriers, Inc. v. Bates, 147 S.W.3d 264, 275-76 (Tex. 2004) (discussing issues implicated by nuisance characterization).

409. ERICKSON & MORRISON, *supra* note 397, at 161 (discussing risks of misidentification or non-identification of compounds because of analytical testing methods and laboratory equipment detection limitations).

ing tort doctrine, a current owner who only recently learned the significance of older contamination may find a state law cleanup cost claim (and with it the incentive voluntarily to remediate contamination caused by others) time-barred by applicable statutes of limitations, even under a discovery rule.[410]

1. Defining Continuing Nuisance—Reasonable Abatability

The definition of "continuing nuisance" in contaminated property cases should reflect the circumstances of the underlying environmental problem. The definition of a "continuing nuisance" should arise from the defendant's failure to remediate the soil and groundwater contamination she caused or to which she contributed. Viewed this way, the defendant is responsible for the continued contamination even after the cessation of her conduct that contaminated the environment (e.g., disposal activity, installation and maintenance of a leaking underground storage tank). It is the unabated contamination, not the acts creating it, that constitutes the continuing interference with the plaintiff's use and enjoyment of property.[411] The defendant thus should have an ongoing duty to abate the contamination; the repeated failure to perform this duty constitutes a separate cause of action that continues to accrue until the defendant performs its duty or the contamination is otherwise abated.

An ongoing duty to abate through remediation assumes that remediation is practicable. Under the proposed paradigm, the definition of a continuing nuisance requires that the contamination reasonably be capable of remediation. "Reasonable abatability" contemplates that (a) the contamination could be abated to an appropriate level; (b) abatement to such a level is technically possible; and (c) the abatement could be accomplished at a reasonable cost.[412]

a. Abatement to an Appropriate Level

An appropriate abatement level could be determined in one of two ways. First, site-specific regulatory agency cleanup levels could set a presumptive level of abatement.[413] The California Court of Appeal accepted such a role for agency approved cleanup levels in *Capogeannis v.*

410. *See supra* notes 248-49 and accompanying text.

411. *See* Mangini v. Aerojet-Gen. Corp. (*Mangini I*), 281 Cal. Rptr. 827, 838 (Cal. Ct. App. 1991) (continuing unabated harm is basis for continuing nuisance claim under California law); Capogeannis v. Super. Ct., 15 Cal. Rptr. 2d 796, 804 (Cal. Ct. App. 1993) ("[T]he continuing nature of nuisance refers to the continuing but abatable damage caused by the offensive condition, not the acts causing the offensive condition to occur." (citation omitted)).

412. In Mangini v. Aerojet-Gen. Corp. (*Mangini III*), 912 P.2d 1220, 1229 (Cal. 1996), the Supreme Court of California embraced a "reasonableness" limitation on an abatability standard for continuing nuisance, concluding that a nuisance was continuing for limitations purposes if it "can be remedied at a reasonable cost by reasonable means." *See supra* notes 267-73 and accompanying text.

413. *See supra* note 271 and accompanying text.

Superior Court.[414] The court rejected the defendants' "essentially semantic argument that because it does not appear that the contamination [leakage from a removed underground storage tank] can ever be *wholly* removed the nuisance must be deemed permanent."[415] Instead, the court was "satisfied to presume that cleanup standards set by responsible public agencies sufficiently reflect expert appraisal of the best that can be done to abate contamination in particular cases. As judges we will not presume to insist on absolutes these agencies do not require."[416] Similarly, the court noted that because "as a practical matter the only impact of the contamination on the Capogeannises is by way of the regulatory agencies demands,"[417] once those "demands have been met, so far as the Capogeannises are concerned the nuisance will be abated."[418]

Second, expert testimony could be used to address site cleanup standards. The characterization of a nuisance as permanent or continuing may come before the court at trial or by summary judgment motion before a regulatory agency has set cleanup standards for the site.[419] If so, the court will require expert evidence regarding site conditions to evaluate the abatability of site contamination. Similarly, a party may turn to expert testimony to rebut any presumption that agency cleanup standards demonstrate that site contamination is abatable.[420]

b. Reasonable Means of Abatement

The concept of "reasonable abatability" should include a two-fold technical component. First, means must exist to accomplish the contemplated remediation. Merely because an agency (or an expert witness) has proposed a cleanup standard does not mean that the standard can be met with current technology. For example, a proposed standard may contemplate reduction of groundwater contamination pollution levels to a specified level of X parts per billion (ppb) for a particular compound. It may be impossible, however, to meet the X ppb standard: the subsurface source of groundwater contamination may not have been lo-

414. 15 Cal. Rptr. 2d 796 (Cal. Ct. App. 1993).
415. *Capogeannis*, 15 Cal. Rptr. 2d at 805.
416. *Id.*
417. *Id.*
418. *Id.*
419. There may be many reasons why a contamination dispute could reach such advanced stages of litigation without establishment of agency cleanup levels. For example, it may take years to fully characterize large or technically complicated sites, and substantial time after completion of site characterization activities to determine site cleanup levels. Moreover, limited resources can substantially delay an environmental regulatory agency's determination of cleanup standards for sites of any size or complexity.
420. In *Mangini III*, the California Supreme Court held that the contamination of plaintiff's property constituted a permanent nuisance in part because there was no evidence of agency cleanup standards or expert testimony proposing appropriate cleanup levels for the site. *Mangini III*, 912 P.2d at 1221.

cated;[421] the technology may not yet exist to clean up the contaminated groundwater to the desired level; or laboratory equipment or available analytical methods may not be capable of determining whether the cleanup standards has been met.[422]

Second, even if the means exist to remediate the site contamination, other environmental regulatory factors may argue against remediation. For example, in *Beck Development Co. v. Southern Pacific Transportation Co.*,[423] the court concluded that excavating oil-contaminated soil, while technically possible, might not be reasonable in light of a regulatory agency's concern that the contamination presented no risk if left in place but would raise other environmental concerns if it was excavated and transported away from plaintiff's property for off-site disposal.[424]

c. Reasonable Cost of Abatement

For contamination to be deemed "abatable," the remediation must be possible at a reasonable cost. The essence of private nuisance is interference with the plaintiff's use and enjoyment of property.[425] Whether the cost of abatement is reasonable, therefore, should be viewed from the plaintiff's perspective. In other words, abatement costs should be balanced not against the value of the property as contaminated or, for that matter, the value of the property after cleanup. Instead, it should be balanced against the value of the remediated property *to the plaintiff* after the cleanup, including the present value of profits that could be generated from the remediated property. As a result, if a parcel of contaminated property has no net value as contaminated, a fair market value of $5,000,000 as remediated, and the capacity to generate an additional $5,000,000 in profit if remediated (e.g., the present value of profits generated from developing the site for single family residences), the baseline for measuring the reasonableness of remediation should be $10,000,000, the overall value of the highest and best use of the property as uncon-

421. For example, at some sites the presence of dense non-aqueous phase liquids (DNAPLs)—in effect, a volumetrically small but high-concentration collection of the contaminant in a subsurface water-bearing soil zone—can serve as a continuing source of groundwater contamination. *See* ERICKSON & MORRISON, *supra* note 397, §7.4 (describing how, as groundwater flows past the DNAPL, a portion of the DNAPL "will dissolve into the groundwater" at a rate dependent on groundwater velocity and DNAPL solubility). Unless the DNAPL can be located and its effect eliminated or significantly reduced, near-source concentrations of groundwater contamination may not be capable of remediation to proposed cleanup levels.

422. *See supra* note 396 and accompanying text.

423. 52 Cal. Rptr. 2d 518 (Cal. Ct. App. 1996).

424. *Beck Dev. Co.*, 52 Cal. Rptr. 2d at 559-60; *see also* 40 C.F.R. § 300.430(f)(ii)(E) (2006) (NCP requires that selection of remedial action include consideration of "the preference for treatment as a principal element and the bias against off-site land disposal of untreated waste."); JENNIFER L. MACHLIN & TOMME R. YOUNG, MANAGING ENVIRONMENTAL RISK: REAL ESTATE AND BUSINESS TRANSACTIONS § 4-103 (2004) (noting NCP preference in connection with preparation of RI/FS for on-site treatment and cost-effectiveness).

425. RESTATEMENT (SECOND) OF TORTS § 821D (1979) (defining private nuisance as "a nontrespassory invasion of another's interest in the private use and enjoyment of land").

taminated. Viewed this way, a $7,000,000 cleanup would represent abatement at a reasonable cost and thus a continuing nuisance.[426]

Why should cost matter at all? Why should a defendant be "rewarded" with a statute of limitations defense when she causes technically abatable harm that is so extensive that the cost of abating it exceeds the value of cleanup to plaintiff? Cost should matter because at some point the harm is so great that the law should no longer permit a private plaintiff to demand that the defendant abate the nuisance in a direct action premised on interference with the use and enjoyment of property. As comment f to section 839 of the Restatement (Second) of Torts provides, "[t]he law does not require the unreasonable or the fantastic, and therefore even though it might conceivably be possible to abate a particular condition, it is not 'abatable' within the meaning of this Section unless its abatement can be accomplished without unreasonable hardship or expense."[427] Moreover, factoring costs to determine whether a nuisance is continuing or permanent reflects a policy choice that preserves a role for statutes of limitations. Without cost as a factor, almost any technically abatable contamination would support a claim for continuing nuisance, all but eliminating the statute of limitations as a defense in environmental nuisance cases. An analysis into the means and cost of abatement would involve a case-by-case analysis regarding site conditions requiring the expenditure of additional party and judicial resources. These resources likely would be expended anyway, however, in fashioning a nuisance remedy,[428] examining the reasonableness of abatement

426. The highest and best use of the property, viewed objectively, would take into account both plaintiff's planned use and any other, more profitable use that could be employed by plaintiff or a future purchaser (who presumably would pay plaintiff for the potential development opportunity). A court may also consider the reason why plaintiff seeks to recover from defendant abatement costs exceeding the value of the property to plaintiff. On the one hand, plaintiff may be required by a regulatory agency to conduct such a cleanup. On the other hand, the plaintiff may choose to remediate for purely business reasons. *See, e.g.*, *Beck Dev. Co.*, 52 Cal. Rptr. 2d at 559 (evaluating plaintiff's private nuisance claim for remediation of a development site in light of regional water board conclusion that contaminated soil could have remained in place). In the former situation, a court may take into account whether the plaintiff has any choice in remediating defendant's pollution as part of a "reasonableness" analysis; in the latter situation, the economics of remediation may assume determinative significance regarding whether the nuisance is permanent or continuing.

427. RESTATEMENT (SECOND) OF TORTS, § 839 (1979) (addressing liability of possessor of land who fails to abate an artificial condition constituting a nuisance). In addition to evidence of technical abatability, courts may also consider the selection of site-specific cleanup levels as relevant to whether abatement can be achieved at a reasonable cost, at least in those instances where the agency is required to take cost-effectiveness into account in setting cleanup levels. *See, e.g.*, 40 C.F.R. § 300.430(f)(ii)(D) (requiring lead agency under NCP to take cost-effectiveness into account in selecting remedial action); *see also supra* note 271 and accompanying text.

428. It is possible that a plaintiff at an older contamination site will be required by a regulatory agency to cleanup contamination of her property to which she made little or no contribution at a cost that exceeds the value of the remediated property to plaintiff. A court may take into account any regulatory directives imposed on plaintiff in deciding whether the cost is reasonable and thus whether the nuisance is continuing. Moreover, in contrast to a direct private nuisance claim based on plaintiff's interest in freedom from a substantial interference with the use and enjoyment of her property, derivative claims (e.g., if state law permits actions for contribution or equitable indemnity upon issuance of a cleanup order or in the event of enforcement action) may be available against the defendant with regard to cleanup costs arising from such a duty imposed on the current owner.

costs sought as damages by the plaintiff,[429] allocating cleanup costs based on a comparative responsibility analysis, and the crafting of abatement injunctive relief.

d. Comparison with Other Continuing Nuisance Standards

The "reasonable abatability" standard rests on a defendant's inaction in the face of a reasonably abatable harm that continues to interfere with a plaintiff's use and enjoyment of property. It also would preserve cleanup cost claims at many older contamination sites at which the statute of limitations would bar claims for permanent nuisance. The reasonable abatability standard is superior to other continuing nuisance standards currently used by courts around the country, which fail to adequately serve the policy considerations that support preservation of cleanup cost claims at older contamination sites.

The continuing harmful conduct standard[430] does nothing to address the problem of a defendant's inaction in the face of the reasonably abatable harmful condition that remains on a plaintiff's property. Moreover, although the discontinuance of a defendant's conduct creates something of a bright line to determine when the permanent nuisance statute of limitations begins to run, it provides no relief at older contamination sites where a defendant's polluting conduct may have ceased years ago but the offensive condition it created remains unabated.

The predictability of future harm standard employed by the Texas courts[431] looks away from both a defendant's conduct and the abatability of pollution, focusing instead on the predictability of future damages. In effect, the predictability of harm standard constitutes a presumption that a nuisance is permanent by limiting "temporary" nuisance actions to those where future damages cannot reasonably be determined and thus a single action would be unable to fully compensate the plaintiff. It creates no incentives for abatement; indeed, the Texas standard creates a perverse incentive for a plaintiff to argue that the nature of a contamination problem cannot yet be characterized and for a defendant to argue that because the problem can be abated by reasonable means and at a cost capable of a current reliable estimate the contamination must constitute a permanent nuisance and plaintiff's claim should be time-barred.

Other types of property damages, e.g., lost rental or other use of the property, would presumably not be available under a derivative theory of liability.

429. *See, e.g.*, RESTATEMENT (SECOND) OF TORTS § 929(1) (1979) (stating that "[i]f one is entitled to a judgment for harm to land resulting from a past invasion and not amounting to a total destruction of value, the damages include compensation for (a) the difference between the value of the land before the harm and the value after the harm, or at his election in an appropriate case, the cost of restoration that has been or may be reasonably incurred, (b) the loss of use of the land, and (c) discomfort and annoyance to him as an occupant.")

430. *See supra* notes 286-300 and accompanying text.

431. *See supra* notes 301-14 and accompanying text.

The courts that use a "changing over time" standard effectively equate a "continuing nuisance" with a "different-looking nuisance."[432] In other words, if the harmful condition changes over time, it is a different nuisance and therefore a new cause of action should accrue. This approach ignores both the defendant's failure to abate harm and ongoing active causation of harm as the basis for a continuing nuisance analysis. It also draws arbitrary remedial lines, permitting a continuing nuisance claim for groundwater contamination (which changes form over time as groundwater continues to migrate) while relegating to permanent nuisance status discrete, easily removed soil contamination (which generally remains immobile and does not change form or environmental impact over time).

The continued presence standard[433] comes closest to serving the goals achieved by the proposed reasonable abatability standard. On its face, however, the continued presence standard does not address either the technical means of abatement or its cost, transforming all soil and groundwater contamination that does not serve a socially beneficial purpose by remaining in place into a continuing nuisance. Such a broad standard would support a nuisance-based cleanup cost claim in almost all contaminated property cases but would virtually eliminate the statute of limitations as a defense at these sites and in theory permit perpetual successive continuing nuisance actions concerning contamination that cannot practicably be abated.[434] If soil or groundwater contamination is not practicably abatable, it properly should be viewed as a permanent problem for which the rules of permanent nuisance should apply.

Finally, the multi-factored balancing approach[435] has the benefit of flexibility. It also has the significant disadvantages of unpredictability and unreliability regarding whether a given contamination problem constitutes a continuing nuisance. Moreover, it shifts the focus of continuing nuisance law from abating contamination to a case-by-case bundling of issues to be weighed in a potentially inconsistent manner from court to court.

In sum, states should re-examine the law of continuing nuisance as applied to soil and groundwater contamination problems by viewing the definition of continuing nuisance through the lens of the public policy that this tort should serve. The proposed paradigm promotes prompt, efficient cleanups by preserving cleanup cost claims based on a defen-

---

432. *See supra* notes 315-24 and accompanying text.

433. *See supra* notes 274-85 and accompanying text.

434. Even if a state court found the continued presence test more attractive than a reasonable abatability approach, a court might decline to embrace a standard that would all but eliminate the statute of limitations (even with a "social utility" exception) as a liability defense and instead defer such a policy choice to the state legislature.

435. *See supra* notes 325-28 and accompanying text.

dant's continued failure to reasonably abate the soil or groundwater contamination she caused or to which she contributed.

2. Rebuttable Presumption of Continuing Nuisance

The proposed paradigm is designed to encourage (a) abatement of soil and groundwater contamination, (b) voluntarily cleanup by current property owners of contamination caused in whole or in part by others, and (c) informal resolution of private cleanup cost allocation disputes among PRPs. The default characterization of a nuisance in the event of evidentiary equipoise can play a substantial role in serving these goals. Accordingly, the proposed paradigm would create a rebuttable presumption[436] that soil or groundwater contamination constitutes a continuing rather than permanent nuisance.[437]

An environmental nuisance framework that presumes soil and groundwater contamination constitute a continuing nuisance would provide a number of benefits. First, abatement-related private nuisance claims (e.g., claims for cleanup cost damages or injunctive relief ordering defendant to abate a nuisance) presumptively would not be time-barred.

Second, equitable relief directed to abating the nuisance presumptively would be available. Under a reasonable abatability standard for

436. Some courts have articulated a preference for finding a continuing nuisance to protect plaintiffs from the statute of limitations, unforeseen future injuries, and to encourage the abatement of nuisances. *See* Capogeannis v. Super. Ct., 15 Cal. Rptr. 2d 796, 804-05 (1993) (holding that issues of material fact whether contaminants from former owner and tenant underground storage tank constituted a continuing or permanent nuisance barred summary judgment for defendants on statute of limitations). This preference, however, often manifests itself by allowing plaintiffs in close or doubtful cases to elect which nuisance theory to pursue. *Id.* at 801. This election, of course, is illusory in connection with older contamination problems where a permanent nuisance claim would be time-barred.

437. Not all soil or groundwater contamination, of course, constitutes a nuisance. *De minimus* concentrations of contaminants on a plaintiff's property, for example, might not constitute a nuisance. A plaintiff thus would still be required to prove that the contamination substantially interfered with the use and enjoyment of her property and that the defendant's negligent, intentional and unreasonable or ultra-hazardous activity conduct proximately caused the contamination. *See supra* note 174 and accompanying text; *see also* Jezowski v. City of Reno, 286 P.2d 257, 260-61 (Nev. 1955) (private nuisance claim requires proof of "material annoyance, inconvenience, discomfort or hurt"); Robie v. Lillis, 299 A.2d 155, 158 (N.H. 1972) (interference with use and enjoyment of property must be substantial); Energy Corp. v. O'Quinn, 286 S.E.2d 181, 182 (Va. 1982) (condition "substantially impairing the occupant's comfort, convenience, and enjoyment of the property" may constitute private nuisance); Brown & Hansen, *supra* note 224, at 664-65 (noting that "many California courts have utilized the common law balancing approach to determine whether a condition or activity constitutes a nuisance under the provisions of [California Civil Code] section 3479. Under this approach, the plaintiff must establish that the harm of a nuisance outweighs the benefits of the defendant's conduct. In addition, the plaintiff must demonstrate that the injury is substantial and not nominal. Whether a particular use of land constitutes a nuisance must be determined on a case-by-case basis in light of all the circumstances. The relevant balancing factors include the priority of the use, the locality and surroundings of the property, the nature and extent of the nuisance and the injury caused thereby, whether the nuisance is continual or occasional, and the number of people affected."). Assuming, however, that the presence of contaminants does constitute a nuisance, under the proposed paradigm it would presumptively constitute a continuing rather than permanent nuisance.

continuing nuisance, soil and groundwater contamination presumptively would be capable of abatement by reasonable means and at reasonable cost.[438] A plaintiff also would be entitled to recover damages consistent with the continuing (i.e., non-permanent) nature of the nuisance sustained during the limitations period (e.g., abatement costs, damages arising from the lost use of the property).[439] A defendant would typically

438. Sites that already are the subject of a regulatory agency cleanup order may present primary jurisdiction concerns. Under the doctrine of primary jurisdiction, a court has the discretion to refer certain matters to a specialized administrative agency. It is "a prudential doctrine under which courts may, under appropriate circumstances, determine that the initial decision-making responsibility should be performed by the relevant agency rather than the courts." Syntek Semiconductor Co. v. Microchip Tech., Inc., 307 F.3d 775, 780 (9th Cir. 2002). Courts may consider a variety of factors in deciding whether to exercise the doctrine, including whether (a) the court is being called on to decide factual issues which are not within the conventional experience of judges or are instead issues of a sort that a court routinely considers; (b) the defendant could be subjected to conflicting orders of both the court and the agency; (c) relevant agency proceedings have been initiated; (d) the agency has proceeded diligently or allowed the proceedings to languish; and (e) the plaintiff is seeking injunctive relief requiring technical or scientific expertise for the courts to craft. Schwartzman, Inc. v. Atchison, Topeka & Santa Fe Ry. Co., 857 F. Supp. 838, 841-43 (D.N.M. 1994) (applying primary jurisdiction doctrine to stay landowner's claim for injunctive relief where EPA had already begun process of initiating remedial investigation and feasibility study). Some courts have deferred consideration of injunctive relief under the primary jurisdiction doctrine where an environmental regulatory agency actively was involved in site characterization or remediation, *see id.*; Liss v. Milford Partners, 39 Conn. L. Rptr. 216 (Conn. Super. Ct. 2005) (unpublished) (denying motion to dismiss claim seeking injunction under Connecticut law that defendants remediate soil and groundwater contamination but postponing further judicial proceedings pending completion of administrative action by state Department of Environmental Protection); White Lake Improvement Ass'n. v. City of Whitehall, 177 N.W.2d 473, 485 (Mich. Ct. App. 1970) (affirming dismissal of complaint seeking injunction to abate nuisance caused by discharge of waste into lake on ground that plaintiff should pursue administrative remedies through state water agency before court should further entertain action to abate nuisance), while others have rejected the primary jurisdiction argument and permitted claims to abate a nuisance, *see, e.g.*, Attorney Gen. of Mich. v. Thomas Solvent Co., 380 N.W.2d 53, 67-68 (Mich. Ct. App. 1986) (doctrine of primary jurisdiction did not require court to defer jurisdiction over contaminated site to EPA in light of limited agency resources, lack of emergency situation, and absence of agency objection); Meinders v. Johnson, 134 P.3d 858, 867 (Okla. Civ. App. 2005) (affirming injunction that defendants remedy surface and subsurface pollution from mineral exploration and production on plaintiff's property where Corporation Commission had not yet exercised jurisdiction over matter).

439. *See supra* note 252 and accompanying text. Some courts have treated a nuisance as continuing for statute of limitations purposes only but nevertheless awarded "permanent" nuisance (i.e., prospective) damages. *See, e.g.*, Cook v. Rockwell Int'l Corp., 358 F. Supp. 2d 1003, 1012 (D. Colo. 2004) (presence on plaintiffs' property of plutonium released from nearby nuclear weapons plant constituted a continuing nuisance and trespass for statute of limitations purposes but could provide basis for prospective, diminution in value damages effectively purchasing an easement for invasion to continue); Beatty v. Washington Metro. Area Transit Auth., 860 F.2d 1117, 1125 (D.C. Cir. 1988) ("[N]uisances may be classified for two distinct purposes, one for the assessment of damages, and the other for the application of the statute of limitations."); *cf.* Briscoe v. Harper Oil Co., 702 P.2d 33, 36-37 (Okla. 1985) (affirming temporary nuisance damage award for cost of repairing well and permanent nuisance damage award for unabatable diminution in farmland property value caused by oil and gas lessee). Such an inconsistent approach to continuing nuisance would all but eliminate the statute of limitations as a defense in environmental nuisance cases and fail to serve any unifying policy except maximizing plaintiff's financial recovery. Under the proposed paradigm, a presumption against prospective damages would keep the policy focus of continuing nuisance actions on remediation and cooperation. Whether stigma damages are more consistent with a permanent or continuing nuisance theory presents a challenging conceptual problem. (A detailed discussion of stigma damages is beyond the scope of this article.). On the one hand, diminution of property value caused by stigma of prior contamination that has been abated arguably is consistent with a continuing nuisance theory and thus could be recoverable. On the other hand, diminution in property value caused by contamination stigma could be viewed as prospective harm

bear sole responsibility for these damages at sites where the defendant was the sole cause of the contamination. At sites where the plaintiff had partial responsibility for the contamination,[440] damages (including the cost of complying with any injunction issued by the court) could be allocated proportionately.[441] By contrast, contamination that cannot reasonably be abated should not be enjoined[442] and would constitute a permanent nuisance.[443]

Third, plaintiffs presumptively would have a right to bring successive actions until the nuisance was abated. In each action, a plaintiff would be able to seek injunctive relief and damages sustained during the preceding limitations period. The right to bring successive actions protects a plaintiff against unforeseen future remediation costs and temporary harm occasioned by a defendant's continued failure to remediate the contamination (e.g., lost property use damages, abatement costs). It also provides powerful incentives for the parties to cooperate in promptly remediating the property. A plaintiff would want to get the site cleaned

---

and thus reserved only for permanent nuisance actions. For example, in *Santa Fe Partnership v. Arco Products Co.*, 54 Cal. Rptr. 2d 214, 224 (Cal. Ct. App. 1996), a California Court of Appeal felt constrained by California Supreme Court precedent holding that prospective damages are unavailable for a continuing nuisance action to hold that stigma damages were unavailable in a continuing nuisance case.

440. Even if a plaintiff did not contribute to site contamination, a court may consider as part of a comparative responsibility analysis whether the plaintiff acted diligently to mitigate harm upon discovery of the contamination or its significance, or instead allowed site conditions to worsen and abatement costs to increase through inattention or indifference.

441. *See generally* Lewin, *supra* note 407, at 1053-69 (advocating nuisance remedial scheme contemplating proportional allocation of nuisance damages and cost of performing abatement injunction based on the comparative responsibility of plaintiff and defendant for creating or maintaining the nuisance). Courts also could freely experiment with alternative remedial schemes consistent with the fundamental nature of the continuing nuisance analytical framework. *See, e.g.*, RICHARD A. POSNER, ECONOMIC ANALYSIS OF LAW 65 & n.8 (4th ed. Aspen 1992) (citing William F. Baxter & Lillian R. Altree, *Legal Aspects of Airport Noise*, 15 J. L. & ECON. 1 (1972), as basis for proposal, in context of noise pollution, of "time-limited easements," limiting successive continuing tort actions to periodic suits (e.g., every 10 years) in order to limit administrative costs and provide extended platform for bargaining in exchange for payment by defendant of prospective damages covering the temporary pollution easement period). For an extensive discussion regarding the evolution of nuisance remedies, see generally Lewin, *supra* note 166.

442. *See, e.g.*, Spaulding v. Cameron, 239 P.2d 625, 629 (Cal. 1952) (directing the trial court to determine whether a landslide onto the plaintiff's property constituted a permanent nuisance; if so, the court should award diminution in property damages and if not, the court should award injunctive relief to abate the nuisance).

443. By filing a timely claim for permanent nuisance, a plaintiff could recover all past, present, and future damages caused by the contamination, in effect representing the price of a permanent easement for the pollution on plaintiff's property. *Cf.* Cook v. Rockwell Int'l Corp., 358 F. Supp. 2d 1003, 1012 (D. Colo. 2004) (presence on plaintiffs' property of plutonium released from nearby nuclear weapons plant constituted a continuing nuisance and trespass for statute of limitations purposes, and could provide basis for prospective, diminution in value damages effectively purchasing an easement for invasion to continue). If the permanent nuisance claim is not timely filed, plaintiff could not recover damages because of the statute of limitations and "because the nuisance is deemed permanent, the plaintiff may not abate the nuisance by an injunction; rather she is limited to her time-barred claim for damages. In effect, defendant has acquired a right to damage, and continue to damage, plaintiff's land at no cost to defendant!" FISCHER, *supra* note 251, §84(d). The presumption of continuing nuisance created by the proposed paradigm would reduce the risk of the latter result.

up promptly to restore the value of her asset and satisfy regulatory obligations. A defendant (who likely would be precluded from re-litigating her liability in a subsequent action) would want to mitigate delay or lost use damages arising from the continued presence of contaminants on plaintiff's property. A defendant thus would be encouraged to actively participate in site remediation to promote a cost-effective cleanup (which the defendant is paying for, in whole or in part) and to avoid an argument in a subsequent action that the defendant should pay punitive damages for intentionally failing to abate a nuisance for which it has been previously found liable.[444]

In the alternative, a defendant who is found liable for continuing nuisance and subject to successive suits until the nuisance is abated has an incentive to negotiate an informal resolution of remediation and cleanup cost allocation issues.[445] For example, the defendant could enter into a settlement agreement establishing a contract-based set of duties to directly conduct or financially participate in the cleanup and avoid future litigation transaction costs. Similarly, the defendant may attempt to negotiate a settlement by paying the plaintiff to liquidate the defendant's continued obligation to remediate the contamination.

The presumption that soil or groundwater contamination constitutes a continuing nuisance would be rebuttable. Accordingly, whether a nuisance ultimately was found to be continuing or permanent would turn on the intersection of three components of the paradigm: (a) the standard (reasonable abatability) by which a court would determine whether the objecting party satisfied its burden; (b) the default presumption that the nuisance is continuing; and (c) the placement of the burden of proof on any party *objecting* to characterization of the nuisance as continuing, i.e., any party contending that the nuisance is permanent. It is to this shifted burden of proof that we now turn.

E. Burden of Proof on the Party Advocating for Permanent Nuisance

The presumption that soil or groundwater contamination constitutes a continuing nuisance is effected by the allocation of the burden of proof. If the evidence regarding the reasonable abatability of contamination conditions is in equipoise, under the proposed presumption the nuisance would be deemed continuing. This presumption could only be overcome if a party—whether defendant *or plaintiff*—who argues that a nuisance is

444. *See, e.g.*, Sumitomo Corp. of Am. v. Deal, 569 S.E.2d 608, 615-16 (Ga. Ct. App. 2002) (affirming as consistent with due process $250,000 continuing nuisance and trespass punitive damage award against upstream developer who knew of damage caused by water leaving its detention ponds but did nothing to abate the flow of water).

445. *See* Capogeannis v. Super. Ct., 15 Cal. Rptr. 2d 796, 804 (Cal. Ct. App. 1993) (finding of continuing nuisance "will tend to encourage private abatement, and perhaps monetary cooperation in abatement efforts, if only to limit successive lawsuits").

permanent can demonstrate that the contamination likely cannot be abated by reasonable means and at a reasonable cost.[446]

Courts that have addressed the issue generally place the burden of proof on a plaintiff to show that a nuisance is continuing.[447] Because the question commonly arises only when a permanent nuisance claim would be time-barred, the burden of proof issue traditionally is framed as whether continuing nuisance characterization is an element of a plaintiff's cause of action[448] or part of a defendant's statute of limitations affirmative defense.[449] Nuisance characterization, however, affects not only the statute of limitations but the right to bring successive nuisance actions and the nature of available relief. The burden of proof, thus, should turn on the policy considerations and consequences arising from nuisance characterization.[450]

Characterizing contamination as a continuing nuisance would promote the paradigm's policy objectives to (a) expand the availability of current landowner common law rights to recover from other PRPs their fair share of cleanup costs, (b) encourage prompt characterization and remediation of contamination, and (c) promote settlement of private cleanup cost disputes. On the other hand, if a nuisance is deemed permanent, landowners would be unable to recover cleanup costs at many older contamination sites and defendants would have no obligation or incentive under common law to help remediate contamination that oc-

---

446. The party advocating for a permanent nuisance would bear both the burden of production (i.e., the obligation to produce sufficient evidence to avoid summary judgment or entry of judgment as a matter of law) and the burden of persuasion at trial.

447. *See, e.g.*, Mangini v. Aerojet-Gen. Corp. (*Mangini III*), 912 P.2d 1220, 1230 (Cal. 1996); *see also supra* notes 337-60 and accompanying text.

448. *See, e.g.*, *Mangini III*, 912 P.2d at 1226 (noting that the parties did not dispute that the characterization of nuisance was an element of plaintiff's cause of action); *Beck Dev. Co.*, 52 Cal. Rptr. 2d at 556 (finding a private nuisance claim time-barred because the plaintiff failed to meet its burden to prove continuing nuisance).

449. *See* Brown & Hansen, *supra* note 224, at 698-99 (arguing that defendants should bear burden of proving that contamination from an underground storage tank constitutes a permanent nuisance, rather than plaintiffs bearing the burden that the site constitutes a continuing nuisance, likening the issue to the defendant's burden of proving an affirmative defense that a claim is barred by the statute of limitations); *see also* Smith, *supra* note 164, at 60-61 (noting that, while plaintiffs typically bear burden of proving damages within continuing tort limitations period, the statute of limitations usually is an affirmative defense).

450. *See* Keyes v. School Dist. No. 1, Denver, Colo., 413 U.S. 189, 209 (1973) (finding of intentionally segregative school board actions shifted to school authorities the burden of showing that their actions as to other segregated schools within the system were not also motivated by segregative intent, observing that "[t]here are no hard-and-fast standards governing the allocation of the burden of proof in every situation. The issue, rather, 'is merely a question of policy and fairness based on experience in the different situations.'" (citation omitted)); David S. Cohen, *The Evidentiary Predicate For Affirmative Action After* Croson*: A Proposal for Shifting the Burdens of Proof*, 7 YALE L. & POL'Y REV. 489, 499 (1989) ("Courts, however, are not hesitant to reallocate both burdens in order to achieve the purposes of the underlying substantive law and to fulfill notions of good public policy."); Martin J. LaLonde, *Allocating the Burden of Proof to Effectuate The Preservation and Federalism Goals of the Coastal Zone Management Act*, 92 MICH. L. REV. 438, 449 (1993) ("Courts often do not hesitate to allocate the burden to realize the purposes of the substantive law and to promote public policy goals.").

curred long ago. Moreover, at sites where a permanent nuisance claim would not be time-barred, an award of prospective permanent nuisance damages would effectively purchase a pollution easement for defendant and undermine the current landowner's incentive to use a damage award to fully remediate her property, leaving the environment polluted and potentially useful property unavailable for fully productive use.[451]

Placing the burden of proof on the party advocating that contamination constitutes a permanent nuisance would eliminate a series of perverse incentives under current practice. First, defendants arguing that a nuisance claim should be time-barred because the contamination constitutes a permanent nuisance currently need only point to the absence of data regarding contamination abatability.[452] Under the proposed paradigm, defendants would need to proffer site characterization data if they wished to show that a nuisance cannot be abated by reasonable means and at reasonable cost.[453] Second, defendants would have an incentive to become actively involved in the site investigation and remediation process, either to control abatement costs for which they may be liable or to demonstrate that they were prepared to participate in nuisance abatement but were deterred from doing so by unreasonable plaintiff demands.[454] Third, plaintiffs seeking to pocket a one-time potential "windfall" of past, present, and future money damages for permanent nuisance rather than use their recovery to remediate their property would have to prove that their property is not subject to reasonable abatement. Plaintiffs would be deterred taking this position by the risks of harming prospects for future site development (e.g., the risk of creating by their own argu-

451. Where the statute of limitations defense is unavailable, the risk of a substantial damages award for permanent nuisance could lead a defendant to argue that a nuisance is continuing. The risk of such an award also could encourage greater care in hazardous substance handling and disposal practice.

452. *See supra* notes 338-56 and accompanying text.

453. The burden of proof regarding nuisance characterization also has a significant impact on litigation timing. A plaintiff bearing the burden of proving that a nuisance is continuing cannot proceed to trial until sufficient site condition data has been generated to meet its burden of production. As a result, a plaintiff may be forced to try to delay trial until the information becomes available. A defendant, on the other hand, currently has the incentive to accelerate trial in the hope that the plaintiff cannot meet his or her burden of proving that contamination is reasonably abatable. Under the proposed paradigm and re-allocated burden of proof, until a site has been adequately characterized the defendant would need to obtain a delayed trial date or else risk a failure of proof that the contamination cannot be reasonably abated. Similarly, a defendant also would have an incentive to become involved in the regulatory process and urge agencies to proceed promptly with approval of site investigation and remediation plans and issuance of cleanup standards. The proposal further would eliminate the risk that a plaintiff would be unable to recover cleanup costs for contamination created by another because an agency failed to generate the site cleanup standards necessary for plaintiff to prove that the contamination was reasonably abatable.

454. Defendants thus could argue that delays or unreasonable demands imposed on them by plaintiffs in abatement activities (e.g., demanding that defendants pay without meaningful input about how the money would be spent, preventing defendants from communicating with regulatory agencies) would support equitable liability defenses of waiver, laches, or estoppel, *see, e.g.*, Jamail v. Stoneledge Condo. Owners Ass'n, 970 S.W.2d 673, 676-77 (Tex. App. 1978) (laches may bar or qualify relief in private continuing nuisance claim), or reduce defendant's share of abatement cost responsibility.

ment a future stigma on property value), compromising potentially available insurance coverage for site remediation, and undermining arguments to regulatory agencies about site-specific, cost-effective remediation strategies.

The proposed re-allocated burden of proof, by increasing the chance that the contamination will be considered a continuing nuisance and thus subject to reasonable abatement, also would encourage a cooperative approach to site characterization and settlement of cleanup cost responsibilities. A defendant may wish to avoid litigation transaction costs associated with the threat of a series of continuing nuisance lawsuits, the risk of an injunction ordering participation in a cleanup, the specter of prospective damages for plaintiff's lost use of the property until the contamination is abated, and the possibility that a plaintiff would demand punitive damages for defendant's alleged willful failure to abate a continuing nuisance. Similarly, a plaintiff would have an incentive to negotiate settlement because of the reduced chance of a large prospective damage award, the risk of regulatory compliance obligations, and the desire to avoid litigation transaction costs.

Finally, the proposed re-allocation of the burden of proof should be accompanied by a presumption that the only remedies available to a plaintiff are those consistent with the abatability of the nuisance.[455] These remedies would include (a) injunctive relief for abatement of contamination and its source (if still actively polluting) and (b) damages consistent with an abatable nuisance suffered by plaintiff during the limitations period,[456] such as abatement costs and damages arising from lost opportunity costs associated with unabated contamination (e.g., lost rents, delayed development profits).[457] Damages that assume permanent

---

455. This presumption would be consistent with the perspective of most courts that permanent nuisance claims permit a plaintiff to recover in one action all past, present, and future damages caused by the nuisance, and that continuing nuisance claims permit injunctive relief directed toward nuisance abatement as well as past damages suffered within the limitation but not future damages based on a nuisance that could be discontinued or abated at any time. *See generally* FISCHER, *supra* note 251, §84[b]-[e].

456. *See, e.g.*, Hoery v. United States, 64 P.3d 214, 219 n.7 (Colo. 2004) (damages available for continuing torts up to time of suit); Anderson v. State, 965 P.2d 783, 792 (Haw. Ct. App. 1998) ("[A] continuous tortious act should not be subject to a limitations period until the act ceases. The doctrine also recognizes that though the statute of limitations is tolled by a continuing tortious act, 'in such a case[,] a recovery may be had for all damages accruing within the statutory period before the action, although not for damages accrued before that period.' (quoting Wong Nin v. City & County of Honolulu, 33 Haw. 379 (1935))); Lyons v. Township of Wayne, 888 A.2d 426, 430 (N.J. 2005) ("One who suffers a continuing nuisance, therefore, is 'able to collect damages for each injury suffered within the limitations period.'" (citation omitted)). To promote efficiency, continuing nuisance damages incurred to the time of trial (rather than to the time the continuing nuisance action was commenced) also could be recoverable. *See, e.g.*, Renz v. 33rd Dist. Agric. Ass'n., 46 Cal. Rptr. 2d 67 (Cal. Ct. App. 1995) (permitting recovery of continuing nuisance damages between commencement and conclusion of action notwithstanding California Supreme Court *dicta* that damages available during limitations period preceding commencement of action).

457. *See, e.g.*, Kathan v. Bellows Falls Village Corp., 223 A.2d 470, 472 (Vt. 1966) (cost of repair rather than diminution in property value is appropriate measure of damages in this continuing

property impairment, such as diminution in property value, would be presumptively unavailable under the proposed paradigm.[458] Such a remedial presumption would further promote the policies of site remediation and voluntary cleanup on which the proposed paradigm is based by keeping the focus of continuing nuisance litigation on cleanup rather than on calculating the price of a permanent nuisance pollution easement.

CONCLUSION: NUISANCE LAW AND THE FUTURE OF PRIVATE CLEANUP COST DISPUTES

The uncertainty regarding the future role of federal law in private cleanup cost disputes caused by *Aviall* has created an opportunity for private nuisance to play a significant role in private cleanup cost disputes. Private nuisance law can fill significant gaps in the CERCLA private enforcement scheme by providing a private cleanup cost remedy at petroleum contamination sites[459] and encourage voluntary remediation by owners of contaminated property by allowing them to obtain from those who caused most or all of the pollution their fair share of cleanup costs.[460] Moreover, private nuisance law can promote litigation and remedial efficiency and flexibility by providing a common body of state law for resolution of all contamination-related claims[461] and encouraging technically sound, prompt, and cost-efficient remediation whether or not in compliance with the often time-consuming and expensive procedural requirements of the NCP.[462]

In its current state, however, private nuisance law remains a theory of unrealized potential inapplicable to many common types of private cleanup cost disputes. In most states, a landowner cannot bring a private nuisance claim to address same property contamination problems. The responsible parties' failure to remediate long-standing contamination continues to substantially interfere with current owner's use and enjoyment of her property, yet in many jurisdictions nuisance claims to obtain cleanup costs from or an abatement injunction against the polluter would

---

trespass case); *see also* FISCHER, *supra* note 251, §84(b)(i) (temporary nuisance damages include diminished rental value and lost use damages).

458. *See supra* note 439 and accompanying text.

459. CERCLA does not provide a cleanup cost remedy for petroleum contamination because petroleum is expressly excluded from the definition of "hazardous substances" covered by the statute. 42 U.S.C.A. § 9601(14) (West 2006). *See supra* note 136 and accompanying text.

460. The current owner of contaminated property is a "covered person" liable under CERCLA section 107(a)(1). 42 U.S.C.A. § 9607(a)(1). Under *Aviall*, the landowner would be unable to bring a section 113(f) contribution action unless it first had been sued under CERCLA or settled with the government, while under pre-*Aviall* case law the current owner as a PRP also would be barred from asserting a section 107(a)(4)(B) cost recovery claim. *See supra* notes 72-83 and accompanying text.

461. Private remedies under CERCLA only address recovery of "necessary costs of response" to a release or threatened release of hazardous substances. 42 U.S.C.A. §§ 9607(a)(4)(B), 9613(f). CERCLA does not authorize private claim injunctive relief, nor does it provide for personal injury damages, lost profit or other economic damages, lost use or diminution in value property damages. *See supra* note 63 and accompanying text. Some or all of these contamination-related damages may be available to a current owner under a continuing or permanent nuisance theory.

462. *See supra* notes 140-45, 429 and accompanying text.

be time-barred. Finally, the doctrinal limitations of current nuisance law are underscored by the dramatic state-by-state variations in the scope of nuisance law, reflecting a patchwork quilt of random environmental protection rather than a coherent body of law responding to the nationwide problem of remediating contaminated properties.

The environmental nuisance paradigm proposed by this article would solve these problems by (a) expanding the scope of nuisance liability to include same property private nuisance disputes, (b) eliminating *caveat emptor* and other common law market-based risk allocation tools as defenses to liability,[463] (c) adopting a reasonable abatability of harm standard for continuing nuisance, (d) creating a rebuttable presumption that a soil or groundwater contamination nuisance was continuing (i.e., the contamination was reasonably abatable), and (e) placing the burden of proof on the party contending that a nuisance was permanent to show that the contamination was not reasonably abatable. This proposed paradigm would allocate cleanup cost obligations on the basis of comparative responsibility for site conditions, encouraging voluntary cleanups by current landowners and the informal resolution of cleanup cost allocation disputes.

This paradigm could be adopted by the states in several ways. First, state appellate courts could embrace the principles underlying the paradigm through the case-by-case evolution of common law. The definition of nuisance under the Restatement (Second) of Torts and most state law is broad enough to encompass same property cleanup cost disputes, and many state appellate courts have not yet directly addressed whether soil or groundwater contamination constitutes a continuing or permanent nuisance.[464]

Second, state legislatures may find that legislation represents a faster and more efficient solution. Some state appellate courts may be unable to implement the paradigm because of *stare decisis* concerns or limitations imposed by existing state statutes. Moreover, the evolution of state private nuisance case law could take years. To avoid these potential obstacles, state legislatures could enact an environmental nuisance statute codifying the elements of this paradigm. Whether accomplished by statute or case law, however, the goal of meaningfully addressing the nation's soil and groundwater pollution problem will be served by each state that chooses to embrace the paradigm.

---

463. An express contractual release constituting a knowing waiver of nuisance-based rights would remain enforceable under the proposed paradigm. *Caveat emptor* and other common law market-based risk allocation tools would remain relevant to fashioning a remedy against a liable defendant. *See supra* notes 399-407 and accompanying text.

464. *See supra* note 364 and accompanying text.

Other bodies of law, of course, remain applicable to private cleanup cost disputes.[465] To the extent that after *Aviall* a landowner can still recover cleanup costs from other PRPs under CERCLA, she may choose to proceed with a cleanup consistent with the NCP in order to take advantage of CERCLA's retroactive, status-based strict liability scheme[466] as well as comparable private claims possibly available under state Superfund laws. Buyers of contaminated property may have contract or fraud-based claims against sellers, newly-discovered contamination may give rise to a negligence claim, and contamination from a neighboring property could provide the basis for a trespass or traditional private nuisance action. None of these legal theories, however, provide the breadth and flexibility that would be available under the proposed private nuisance paradigm.[467]

A modernized law of private nuisance would assure the availability of a cleanup cost remedy to the owner of property contaminated in whole or in part by others. The sea change in contaminated property law occasioned by *Aviall* has shined a spotlight on the need to modernize the law of private nuisance. State courts and legislatures should seize the opportunity to bring a revitalized law of private nuisance law back from the margins as a significant rule of decision in private cleanup cost disputes.

465. *See supra* note 363.

466. CERCLA should not broadly preempt direct actions under nuisance or other state law theories for cleanup cost damages or abatement injunctive relief (at least in the absence of an EPA or court order issued under RCRA or CERCLA). For a discussion of CERCLA preemption of state law direct and derivative claims, see Aronovsky, *supra* note 5, at 90-104.

467. *See supra* notes 150-61 and accompanying text.

ESTATE TAXATION OF REDEMPTION AGREEMENTS: THE TREASURY LOSES "CONTROL"

JAMES M. DELANEY†

TABLE OF CONTENTS

† Associate Professor, University of Wyoming College of Law; University of Florida College of Law, L.L.M. (Taxation); Gonzaga University School of Law, J.D.; University of Washington, B.A. (Economics).

INTRODUCTION

For many years equity interest owners in closely held entities have engaged in buy-sell agreements. Buy-sell agreements provide for, among other things, restrictions on the transfer of equity interests upon the occurrence of certain triggering events such as the death, disability, or retirement of an equity holder. Parties to buy-sell agreements experience varying income and estate tax consequences depending on the type of agreement, form of business entity associated with the agreement, the manner in which an equity holder's interest is transferred, and the source of the funds that are used to purchase the interest. In many cases, an insurance policy on the life of an equity holder serves as a source for the funds that will be used to acquire a decedent equity holder's interest upon his or her death. Though complex, the federal tax consequences to business entities and equity holders that engage in such buy-sell arrangements have largely been settled.

However, with the advent of the limited liability company (the "LLC") and with two federal circuit courts of appeal issuing opinions that impact how estates of deceased corporate shareholders are taxed, the manner in which estate tax is imposed on transfers subject to certain common buy-sell arrangements has been called into question.[1] Further, an analysis of the policy underlying the applicable sections of the Internal Revenue Code[2] (the "Code") governing such transactions reveals very little consistency between the relevant Code sections and applicable Treasury Regulations.[3] Not surprisingly, there is also little consistency in the manner in which such sections are applied by the courts and ad-

1. *See* Estate of Blount v. Comm'r, 428 F.3d 1338, 1339-40 (11th Cir. 2005); Estate of Cartwright, 183 F.3d 1034, 1035 (9th Cir. 1999).

2. All section references to the Code are references to Internal Revenue Code of 1986 as amended from time to time.

3. *See infra* Part II.B.1.a.

ministered by the Internal Revenue Service (the "IRS" or the "Service").[4]

Current provisions of the Code and regulations have historically been interpreted to require the value of insurance proceeds received on a life insurance policy owned by a corporation to be ratably included in the estate of the insured decedent shareholder via an increase in value stock of the entity.[5] Under this interpretation, a deceased shareholder's gross estate would reflect a proportionate increase in the value of the stock due to the company's receipt of the insurance proceeds.

However, the Ninth and recently the Eleventh Circuit Courts of Appeal have issued opinions which impact the manner in which stock of a closely held corporation subject to a buy-sell agreement is valued for estate tax purposes.[6] The Ninth and Eleventh Circuit Courts of Appeal have held, among other things, that in determining the value of a corporation for purposes of determining the value of a deceased shareholder's interest, the value of insurance proceeds received by the corporation must be offset against the corporation's obligation to redeem the shareholder's stock under the buy-sell arrangement.[7] The holdings remove the value of insurance proceeds from the value of the corporation which, in turn, reduces the value of each shareholder's equity interest. While this interpretation may be logical, it is misplaced in calculating the value of a shareholder's stock in a redemption transaction for estate tax purposes.

In contravention to the apparent intent of the Code and historical application of current regulations, the two opinions appear under certain circumstances to allow the value of insurance proceeds received by a closely held corporation to completely escape estate taxation.[8] This outcome thwarts the Code's overall goal of including either all or a portion of the value of such insurance proceeds in the insured decedent's estate.[9] The opinions are also inconsistent with the IRS's interpretation of its own regulations that govern the manner in which insurance proceeds received by a corporation are included in the value of a deceased shareholder's stock.[10]

Directly related to the above problem and due to the increased popularity of limited partnerships and limited liability companies, taxpayers have increasingly sought to implement buy-sell agreements within partnership or LLC structures. Although the regulations provide guid-

4. *See infra* Part II.B.1.a.iii.

5. *See* I.R.C. §§ 2031, 2042; Treas. Reg. §§ 20.2031-2(f) (as amended in 2006), 20.2042-1(c)(6) (as amended in 1979); *see also infra* Part II.B.1.a; Part II.B.1.b.

6. *Estate of Blount*, 428 F.3d at 1339-40; *Estate of Cartwright*, 183 F.3d at 1035.

7. 428 F.3d at 1346; 183 F.3d at 1038.

8. *See infra* Part II.B.1.a.iii.(b).

9. *See* I.R.C. § 2042(a)(2).

10. *See infra* Part II.B.1.a.iii.(c).

ance on the tax consequences of insurance proceeds received by a corporation, no regulatory guidance is available in relation to receipt of insurance proceeds in a partnership structure.[11] Taxpayers and commentators are now inquiring as to the tax consequences of the receipt of insurance proceeds by a limited partnership or limited liability company. Without guidance from the Treasury on the use of an insurance funded buy-sell arrangement in an LLC or limited partnership structure, taxpayers and their advisors are left to interpret a small number of old cases and rulings to determine the tax consequences of various proposed structures. Further, existing rulings by the IRS in the partnership area are inconsistent with the manner in which the regulations treat corporations and shareholders under similar circumstances.[12] Given the increasing number of closely held companies that are being formed in the United States and the important business and estate planning goals that are served by such buy-sell agreements, the tax consequences of such arrangements should be clarified.[13]

In an effort to address these problems, this article first provides an overview of the manner in which buy-sell arrangements are structured with a specific focus on using insurance proceeds to fund the arrangement.[14] The article reviews the historical manner in which corporations and shareholders who are parties to buy-sell agreements are taxed where the agreement is funded by life insurance proceeds.[15] An argument is made that the recent holdings of the Ninth and Eleventh Circuit Courts of Appeal inappropriately allow taxpayers to exclude all of the value of insurance proceeds payable to a closely held corporation from the gross estate of a deceased shareholder.[16]

The article then analyzes the application of the Code, regulations, and rulings applicable to the receipt by partnerships and LLCs of insurance proceeds of a policy on the life of a member of partner where a buy-

---

11. *See, e.g.*, Treas. Reg. § 20.2042-1(c)(6).

12. *See, e.g.*, I.R.S. Gen. Couns. Mem. 39034 (Sept. 21, 1983), *available at* 1983 GCM LEXIS 83,*1-2; Rev. Rul. 83-147, 1983-2 C.B. 158.

13. For data and statistics on the increasing number of LLC formations over the last several years see TaxProf Blog, http://taxprof.typepad.com/taxprof_blog/govt_reports/index.html which shows the following statistical graphic that compares the cumulative increase in LLC formations to formations of corporations in several large states:

	Cumulative +/-		2005		2004		2003		2002	
State	LLC	Corp	LLC	Corp	LLC	Corp	LLC	Corp	LLC	Corp
CA	87.1%	20.1%	70,024	107,923	58,097	103,325	45,274	93,696	37,429	89,880
FL	237.9%	23.7%	130,558	175,698	100,070	177,490	62,406	168,080	38,639	142,036
NY	72.1%	(11.4%)	58,847	82,300	47,967	84,434	40,768	83,273	34,193	92,929
OH	60.3%	(18.5%)	42,594	14,921	38,765	16,386	31,147	16,601	26,575	18,299
PA	109.6%	4.6%	27,885	23,107	23,752	23,156	16,472	20,943	13,302	22,096
TX	81.3%	(27.3%)	59,076	40,945	49,677	42,302	35,285	55,107	32,593	56,319

The above graphic was synthesized by Larry Ribstein from data obtained from the International Association of Commercial Administrators, Ideoblog, http://busmovie.typepad.com/ideoblog/2006/05/the_data_is_out.html.

14. *See infra* Part I.

15. *See infra* Part II.

16. *See infra* Part II.B.1.

sell agreement is in place.[17] The article focuses on areas in which the cases, rulings, statutes, and regulations either inaccurately or inadequately provide guidance to taxpayers. A conclusion is reached that the guidance in relation to corporations is inconsistent with treatment accorded to partners of partnerships and members of LLCs. In this regard, the article proposes that the regulations applicable to corporations, with suggested amendments, should apply to limited liability companies and limited partnerships in the same manner as they apply to corporations.[18]

Finally, the article questions whether the case precedents and regulations now in force protect the overall goals of the Code as originally enacted by Congress.[19] The author concludes that the Eleventh Circuit Court of Appeal's recent decision is arguably inaccurate. At the same time, the author acknowledges the ambiguities in the current regulations as interpreted by the Courts of Appeal and proposes amendments that will assist courts and taxpayers in interpreting them. The amendments, if adopted, seek to clarify the regulations to make them consistent with the intent of the Code and apply equally to corporations and entities that are treated as partnerships for tax purposes.

I. OVERVIEW OF BUY-SELL ARRANGEMENTS

A buy-sell agreement is a contract pursuant to which a corporate shareholder, partner, or member of an LLC may (or must) offer his or her equity interest for sale to the company or the other equity holders upon the occurrence of certain triggering events.[20] Triggering events may include, among others, the equity holder's death, disability, retirement, or termination.

A buy-sell arrangement generally restricts the sale or transfer of an ownership interest in the entity to certain related parties or unrelated parties in general. For example, the equity holders of a closely held entity may wish to limit future ownership to individual family members who are perceived as capable of participating in management and operations of the entity. The terms of a buy-sell agreement also may include a method for determining a price for a unit of interest in an entity. By limiting ownership and setting a method of valuation, a buy-sell agreement can provide a method that is acceptable to all the equity holders of transferring control and continuing the existence of an entity upon the death, disability, or termination of employment of a single equity holder.

A buy-sell arrangement can create a market for a closely held equity interest. Without such an arrangement, little, if any, market may exist for

17. *See infra* Part II.B.1.b.
18. *See infra* Part III.
19. *See infra* Part III.
20. *See* Jonathan E. Strouse, *Redemption and Cross-Purchase Agreements: A Comparison*, THE PRACTICAL ACCOUNTANT, Oct. 1991, at 44.

an interest in a closely held entity. With a buy-sell arrangement in place, a minority interest in a closely held company may be transferred to other equity holders, family members or redeemed by the company. A buy-sell agreement will often set the price for the purchase.

A buy-sell agreement can also provide a source of funds upon a purchase or redemption of a deceased shareholder's ownership interest. For example, where a corporation, partnership, or LLC becomes a party to a buy-sell agreement and is required under certain circumstances to purchase the interest of an equity holder, the company may be authorized under the arrangement to purchase life insurance on each of the equity holders in order to fund the purchase of a deceased equity holder's interest. In the event the equity holder passes away, the entity will receive the insurance proceeds and use such proceeds to purchase the decedent's equity interest from his or her estate.

Of course, buy-sell arrangements differ in structure and form. The two most common forms of buy-sell agreements are cross-purchase agreements and redemption agreements.[21] The two structures result in different federal income and estate tax consequences to the entity and its equity holders. Depending on the facts and circumstances, the equity holders of an entity will select the structure that is most efficient from an economic, management, income, and estate tax perspective. This section initially focuses on the form of typical structures that use insurance proceeds to fund buy-sell agreements and then analyzes federal tax consequences of each of the basic structures.

A. Cross-Purchase Agreements

A cross-purchase agreement is a contract among the equity holders of a company. Under the terms of a cross-purchase agreement, each equity holder agrees or has an option to purchase some or all of the stock offered by the other equity holders.[22] The corporation, partnership, or LLC is typically not a party to a cross-purchase agreement. In a cross-purchase arrangement, the remaining equity holders must purchase, for example, another equity holder's interest upon a triggering event (e.g., death).

In a cross-purchase arrangement, each equity holder is authorized to acquire an insurance policy on the lives of each of the other equity holders.[23] Each shareholder pays the premiums on the policies that he or she has acquired on the lives of the other equity holders. Upon the death of

21. There are also variants on these two basic forms of the agreements that may include, among others, agreements between companies and their equity holders under which the equity holder may offer his or her interest initially to the company and thereafter to the other equity holders. Alternatively, the equity holder may first offer his or her interest to the other equity holders and thereafter to the company.

22. Strouse, *supra* note 20, at 44.

23. *See, e.g., id.* at 44-45.

an insured equity holder, the surviving equity holder will receive the life insurance proceeds as beneficiary on the policy ensuring the life of the decedent equity holder.[24] The surviving equity holders will then use the insurance proceeds to satisfy their obligation under the cross-purchase agreement to acquire the deceased equity holder's interest.

1. Tax Consequences to the Selling Equity Holder

There are a number of income and estate tax consequences to the equity holders in a cross-purchase agreement which differ depending upon whether the company is a C corporation, S corporation, or an entity treated as a partnership for tax purposes (e.g., a general partnership, limited partnership or LLC). Generally, there are no material tax consequences to the company when a cross-purchase agreement is executed and equity interests are exchanged among the equity holders under the terms of the agreement.

Upon the death of a shareholder, the estate will receive a basis step up to fair market value in relation to the deceased shareholder's stock.[25] Where insurance proceeds are used to fund a cross purchase obligation, the remaining shareholders use the policy proceeds to purchase a deceased shareholder's equity interest from the decedent's estate. Corporate stock generally fits into the definition of a capital asset.[26] Thus, any gain recognized on the sale of the corporate stock is generally subject to the lower tax rates applicable to capital gains.[27] Under these circumstances, when the remaining shareholders purchase the decedent's equity interest at fair market value from decedent's estate, the estate receives the stock with a basis equal to its fair market value and the estate will not realize any gain upon the death of the decedent.[28]

The receipt of the insurance proceeds by the remaining shareholders will not be included in deceased shareholder's estate unless the estate held incidents of ownership in the policy.[29] Where the remaining shareholders own the policies, the remaining shareholders should have all incidents of ownership in the policy insuring decedent's life. Neither the decedent nor the decedent's estate should hold incidents of ownership as contemplated under Code section 2042.

---

24. *Id.* at 45.

25. I.R.C. § 1014(a)(1). There is an exception to the above in relation to a shareholder of an S corporation. The estate of a deceased S corporation shareholder will not receive a step up in the basis to fair market value upon the shareholder's death to the extent of the shareholder's interest in unrealized receivables of the corporation.

26. I.R.C. § 1221. An exception exists where the shareholder is a "dealer" in stock.

27. *Id.*; *see also* I.R.C. §§ 1222, 1223.

28. This conclusion assumes that there was not a meaningful change in the value of the stock between the time that the decedent shareholder died and the time that the estate sells the stock to the remaining shareholders. In the event that the value of the stock changes in the interim period, gain or loss may be realized and recognized by the estate.

29. *See* I.R.C. § 2042.

The tax consequences to a selling partner or member of an LLC upon the sale of units subject to a cross-purchase agreement are similar to the tax consequences to a corporate shareholder. Similar to a corporate shareholder, any gain or loss realized by a partner upon a sale or exchange of an interest in a partnership or LLC is generally characterized as gain or loss from the sale of a capital asset.[30] Under these circumstances, the selling deceased partner or member's estate should experience a step up in the basis of the ownership interest and realize capital gain, if any, on the exchange.[31]

2. Tax Impact on the Cross-Purchase Buyer

A purchasing shareholder or purchasing partner's initial basis in an interest acquired from another shareholder or partner is equal to its cost.[32] Generally, the cost is equal to fair market value of the equity interest purchased from decedent's estate. A transfer of a partnership interest ordinarily will not terminate the partnership unless fifty percent or more of the total interest in capital and profits is sold or exchanged in a twelve-month period.[33]

3. Drawbacks of the Cross-Purchase Agreement

One drawback of the cross-purchase agreement exists when there are a large number of equity holders that are parties to the buy-sell arrangement. If, for instance, there are just two equity holders, two life insurance policies are sufficient; one policy is beneficially owned by A insuring the life of B and one policy is beneficially owned by B insuring the life of A. However, when there are a large number of equity holders, the number of policies required may become impractical. The number of policies required is equal to i (the number of equity holders) multiplied by $(i - 1)$.[34] For example, if a company has 6 equity holders, the number of required insurance policies would be 30 [calculated as follows: 6 x (6-1)]. Under these circumstances, the cost and administrative burden of managing the policies may become problematic. For this and various other reasons, equity holders may instead structure their buy-sell arrangement as a redemption agreement.

---

30. I.R.C. § 741.

31. *See* Treas. Reg. § 1.742-1 (1960). Exceptions exist in relation to the above general rules where the partnership holds specific types of assets. Similar to the rule applicable to the estate of an S corporation shareholder, the estate of a deceased partner will not receive a step up in the basis to fair market value upon the partner's death to the extent the value of the interest is decreased by items of income in respect of a decedent (items of "IRD"). *See* I.R.C. § 691; *see also* WILLIAM S. MCKEE ET AL., FEDERAL TAXATION OF PARTNERSHIPS AND PARTNERS ¶ 23.04[1] (2006). Items of IRD include, among other things, certain partnership receivables and the deceased partner's distributive share of partnership income for the period ending with his death. *Id.* (citing I.R.S. Priv. Ltr. Rul. 91-02-018 (Oct. 12, 1990)).

32. I.R.C. § 1012.

33. I.R.C. §708(b); *see also* MCKEE ET AL., *supra* note 31, ¶ 15.02[2](d).

34. Strouse, *supra* note 20, at 45.

Another practical problem with cross-purchase agreements can arise when a new shareholder or partner enters into the agreement. The new shareholder or partner must obtain an insurance policy on the life of each of the other equity holders.[35] Obtaining an insurance policy on an older shareholder or partner who is in poor health or of advanced age can be expensive or hard.[36]

Finally, where insurance proceeds are received under a policy by remaining shareholders or partners, there is no guaranty that such proceeds will be used to purchase the decedent shareholder or partner's equity interest.[37] The proceeds may instead be expended in an unrelated fashion by the remaining shareholders or the proceeds may be subject to creditors' claims.[38]

B. Redemption Agreements

A redemption agreement is a contract between the equity holders of the company and the company itself. Under a typical redemption agreement, a corporation, partnership, or LLC agrees to purchase or redeem stock or units offered by the shareholders, partners, or members.[39] If, pursuant to the agreement, the company funds the purchase of the equity interests with proceeds from life insurance, the company generally purchases only one policy on the life of each equity holder.[40] Thus, for example, if six equity holders exist, the company need only purchase six policies.[41]

One benefit of a redemption agreement is that the premiums will be paid by the company which may oversee and confirm that the policy is maintained.[42] Since the company pays the premiums on the life insurance policies, the cost of insurance is borne by the shareholders in proportion to their equity interests.

Redemption agreements have specific federal income and estate tax consequences to the entities and equity holders that engage in such agreements. The federal income tax consequences to the parties involved in a corporate redemption agreement depend on whether the entity that is a party to the agreement is a C corporation, partnership, or LLC treated

35. MCKEE ET AL., *supra* note 31, ¶ 23.07[1].
36. *Id.*
37. *Id.*
38. *Id.*
39. Strouse, *supra* note 20, at 45.
40. *Id.*
41. *Id.* at 46. If, in the alternative, a cross-purchase structure had been implemented here, the company would have been required to purchase 30 policies under the formula. *See supra* note 34 and accompanying text.
42. *Id.* at 45.

as a partnership for tax purposes. The analysis below is intended to cover the basic federal tax consequences of a redemption agreement.[43]

1. Tax Impact on Corporation & Shareholders

Implementation of a redemption agreement between a C corporation and its shareholders requires analysis of whether the distribution of the insurance proceeds will be treated as a dividend distribution under Code section 301 or a taxable exchange between the corporation and the shareholder under section 302.[44] The complexity of sections 301 and 302 generally causes redemption agreements to be viewed as more complex than a cross-purchase buy-sell arrangement which is among only the shareholders. It can be time consuming and costly to determine the tax consequences of a corporate distribution.

Upon the death of a shareholder who is a party to a corporate redemption agreement, the remaining shareholders may be faced with the difficult task of determining whether the distribution in redemption is a dividend or consideration received in sale or exchange. In the event that the distribution is characterized as a dividend, the remaining equity holders will receive no step up in the basis of their equity interests upon the redemption.[45]

a. Tax Considerations: Sale or Exchange Treatment

A redemption qualifies as a sale or exchange of a shareholder's stock under four different circumstances: (1) if it is not essentially equivalent to a dividend,[46] (2) if the distribution is substantially disproportionate with respect to the shareholder,[47] (3) if redemption completely terminates the shareholder's interest in the corporation,[48] or (4) the redemption is from an unincorporated shareholder and is in partial liquida-

43. For a more exhaustive analysis, see also HOWARD M. ZARITSKY, STRUCTURING BUY-SELL AGREEMENTS: ANALYSIS WITH FORMS ch. 8 (2005).

44. I.R.C. §§ 301, 302.

45. I.R.C. § 301.

46. I.R.C. § 302(b)(1). In general, in order to qualify under § 302(b)(1) as a redemption that is essentially equivalent to a dividend, the redemption must result in a meaningful reduction of the shareholder's proportionate interest in the corporation. United States v. Davis, 397 U.S. 301, 313 (1970); *see also* Rev. Rul. 85-106, 1985-2 C.B. 116. For this purpose, the attribution rules of § 318 of the Code apply. *Davis*, 397 U.S. at 307. In determining whether a reduction in interest is "meaningful," the most significant rights are: (1) the right to vote and thereby exercise control; (2) the right to participate in current earnings and accumulated surplus; and (3) the right to share in net assets upon liquidation. *See* Rev. Rul. 85-106, 1985-2 C.B. 116 (citing Rev. Rul. 81-289, 1981-2 C.B. 82).

47. I.R.C. § 302(b)(2). In general, a redemption is substantially disproportionate within the meaning of § 302(b)(2) if the shareholder's interest in outstanding common voting and nonvoting common stock after the redemption is less than 80% of the shareholder's interest before the redemption, and if, after the redemption, the shareholder owns less than 50% of the combined voting power of all shares. I.R.C. § 302(b)(2)(C).

48. I.R.C. § 302(b)(3). In general, a redemption qualifies as a complete termination under § 302(b)(3) if a shareholder's interest terminates all interests in the corporation as a result of the redemption. For this purpose, the attribution rules of § 318 of the Code apply. *See* I.R.C. § 302(c)(2).

tion of the distributing corporation.[49] To the extent that the distribution in redemption meets any one of these specific circumstances, the distribution will qualify for capital gains rather than ordinary income treatment.[50] Again, the determination of whether a redemption is a sale or exchange under Code section 302(b) is complex. A complete discussion of each of the four circumstances is beyond the scope of this article.[51] Nevertheless, classification of a distribution in redemption as a sale or exchange, as opposed to a dividend, is important primarily because of the ability of the shareholder to benefit from his or her basis in the shares.

For example, assume A and B are shareholders of C Corp. (which has substantial earnings and profits). A, B and C Corp engage in a redemption agreement pursuant to which A's stock is to be redeemed for $100,000 ($1.00 per share as of the date of death; 100,000 shares) upon A's death. If the redemption is taxed as a sale or exchange of the stock, the shareholder decedent's estate should recognize no gain due to the step up in basis to fair market value on the date of the decedent's death.[52]

Thus, the main reason for avoiding dividend treatment is to take advantage of the tax-free return of basis. Prior to the imposition of a lower tax rate upon dividends, historically, the disparity between the ordinary tax rate imposed on dividend income and capital gains provided an even greater benefit to sale or exchange treatment upon redemption.[53]

b. Tax Considerations: Dividend Treatment

If a distribution in redemption of stock by a corporation does not meet the requirements of section 302(b), the distribution generally will be classified as a distribution of a dividend to the extent of the corporation's earnings and profits.[54] If a distribution exceeds earnings and profits, it will represent a return of capital to the extent of the shareholder's basis in its stock.[55] Finally, to the extent the distribution exceeds the shareholder's basis in its stock, the distribution will be treated as gain from the sale or exchange of property.[56] Due to the fact that an estate will generally receive a step up in the basis of the decedent's assets to fair market value on the date of the decedent's death, characterization of the distribution as a dividend can be very costly from a tax perspective as compared to sale or exchange treatment which results in little or no tax-

49. I.R.C. § 302(b)(4).

50. *See Davis*, 397 U.S. at 305.

51. For a thorough analysis of the application of these rules, see BORIS I. BITTKER & JAMES S. EUSTICE, FEDERAL INCOME TAXATION OF CORPORATIONS AND SHAREHOLDERS ch. 9 (2005) (covering redemptions of corporate stock).

52. *See* I.R.C. § 302(a).

53. If the redemption is taxed as a "dividend" pursuant to § 301(a) and § 302(d), the estate will be taxed at net capital gain rates pursuant to §1(h)(11). *See* I.R.C. § 1(h)(11) (generally providing that qualified dividend income is subject to tax at the net capital gain rates).

54. I.R.C. §§ 301, 302(c)(1).

55. I.R.C. § 301(c)(2).

56. I.R.C. § 301(c)(3)(A).

able gain. This is because, where there are substantial earnings and profits, the full amount of the distribution in redemption will be treated as a dividend subject to tax. Whereas, in a redemption treated as a sale or exchange, gain subject to tax on the distribution will be reduced by the amount of basis that is recovered.[57]

2. Tax Impact on Partnerships, LLCs, Partners, and Members

Generally, the payment by a partnership to a partner in complete liquidation of a partner's interest does not result in a tax consequence to the partnership.[58] Further, assets may generally be distributed in kind tax-free.[59] However, liquidation payments to partners are governed by section 736.

Like the rules that apply to corporate shareholder distributions, the tax consequences to a partner (or member of an LLC which is treated as a partnership for tax purposes) can be quite complex from a tax perspective. Initially, section 736 divides liquidation payments into two categories meeting the definitions of section 736(a) and section 736(b). Payments classified as section 736(a) payments are further subdivided into two categories which include payments of a distributive share of partnership income and payments more in the nature of guaranteed payments.[60] The section 736(b) category generally includes payments to the partner for his or her share of the partnership property.[61] Payments to a partner in relation to partnership property are further subject to either ordinary or capital characterization depending upon the nature of the property.

For example, payments for substantially appreciated property and unrealized receivables generally will attract ordinary income treatment.[62] Cash payments for other types of property such as goodwill, for example, will attract capital gain treatment.[63] Further, tax treatment may differ depending upon whether the partner was a partner in a partnership in which capital was an income-producing factor versus a partnership in which capital was not an income-producing factor.[64]

II. FEDERAL TAX CONSIDERATIONS OF USING INSURANCE TO FUND BUY-SELL ARRANGEMENTS

In addition to general business and tax considerations, there are specific income and estate tax rules that may affect the decision to use insurance proceeds to fund either a cross-purchase agreement or redemp-

57. *See* I.R.C. § 1001(a).
58. I.R.C. § 736.
59. 33A AM. JUR. 2d *Federal Taxation* ¶ 10376 (2006).
60. *See* I.R.C. § 736(a).
61. *See* I.R.C. § 736(b).
62. *See* I.R.C. § 751.
63. *See* I.R.C. §§ 731, 741.
64. *See, e.g.*, MCKEE ET AL., *supra* note 31, ch. 22.

tion agreement. The Code accords special income tax treatment to insurance proceeds that are payable upon the death of the insured.[65] Additionally, the estate tax sections of the Code also provide specific rules that define when a decedent's gross estate must include insurance proceeds.

A. Income Taxation of Life Insurance Proceeds

Code section 101(a) generally provides that gross income does not include amounts received under a life insurance contract if such amounts are paid by reason of death of the insured.[66] Under this rule, a beneficiary of a life insurance policy that receives insurance proceeds upon the death of the insured may exclude the insurance proceeds from gross income.

For example, assume A and B are the only shareholders of C Corp. Inc. They have engaged in a cross-purchase agreement (the "Agreement") under which, upon the death of either A or B, the other will buy the decedent's share for $100,000. Pursuant to the Agreement, A and B each purchase a life insurance policy on the life of the other and each of them intends to use the proceeds to purchase the other's interest upon the death of the other party. When A dies, B will receive the $100,000 from the insurer and may exclude the full amount of the proceeds from income.[67]

Exceptions to the section 101(a) general exclusionary rule apply under certain circumstances. If a life insurance contract or any interest therein is transferred for valuable consideration, the exclusion from gross income is limited to an amount equal to the sum of the actual value of the consideration plus any premiums paid by the transferee.[68] The phrase "transfer for valuable consideration" is defined as any absolute transfer-for-value of a right to receive all or part of the proceeds of a life insurance policy.[69] Thus, the creation, for value, of an enforceable contractual right to receive all or a part of the proceeds of a policy may constitute a transfer for valuable consideration of the policy or an interest therein.[70]

---

65. *See* I.R.C. § 101(a)(1).

66. *Id.* The tax-free receipt of proceeds can be likened to a free step up in basis. In effect, I.R.C. §101(a)(1) operates the same as a step up in basis. *See* I.R.C. § 101(a)(1). I.R.C. § 1022, which would replace I.R.C. § 1014 at the time of estate tax repeal under current law, provides for carryover basis. *See* I.R.C. § 1022(b)(2)(C). As such, life insurance could be viewed as even more tax beneficial than other properties.

67. *See* I.R.C. § 101(a)(1); *see also* ZARITSKY, *supra* note 43, ¶ 8.02[1][a].

68. I.R.C. § 101(a)(2). Also included with the amount paid and any future premiums are "other amounts" which phrase includes interest paid or accrued by the transferee on indebtedness with respect to such contract or any interest therein if such interest is not allowable as a deduction under § 264(a)(4). *Id.*

69. Treas. Reg. § 1.101-1(b)(4) (as amended in 1982).

70. *See id.*

For example, A pays a premium of $500 for an insurance policy in the face amount of $1,000 upon the life of B, and A subsequently transfers the policy to C for $600. Thereafter, C does not make any additional payments in relation to the policy and receives the proceeds upon the death of B. The amount which C can exclude from his income is limited to $600. The $400 of the proceeds which exceed the $600 purchase price must be included by C as income.

B. Estate and Transfer Taxation of Life Insurance

In addition to income tax consequences, the impact of using life insurance proceeds to fund a buy-sell agreement requires analysis of the estate tax consequences. Section 2031 operates in conjunction with section 2033 and generally defines "gross estate" as including the value of all of a decedent's property (tangible or intangible and wherever situated) at the time of death.[71] Broadly, section 2033 requires the value of all of a decedent's property be included in his or her gross estate at the time of death.[72] Although sections 2031 and 2033 each relate to value and inclusion, section 2033 identifies includible interests whereas section 2031 generally addresses valuation.[73]

Identification of specific interests includible in a decedent's gross estate is addressed in sections 2033 through 2046.[74] Included within that range of sections is section 2042 which provides rules in regard to inclusion of the value of insurance proceeds in a decedent's gross estate.[75]

1. Section 2042: Incidents of Ownership

Section 2042 operates under certain circumstances to require inclusion in a gross estate of the value of insurance proceeds payable upon a decedent's death.[76] Taxpayers generally attempt to structure their estates in a fashion that escapes application of section 2042 thereby excluding life insurance proceeds from their gross estates. Fundamentally, section 2042 is a specific inclusion section that requires insurance proceeds receivable by an executor on the life of a decedent to be included in his or her gross estate.[77] There is no question that insurance proceeds from a policy on the life of the decedent that are actually received by an executor must be included in a decedent's estate.

Less intuitive is section 2042's requirement under certain circumstances that amounts payable to beneficiaries other than the decedent's

71. I.R.C. § 2031(a); *see also* I.R.C. § 2033 ("The value of gross estate shall include the value of all property to the extent of the interest therein of the decedent at the time of his death.").

72. *See* I.R.C. § 2033.

73. *See* RICHARD B. STEPHENS ET AL., FEDERAL ESTATE & GIFT TAXATION ¶ 4.01 (2006), *available at* 1999 WL 1031606.

74. *Id.*; *see also* I.R.C. §§ 2033-2046.

75. *See* I.R.C. § 2042.

76. *See id.*

77. I.R.C. § 2042(1).

estate under policies on the life of the decedent must also be included in the decedent's gross estate. Where the decedent possesses "incidents of ownership" in relation to the insurance policy, proceeds payable to a third party must be included in the decedent's gross estate.[78] The term "incidents of ownership" is not limited in its meaning to ownership of the policy in a technical legal sense.[79] Incidents of ownership include *any* economic interest or benefit from an insurance policy.[80]

The regulations specifically provide that "incidents of ownership" include the power to change the beneficiary, to surrender or cancel the policy, to assign the policy or revoke an assignment, to pledge the policy for a loan, or to borrow against the cash surrender value of the policy.[81] This list, though illustrative, is not exhaustive.[82] Incidents of ownership broadly encompass most every right that is retained in a policy.[83] In the obvious example, if an insured transfers a policy to another but retains the right to change the beneficiary on the policy, retention of such right will result in the proceeds being included in the insured's estate.[84]

A number of cases and rulings address less obvious examples and add to the definition of the phrase "incidents of ownership" for purposes of applying the Code and regulations. In Revenue Ruling 61-123,[85] the IRS ruled on a fact pattern in which an airline passenger purchased an accident insurance policy on his life prior to his death in an airplane crash.[86] After filling in the beneficiary designation, he mailed the policy insuring his life to the beneficiary and boarded the plane.[87] The proceeds were not payable to the decedent's estate. However, under these circumstances, the IRS ruled that the proceeds of the policy were includible in the decedent's gross estate even though as a practical matter it was impossible for him to exercise any incidents of ownership while the plane was in flight.[88]

Similarly, in *Commissioner v. Estate of Noel*,[89] the decedent's spouse purchased an accidental death insurance policy on his life several hours prior to a plane crash, which killed the insured husband.[90] Not-

78. I.R.C. § 2042(2).
79. Treas. Reg. § 20.2042-1(c)(2) (as amended in 1979).
80. *See* Treas. Reg. § 20.2042-1(c); Chase Nat'l Bank v. United States, 278 U.S. 327, 335 (1929); *see also* ZARITSKY, *supra* note 43, ¶ 8.02[4].
81. Treas. Reg. § 20.2042-1(c); *see also* ZARITSKY, *supra* note 43, ¶ 8.02[4].
82. H.R. REP. NO. 2333-77, § 404 (1942), *reprinted in* 1942-2 C.B. 372, 491; *see also* STEPHENS ET AL., *supra* note 73, ¶ 4.14[4][a].
83. *See* STEPHENS ET AL., *supra* note 73, ¶ 4.14[4][a], *available at* 1999 WL 1031619.
84. *See* I.R.C. § 2042(2). This analysis is similar to the analysis applied in sections 2036 through 2038 relating to transfers of property where an interest is retained by the grantor of the property. *See* I.R.C. §§ 2036-2038.
85. Rev. Rul. 61-123, 1961-2 C.B. 151.
86. *Id.*
87. *Id.*
88. *Id.*
89. 380 U.S. 678 (1965).
90. *Estate of Noel*, 380 U.S. at 679-80.

withstanding that the decedent was in transit and could not actually have exercised any incidents of ownership, the Supreme Court held that the proceeds must be included in the decedent's estate.[91] The holding in *Noel* and the outcome in Rev. Rul. 61-123 indicate a tendency on the part of the Court and the IRS to find and apply incidents of ownership in an insurance policy where an individual has the right to change the beneficiary even though the individual had no real control over the beneficiary designation.

The right to change a beneficiary designation is not the only attribute that may result in incidents of ownership. Use of an insurance policy as collateral will also constitute "incidents of ownership." In *Estate of Krischer v. Commissioner*,[92] the decedent obtained loans, renewals of loans, or additions to existing loans from a financial institution.[93] At inception, each loan was secured by either one or both of two life insurance policies.[94] Renewals or additions to such loans were expressly made subject to an assignment of those policies as collateral.[95] The Commissioner determined that the decedent possessed incidents of ownership in the policies by virtue of his power to pledge them as collateral for past and future loans.[96] Attribution of incidents of ownership to the decedent caused the proceeds to be includible in decedent's gross estate pursuant to section 2042(2).[97] The Court in *Estate of Krischer* agreed.[98] Thus, notwithstanding that the policy owner irrevocably transfers a policy, if the owner retains the right to use the policy as collateral, the owner will continue to have incidents of ownership under section 2042(2).

In each of the above cases or rulings, the proceeds of the life insurance policy were payable to a third party and not the decedent insured's estate. Under these circumstances, section 2042 nevertheless operated to include the proceeds in the decedent's gross estate. The courts and the IRS have been very liberal in attributing incidents of ownership to a decedent where the decedent has retained even the slightest rights to exercise control over the policy.

However, the power to substitute one life insurance policy for another of equal value appears not to be an incident of ownership. In *Estate of Jordahl v. Commissioner*,[99] the decedent created a trust and named himself as one of three trustees.[100] The corpus of the trust included insurance policies on the decedent's life and other income-

91. *Id.* at 683-84.
92. 32 T.C.M. (CCH) 821, 1973 Tax Ct. Memo LEXIS 117 (1973).
93. *Id.* at *3.
94. *Krischer*, 1973 Tax Ct. Memo Lexis 117, at *3.
95. *Id.*
96. *Id.* at *9.
97. *Id.*
98. *Id.* at *13.
99. 65 T.C. 92 (1975).
100. *Jordahl*, 65 T.C. at 92.

producing assets.[101] The trustees were instructed to pay the premiums out of the assets of the trust and to pay over any remaining income to the decedent.[102] At no time was income insufficient to pay the premiums.[103] On his death, his daughter was to receive the income until she reached age 50 at which time she was to receive the principal.[104] The decedent retained the power to substitute securities, property, and policies "of equal value" for those transferred to the trust.[105] The court held that the insurance proceeds were not includible because the right to substitute other policies "of equal value" did not give him a right to the "economic benefits" of the policies and because his powers as trustee were strictly limited.[106]

a. Incidents of Ownership Held by Corporations

i. Current Treasury Regulations

Incidents of ownership on a policy held by a corporation can also be attributed to an individual shareholder of the corporation. Insurance proceeds payable to a third party other than the corporation must be included in an insured shareholder's estate if the insured shareholder can actually exercise any incident of ownership in the policy.[107] Thus, an insured shareholder will be attributed incidents of ownership held by the corporation where the insured owns a controlling interest in the corporation and the proceeds are payable to a beneficiary other than the corporation.[108] However, a shareholder may avoid being attributed incidents of ownership in an insurance policy possessed by the corporation under certain circumstances. Regulations under section 2042 provide an exception to this general rule of inclusion where a corporation's incidents of ownership will *not* be attributed to a decedent *through stock ownership* to the extent that the proceeds are payable directly to the corporation.[109]

The above analysis under the regulations first hinges on whether the corporation or a third party receives the insurance proceeds. More specifically, where a corporation is the beneficiary and recipient of proceeds of a life insurance policy on the life of a decedent shareholder, the corpo-

---

101. *Id.*
102. *Id.*
103. *Id.*
104. *Id.*
105. *Id.*
106. *Id.* at 101 (citing I.R.C. § 2042(2)).
107. Treas. Reg. § 20.2042-1(c)(6) (as amended in 1979). An example is given in the regulations providing that if the decedent is the controlling stockholder in a corporation and the corporation owns a life insurance policy on his or her life, the proceeds of which are payable to the decedent's spouse, the incidents of ownership held by the corporation will be attributed to the decedent through his or her stock ownership and the proceeds will be included in his or her estate under section 2042. *Id.*
108. *Id.*; *see also* KATHRYN G. HENKEL, ESTATE PLANNING & WEALTH PRESERVATION: STRATEGIES & SOLUTIONS ¶ 12.02[1][b][i] (2005).
109. Treas. Reg. § 20.2042-1(c)(6).

ration's incidents of ownership will *never* be attributed to the decedent shareholder through his or her stock.[110] This rule operates notwithstanding the fact that a majority shareholder would have actual control over any insurance policy held by the corporation. Applying this rule in isolation appears to allow the insurance proceeds to escape inclusion in a majority shareholder's estate and escape the direct requirements of section 2042 of the Code. However, as discussed extensively below, estates of majority shareholders have historically reported the value of the proceeds indirectly via an increase in the value of the decedent shareholder's stock under the general inclusion rule of section 2031.

On the other hand, where the insurance proceeds are not payable to or for the benefit of the corporation, the question arises as to whether incidents of ownership held by the corporation will be attributed to an insured shareholder. Where a decedent shareholder owns a controlling interest and the proceeds are payable to a third party (e.g., not the corporation), incidents of ownership held by the corporation are attributed to the decedent shareholder's estate under section 2042.[111] The attribution of incidents under these circumstances is based upon the fact that a majority shareholder has authority to actually exercise control over the policy. This, in turn, causes the insurance proceeds to be included in the majority shareholder's estate notwithstanding that the proceeds are payable to a third party. However, if a decedent shareholder owns only a *minority* interest, no incidents are attributable and no proceeds are directly included in the decedent shareholder's estate.[112]

Operation of these rules is demonstrated in the following example provided in the regulations under section 2042. Assume the decedent was the controlling stockholder in a corporation. Further, assume that the corporation owned a life insurance policy on the decedent shareholder's life and the proceeds were payable to the decedent's spouse, a third party. The incidents of ownership held by the corporation are attributed to the decedent through his stock ownership and all of the proceeds are includible in the decedent's gross estate under section 2042.[113] However, if in this example the policy proceeds had been payable forty percent (40%) to the decedent's spouse and sixty percent (60%) to the corporation, only forty percent of the proceeds would be included in decedent's estate under section 2042.[114] While the example in the Regulation does not specifically address the remaining sixty percent of the insurance proceeds that went to the corporation, it can be inferred that the value of the corporate stock in the decedent's estate will increase due to the corporation's receipt of the additional sixty percent. Further, it can

110. *Id.*
111. *See id.*
112. *Id.*
113. *Id.*
114. *Id.*

also be inferred that the increased value of the decedent's stock must be included in the decedent's gross estate under section 2031.[115]

In sum, the regulations appear to require inclusion of at least a ratable portion of the value of the insurance proceeds in a shareholder's estate where the proceeds are payable to the corporation under section 2031. Where the corporation receives the insurance proceeds, the Regulations operate on the presumption that the value of the corporate stock included in a decedent's estate under section 2031 will increase proportionately upon the corporation's receipt of the insurance proceeds. Conversely, where policy proceeds are payable to a party other than the corporation (or the insured shareholder's estate) and corporate incidents are attributable to the decedent shareholder, the decedent shareholder's estate is required to include all of the proceeds in his or her gross estate under section 2042. These rules reflect the close relationship that exists between section 2042 and section 2031.

ii. Current Case Law

The above analysis under the regulations is not, however, consistent with recent holdings of the Ninth Circuit Court of Appeals in *Estate of Cartwright v. Commissioner*[116] and the Eleventh Circuit Court of Appeals in *Estate of Blount v. Commissioner.*[117] In *Cartwright*, the decedent, Mr. Cartwright, was a majority shareholder in the CSB law firm ("CSB"), a California professional corporation.[118] Mr. Cartwright died in 1988 owning a majority 71.43% of the shares of CSB.[119] CSB held two insurance policies on the life of Mr. Cartwright which, upon his death, paid a total of $5,062,029 in proceeds.[120] Pursuant to a redemption agreement in place at the time of Mr. Cartwright's death between CSB and its shareholders, CSB paid the insurance proceeds that it received to Mr. Cartwright's estate in redemption of Mr. Cartwright's shares.[121] CSB treated $4,080,256 of the total payment as non-employee compensation and the remaining approximately one million dollars as a payment in redemption of Mr. Cartwright's stock.[122] However, the estate treated the full amount received as paid in redemption of Mr. Cart-

115. *See* T.D. 7312, 1974-1 C.B. 277, 1974 IRB LEXIS 835, at *3 (1974) (providing that where a corporation is the beneficiary of any portion of the proceeds of a life insurance policy, there is no need for that portion to be included in the gross estate under section 2042 because it directly affects the value of the stock that is included in the decedent's gross estate). *See infra* Part II.B.1.a.iii.

116. 183 F.3d 1034 (9th Cir. 1998).

117. 428 F.3d 1338 (11th Cir. 2005).

118. *Cartwright*, 183 F.3d at 1035; *see also* Estate of Cartwright v. Comm'r, 71 T.C.M. (CCH) 3200, 1996 Tax Ct. Memo LEXIS 299, at *3 (1996) (indicating that they incorporated the firm of Cartwright, Saroyan, Martin & Sucherman, Inc. (CSB), as a professional corporation under California law).

119. *Cartwright*, 183 F.3d at 1036.

120. *Id.*

121. *Id.*

122. *Id.*

wright's stock and none of the payment as non-employee compensation.[123] The IRS disagreed with the estate and determined that the $4,080,256 was compensation and that the estate owed $1,142,472 for its tax deficiency.[124] Consistent with CSB's treatment, the IRS treated the additional approximately one million dollars as the redemption amount.

The first issue the court addressed was whether the payment to Mr. Cartwright's estate was made solely for redemption of the stock or whether it was in part for Mr. Cartwright's stock and in part for his claim for compensation due.[125] Mr. Cartwright's estate contended that the full $5 million should be treated as paid in exchange for the redemption of the decedent's stock.[126] In affirming the Tax Court, the Ninth Circuit concluded that approximately one million dollars of the insurance proceeds constituted payment for Cartwright's stock and approximately four million dollars constituted compensation for services.[127] Under the Ninth Circuit's holding, approximately four-fifths of the payment constituted a liability of the company that related to compensation for services rendered before the decedent shareholder's death.

The court next addressed whether the five million dollars of insurance proceeds should have been included as an asset of CSB for purposes of valuing CSB stock held by Mr. Cartwright's estate under section 2031.[128] Affirming the Tax Court again and quoting the regulations under section 2031, the Ninth Circuit agreed that the Tax Court appropriately excluded *all* of the life insurance proceeds in valuing stock.[129] The court indicated that consideration must be given to non-operating assets including proceeds of life insurance policies payable to or for the benefit of the company, to the extent such non-operating assets have not been

123. *Id.*
124. *Id.*
125. *Id.*
126. *Id.*
127. *Id.* at 1036-38.
128. *Id.* at 1037-38.
129. *Id.* at 1038. Treas. Reg. § 20.2031-2(f)(2) provides that the fair market value of shares of stock should be determined by considering the company's net worth, prospective earning power and dividend-paying capacity, and other relative factors. The regulation states that:

> Some of the "other relevant factors" . . . are: The goodwill of the business; the economic outlook in the particular industry; the company's position in the industry and its management; the degree of control of the business represented by the block of stock to be valued; and the values of securities of corporations engaged in the same or similar lines of business which are listed on a stock exchange. However, the weight to be accorded such comparisons or any other evidentiary factors considered in the determination of a value depends upon the facts of each case. *In addition to the relevant factors described above, consideration shall also be given to non-operating assets, including proceeds of life insurance policies payable to or for the benefit of the company, to the extent such non-operating assets have not been taken into account in the determination of net worth, prospective earning power and dividend-earning capacity.* Complete financial and other data upon which the valuation is based should be submitted with the return, including copies of reports of any examinations of the company made by accountants, engineers, or any technical experts as of or near the applicable valuation date.

Treas. Reg. § 20.2031-2(f)(2) (emphasis added).

taken into account in determining net worth.[130] The court reasoned that the proceeds received by CSB from the insurance policy would not necessarily affect what a willing buyer would pay for CSB's stock because the insurance proceeds were offset dollar-for-dollar by CSB's obligation to pay out the entirety of policy benefits to Cartwright's estate.[131] Thus, the Court classified the insurance proceeds as being the kind of ordinary non-operating asset that should not be included in the value of CSB under the Treasury Regulations. Neither the Tax Court nor the Ninth Circuit Court of Appeals made any distinction between the approximately four million dollar payment related to compensation and the approximately one million dollars remitted to the estate in redemption of the decedent's shares. By excluding all of the insurance proceeds from the value of CSB, no distinction was made between the four million dollar obligation for services provided by the decedent prior to death and the one million dollar obligation to redeem the decedent's shares. Further, by failing to make any distinction between the two obligations, the holding in *Cartwright* allowed the estate to value its shares of CSB without proportionately increasing the value of the decedent's shares by a ratable portion of the one million dollars of insurance proceeds paid to the estate in the redemption.

On October 31, 2005, the Eleventh Circuit addressed a similar issue in *Estate of Blount v. Commissioner.*[132] In *Estate of Blount*, the decedent, Mr. Blount, was one of two shareholders of Blount Construction Company ("BCC"), a corporation formed in the state of Georgia.[133] The two shareholders, Mr. Blount and Mr. Jennings entered into a stock redemption agreement that required BCC to purchase the stock on the death of the holder at a price agreed upon by the parties.[134] In the early 1990s, BCC purchased insurance policies providing roughly three million dollars in order to fulfill its commitments to purchase the shareholders' stock under the redemption agreement.[135]

In 1996, Mr. Blount was diagnosed with cancer, and his doctor predicted that he had only a few months to live.[136] When Blount died in 1997, he owned roughly 83% of BCC, a clear majority interest.[137] In accordance with the redemption agreement, BCC paid Mr. Blount's estate four million dollars.[138] Mr. Blount's estate reported the value of the shares at four million dollars.[139] However, the IRS determined that the

130. *Cartwright*, 183 F.3d at 1038 (citing Treas. Reg. § 20.2031-2(f)(2)).
131. *Id.* at 1038.
132. 428 F.3d at 1339.
133. *Id.* at 1340.
134. *Id.*
135. *Id.*
136. *Id.*
137. *Id.* at 1341.
138. *Id.*
139. *Id.*

value of the stock was $7,921,975 and that the taxpayer had undervalued the stock by approximately four million dollars.[140]

Based upon expert testimony, the Tax Court concluded that the value of the company was approximately $6.75 million.[141] The Tax Court then added the insurance proceeds of $3.1 million to the value of the company to arrive at $9.85 million as the fair market value of the stock.[142] On review, the Eleventh Circuit concluded that the $6.75 million valuation for BCC was not erroneous.[143] However, the court of Appeals found that the inclusion by the Tax Court of the additional $3.1 million in insurance proceeds was in error.[144]

In declining to include the insurance proceeds in the value of BCC, the Eleventh Circuit agreed with the reasoning in *Cartwright* finding it to be "persuasive and consistent with common business sense."[145] Like the *Cartwright* court, the *Blount* court quoted regulations under section 2031 stating that in valuing the corporate stock, "consideration shall also be given to non-operating assets, including insurance proceeds of life insurance policies payable to or for the benefit of the company, to the extent that such non-operating assets have not been taken into account in the determination of net worth."[146] Focusing on the last clause of the quoted section of the regulation, the court concluded that the language limiting inclusion to the extent that such assets have not been taken into account in determining net worth precludes inclusion of the value of the insurance proceeds received by BCC.[147] The court reasoned that insurance proceeds are not the kind of ordinary non-operating asset that should be included in the value of the corporation.[148] Further, to the extent the insurance proceeds are required to be used to redeem the decedent shareholder's interest in the corporation, the proceeds are offset dollar-for-dollar by the corporation's obligation to purchase decedent's stock under the redemption agreement.[149]

In a footnote, the court indicated that the Commissioner argued that this interpretation frustrates the clear intent of Congress to include corporate owned life insurance in the estate of its sole shareholder.[150] However, in the same footnote, the court stated that "the legislative history relied upon by the Commissioner indicate[s] only that Congress believed

140. *Id.*
141. *Id.* at 1342.
142. *Id.*
143. *Id.*
144. *Id.*
145. *Id.* at 1345.
146. *Id.*
147. *Id.* (citing *Estate of Cartwright*, 183 F.3d at 1038; Huntsman v. Comm'r, 66 T.C. 861, 875 (1976)).
148. *Id.* at 1346.
149. *Id.*
150. *Id.* at 1345 n.6.

that a sole shareholder was deemed to have the incidents of ownership possessed by his corporation on insurance policies on his life."[151] The footnote goes on to briefly interpret Congressional intent:

> [T]he [R]egulations now provide that the incidents of ownership held by a corporation are not to be attributed to its shareholder, and no indication is included in the committee reports that Congress intended property owned by a decedent to be includable in his gross estate at other than its fair market value.[152]

The court further supported this questionable conclusion by indicating that "[t]o suggest that a reasonably competent business person, interested in acquiring a company, would ignore a $3 million liability strains credulity and defies any sensible construct of fair market value."[153]

iii. Applying the Case Doctrine Produces a Different Outcome than Application of the Code and Regulations

(a) Under the Cases

The opinions of the Ninth and Eleventh Circuits in *Cartwright* and *Blount* each addressed valuation of corporate stock within a decedent shareholder's estate in connection with a corporate redemption agreement. In both cases, the corporation in question owned a life insurance policy insuring the life of the majority shareholder, and the corporation received the insurance proceeds upon the majority shareholder's death. Both opinions focused upon determining stock value for purposes of including such value in the decedent shareholder's gross estate for purposes of section 2031. Each opinion concluded that the insurance proceeds received by the company should not affect the value of the stock in the decedent's estate. Each court reasoned that the insurance proceeds were to be offset dollar-for-dollar against an existing obligation of each corporation.

However, there is at least one factual distinction between the two cases. In *Blount*, all of the amounts received by the decedent shareholder's estate were in return for the stock redeemed by the corporation.[154] In *Cartwright*, however, only approximately one-fifth of the amounts received by the decedent shareholder's estate were received in return for the decedent shareholder's stock.[155] The remaining four-fifths of the amounts received were for services provided by the decedent shareholder prior to his death. The large majority of the amounts re-

151. *Id.*
152. *Id.*
153. *Id.* at 1346.
154. *See id.* at 1339-41.
155. *Estate of Cartwright*, 183 F.3d at 1036, 1038.

ceived by the decedent shareholder's estate represented an accrued compensation liability of the corporation not associated in any way with the obligation of the corporation to redeem the decedent shareholder's stock.

Due to the perceived dollar-for-dollar offset of these liabilities, each of the opinions results in the exclusion of insurance proceeds in valuing the corporate stock. Without an increase in the value of the stock related to the corporation's receipt of insurance proceeds, the estates of the two decedent majority shareholders effectively were allowed to exclude the value of any portion of the insurance proceeds in their respective estates. This outcome stands regardless of the fact that in each case, the majority shareholder actually had control over his respective corporation and, therefore, control over all of the incidents of ownership of the policies held by the corporations.

(b) Under the Regulations

Application of current regulations under sections 2042 and 2031 to the facts in each of these cases results in a different outcome than arrived at by the *Blount* and *Cartwright* courts. Because CSB and BCC received the proceeds of the life insurance policies on Cartwright and Blount respectively, the regulations specifically indicate that neither shareholder is attributed corporate incidents of ownership.[156] Therefore, neither of the decedent shareholders' estates would be attributed incidents of ownership, nor would either estate be required to include the insurance proceeds in gross estate under section 2042.

However, in determining the value of the stock under section 2031, Regulations section 20.2031-2(f) requires that consideration be given to non-operating assets, including proceeds of life insurance policies payable to or for the benefit of the company.[157] Contrary to the interpretations of the Circuit Courts of Appeal in *Cartwright* and *Blount,* the intent of this rule is to reflect the insurance proceeds received by the corporation as an asset that proportionately increases the value of the deceased shareholder's stock. The decedent shareholder's stock value for purposes of calculating gross estate must increase proportionate to the value of the insurance proceeds received by the company.

This conclusion is supported by the language in section 2042 and Regulations section 20.2042-1(c)(6). As previously discussed, section 2042 requires a decedent's estate to include all "amount[s] receivable by all other beneficiaries," under a policy on the life an insured shareholder with respect to which the shareholder possessed incidents of ownership.[158] Given that the definition of incidents of ownership is broad and includes "any" incidents of ownership, Congress must have intended that

156. Treas. Reg. § 20.2042-1(c)(6).
157. Treas. Reg. § 20.2031-2(f)
158. I.R.C. § 2042(2).

incidents of ownership be attributed to a majority shareholder of a corporation that owns a life insurance policy in the shareholder's life.[159] It is unlikely that Congress intended that a decedent majority shareholder be allowed to escape inclusion in his or her estate of life insurance proceeds received by his or her controlled corporation.

While the regulations under section 2042 clearly provide an exception to the inclusion of the proceeds *where the corporation receives the proceeds*, the regulations nevertheless attempt to further Congress' clear intent by requiring inclusion of a ratable portion of the proceeds in a decedent shareholder's gross estate as follows:

> In the case of economic benefits of a life insurance policy on the decedent's life that are reserved to a corporation of which the decedent is the sole or controlling stockholder, the corporation's incidents of ownership will not be attributed to the decedent through his stock ownership to the extent the proceeds of the policy are payable to the corporation . . . . ***See § 20.2031-2(f) for a rule providing that the insurance proceeds of certain life insurance policies shall be considered in determining the value of decedent's stock.***[160]

The first sentence is the exception to the general rule of inclusion that prevents corporate incidents from being attributed to a sole or majority shareholder where the policy proceeds are payable to the corporation. If the goal of section 2042 is to include the proceeds in a decedent's estate where a decedent has incidents of ownership, why are the incidents of ownership in the case of a majority shareholder specifically not attributed under the above regulation? The answer lies in the third sentence in the above quoted portion of the regulation. The third sentence, highlighted in italics and bold, refers to regulations under subsection 20.2031-2(f), which is specifically discussed by both the *Cartwright* and *Blount* courts. This regulation provides factors and guidelines for valuing stock that is not publicly traded, like the stock of CSB and BCC.[161] Under this regulation, consideration must be given to non-operating assets, including proceeds of life insurance policies payable to or for the benefit of the company, to the extent such non-operating assets have not been taken into account in the determination of net worth.[162] Read in conjunction with the first sentence of the above quoted language, section 20.2042-1(c)(6) is attempting to require inclusion of a ratable portion of the proceeds in the decedent majority shareholder's estate via a perceived increase in the value of the stock upon the corporation's receipt of the proceeds.

---

159. *See* Treas. Reg. § 20.2042-1(c)(2).
160. Treas. Reg. § 20.2042-1(c)(6) (emphasis added).
161. Treas. Reg. § 20.2031-2(f).
162. *Id.*

The Treasury's statements in 1974 in the preamble to the amendments to regulations under sections 2031 and 2042 support this reading of the regulation.[163] In 1974, Treasury Decision 7312 was published amending the regulations under section 2042 to add, among other things, the current version of Regulation section 20.2042-1(c)(6) and to specifically amend paragraph (f) of Regulation section 20.2031-2.[164] The preamble to the 1974 amendments provides the following explanation:

> Under section 2042 of the Code, if a decedent died possessed of any incidents of ownership in a life insurance policy on his life, the entire proceeds of the policy will be included in his gross estate for estate tax purposes. The term "incidents of ownership" is described in § 20.2042-1(c)(2) as including "a power to change the beneficiary reserved to a corporation of which the decedent is sole stockholder."[165]
>
> A problem was presented in Revenue Ruling 71-463, 1971-2 C.B. 333, which ruling was later withdrawn by Revenue Ruling 72-167, 1972-1 C.B. 307, as to whether a controlling stockholder should be treated as a "sole stockholder" for purposes of section 2042. The position taken in the proposed rules is that a controlling stockholder should be so treated. ***However, where a corporation is the beneficiary of any portion of the proceeds of a life insurance policy, there is no need for that portion to be included in the gross estate under 2042 because it directly affects the value of the stock that is included in the decedent's gross estate.***
>
> Accordingly, §20.2042-1(c) is amended by this document to provide that, with respect to proceeds of corporate-owned life insurance which are payable to either the corporation or a third party for a valid business purpose, incidents of ownership held by the corporation will not be attributed to the decedent through his stock ownership.[166]

Thus, where the corporation is the beneficiary of any portion of the policy proceeds, there is no need for that portion to be included in the gross estate under section 2042 or the regulations thereunder because it directly affects the value of the stock that is included in the decedent's gross estate under the section 2031 Regulations.[167] The intent of the 1974 change in the regulations was to remove the insurance proceeds from section 2042's inclusion requirement because the proceeds directly increase the value of the stock that is included in the decedent's gross estate under section 2031.[168] Without the exclusionary provision in the section 2042 regulations, the insurance proceeds would conceivably be included in a decedent shareholder's estate twice. The insurance pro-

163. *See* T.D. 7312, 1974-1 C.B. 277, 1974 IRB LEXIS 835, *1-4 (1974).
164. *Id.* at *5-10.
165. *Id.* at *2.
166. *Id.* (emphasis added).
167. *Id.* at *1-2.
168. *Id.*

ceeds would be included once via section 2042 due to the fact that the shareholder actually has incidents of ownership in the policies held by the corporation. The insurance proceeds would be included a second time under section 2031 due to the increase in the value of the shares held by the decedent shareholder's estate when the insurance proceeds are paid to the corporation. In an effort to prevent double inclusion in the decedent shareholder's estate of the value of the proceeds received by the corporation, the insurance proceeds included under section 2042 are removed by an amendment to the regulations leaving a single proportionate inclusion under section 2031 when valuing the stock held by the decedent shareholder at the time of his or her death.[169]

Without discussing the intent of section 2042 and the regulations thereunder, the opinions of the Ninth and Eleventh Circuits seem incomplete. A more complete analysis may have first acknowledged that each court, in effect, included the insurance proceeds in valuing the stock of each corporation but then removed the value of the insurance proceeds by offsetting the proceeds dollar-for-dollar against the perceived obligations to the decedent shareholder's estates. In short, by removing the value of the insurance proceeds from the value of the stock, the insurance proceeds have been completely excluded from the decedent majority shareholder's estate, notwithstanding that the decedent had incidents of ownership over the policy. Without a discussion of the regulatory exclusion of insurance proceeds and the history behind the exception, the explanation as to why the value of the insurance proceeds was offset dollar-for-dollar by each corporation's obligation to the decedent's shareholder's estate is at best incomplete and in both instances possibly inaccurate.

(c) The Analysis of the *Blount* and *Cartwright* Courts is Incorrect from an Accounting and Financial Perspective

In addition to incorrectly applying the Code and regulations to the circumstances in *Blount* and *Cartwright*, the conclusion reached by each of the courts that the obligation of BCC and CSB to redeem the shareholder should be offset dollar-for-dollar against the insurance proceeds is incorrect from an accounting and financial perspective. While a third party might value a target corporation by treating a redemption obligation as a liability, it is not at all appropriate for a shareholder to reduce the value of the company by the value of his or her shares.

The *Blount* court summarized its position when it stated that "[t]o suggest that a reasonably competent business person, interested in acquiring a company, would ignore a $3 million liability strains credulity

169. Treas. Reg. §§ 20.2042-1(c)(6), 20.2031-2(f); *see also* T.D. 7312, 1974-1 C.B. 277, 1974 IRB LEXIS 835, *2-3 (1974).

and defies any sensible construct of fair market value."[170] While it may be logical from an unrelated third-party purchaser's perspective to include the liability associated with a redemption agreement in its purchase price analysis, neither the *Cartwright* court nor the *Blount* court compared a redeemed shareholder's perspective to a third-party purchaser's perspective. However, the Tax Court's opinion in *Blount* (later reversed by the Eleventh Circuit) did contain specific and persuasive analysis regarding the effect of a redemption obligation on insurance proceeds received by a corporation.[171] In the Tax Court, Mr. Blount's estate argued that the court should treat BCC's enforceable $4 million dollar redemption obligation as a corporate liability in determining the value of the decedent shareholder's shares.[172] In making this argument, the estate recognized the liability would operate to offset the value of the insurance proceeds received by the corporation. By lowering the value of the company, the proportionate value of the decedent's stock would also be reduced. In declining to allow the offset, the Tax Court reasoned that the redemption obligation should not be treated as a value depressing corporate liability when the very shares that are the subject of the redemption obligation are being valued.[173] The Tax Court's holding would have resulted in a proportionate increase in the value of Mr. Blount's stock that would have increased Mr. Blount's gross estate for estate tax purposes.

In addressing the estate's argument, the Tax Court noted the distinction between a third-party purchase of the shares and the corporation's redemption of its shares from a shareholder and provided the following:

> A simplified example will illustrate the fallacy behind the estate's contention that BCC's obligation to redeem decedent's shares should be treated as a liability offsetting a corresponding amount of corporate assets. Assume corporation X has 100 shares outstanding and two shareholders, A and B, each holding 50 shares. X's sole asset is $ 1 million in cash. X has entered into an agreement obligating it to purchase B's shares at his death for $ 500,000. If, at B's death, X's $ 500,000 redemption obligation is treated as a liability of X for purposes of valuing B's shares, then X's value becomes $ 500,000 ($ 1 million cash less a $ 500,000 redemption obligation). It would follow that the value of B's shares (and A's shares) is $ 250,000 (i.e., one half of the corporation's $ 500,000 value) upon B's death. Yet if B's shares are then redeemed for $ 500,000, A's shares are then worth $

170. *Estate of Blount*, 428 F.3d at 1346.

171. Estate of Blount v. Comm'r, 87 T.C.M. (CCH) 1303, 2004 Tax. Ct. Memo LEXIS 117, *78–86 (2004), *aff'd in part and rev'd in part*, 428 F.3d 1338 (11th Cir. 2005).

172. *Id.* at *78.

173. *Id.* at *80.

> 500,000—that is, A's 50 shares constitute 100-percent ownership of a corporation with $ 500,000 in cash.[174]

The Tax Court then pointed out that it could not be correct that B's one half interest in one million dollars in cash was only worth $250,000, nor could it be correct that A's one half interest in the remainder shifted from a value of $250,000 pre-redemption to a value of $500,000 post-redemption.[175] Further, the error in relation to the valuation of B's shares in the example was the recognition of the corporation's redemption obligation as a claim on the corporate assets when valuing the same shares that would be redeemed with those assets.[176] By contrast, a hypothetical third-party buyer of A's shares would pay $500,000 for A's shares whether the redemption obligation existed or not.[177] This is because a third-party purchaser would take account of both the liability arising from the redemption obligation and the shift in the proportionate interest of A's shares as a result of the redemption.[178]

Treating the corporation's obligation to redeem a shareholder (or a shareholder's estate) as a liability in valuing the company for purpose of redemption would result in valuing the corporation in its post redemption configuration.[179] There is an important distinction between valuing stock to be redeemed from a shareholder and valuing stock prior to a purchase by a third party. With respect to valuing stock to be redeemed, the shareholder (or shareholder's estate) will always seek to receive his or her ratable share of the value of the company *previous* to the redemption.[180] Whereas, a third unrelated party would analyze the purchase of the stock held by the remaining shareholders based upon the value of the shares that remain *after* the redemption takes place.

So, why is it that the Tax Court decided in *Cartwright* that CSB's obligation to the decedent shareholder should offset the insurance proceeds received by the corporation while the Tax Court's decision in *Blount* denied a similar offset? That too was logically explained by the

---

174. *Id.* at *81-82 (footnotes omitted).
175. *Id.* at *82.
176. *Id.* at *82-83.
177. *Id.* at *83 n.36.
178. *Id.* at *83.
179. *Id.* at *79-80.
180. *See id.* at *80-83. In a more extreme example, it would make no sense that a 99% shareholder would base the value of his or her redemption on post-redemption value of 1%. On the other hand, a hypothetical third party willing buyer looking to purchase all of the stock of BCC prior to redeeming the decedent shareholder's estate would correctly consider the corporation's liability in determining the value of the shares. This is so because the obligation to redeem the decedent shareholder's estate would continue after the third party purchaser's acquisition of the stock. Such a hypothetical purchaser's stock value in BCC would decrease proportionately when the redemption occurred. Using the same example, where a third party seeks to purchase a company that is obliged to redeem a 99% shareholder. It would make no sense for the third-party acquirer to pay more than 1% of the pre-redemption value of the company. A third-party buyer will seek to value what he or she will have at the moment *after* purchase. Whereas, a shareholder that is to be redeemed will seek to value what he or she has a moment *before* the purchase.

Tax Court in its decision in *Blount*. As previously indicated, four-fifths of the liability in *Cartwright* was related to personal services provided by Mr. Cartwright to CSB before his death. This four-fifths portion of the liability was not a corporate obligation to redeem its stock.[181] Rather, it was a liability for services performed—not meaningfully different from any other liability of the corporation that would be netted against assets to ascertain the net value of the company.[182] Upon finding that approximately four-fifths of the obligation in *Cartwright* was not a redemption, the Tax Court concluded that the whole liability in Cartwright should be offset dollar-for-dollar against the insurance proceeds. Whereas, the Tax Court properly found that the obligation in *Blount* was a redemption, as opposed to an ordinary liability, which should not be an offsetting liability against the insurance proceeds.[183]

While this analysis explains why four-fifths of the liability in *Cartwright* should be a liability that is offset against the assets of the company (including any insurance proceeds received), the Tax Court failed to completely clear the air in relation to the remaining one-fifth of the liability related solely to redemption of Mr. Cartwright's shares held by his estate. The remaining one-fifth of the liability in *Cartwright* is indistinguishable from the whole liability in *Blount*. Recognizing this, the Tax Court conceded in its decision in *Blount* that the remaining one-fifth portion of the liability in *Cartwright* constituted an obligation to redeem stock.[184] Thus, it might have been appropriate for the Tax Court to have also conceded that the remaining one-fifth of the liability in *Cartwright* should not have been an obligation that was recognized in valuing the corporation.

iv. Summary: Incidents of Ownership Held by Corporations

Based upon the above analysis, it is clear that the manner in which the Treasury and the Tax Court value a decedent's ownership equity interest in a corporation are at odds with the manner in which the *Cartwright* and *Blount* courts have interpreted the Code and relevant Regulations. The Code generally requires a decedent to include in his or her gross estate the value of policy proceeds payable upon the decedent's death to the extent that the decedent held incidents of ownership in the insurance policy. However, the regulations allow a decedent majority shareholder's estate to exclude the value of insurance proceeds received by the decedent's wholly or majority owned corporation notwithstanding

181. *Id.* at *84.
182. *Id.*
183. *See id.* at *85-86; *see also* Huntsman v. Comm'r, 66 T.C. 861, 874 (1976) (indicating that insurance proceeds are treated like any other nonoperating asset when determining a closely held corporation's value); *Estate of Blount*, 2004 Tax Ct. Memo LEXIS 117, at *77.
184. *Estate of Blount*, 2004 Tax Ct. Memo LEXIS 117, at *84-85.

that the decedent could have exercised actual control over the insurance policy. The exclusionary exception provided by the regulations was designed to prevent double inclusion of the value of the insurance policies in the decedent's estate. By allowing an exclusion *where the corporation receives the policy proceeds*, the regulations contemplate inclusion of a single ratable share of the insurance proceeds via an increase in the decedent's stock interest. Thus, in order to include the value of the decedent's interest in the corporation, the value of the insurance proceeds must at least be included in the value of the company one time.

However, the Ninth Circuit's and the Eleventh Circuit's decisions in *Cartwright* and *Blount* thwart the above analysis. The *Cartwright* and *Blount* decisions require a dollar-for-dollar offset of the insurance proceeds received by a corporation against the corporation's obligation to redeem the decedent shareholder. This dollar-for-dollar offset reduces the overall value of the corporation when valuing the deceased shareholders shares. The effect of these holdings is to allow decedent sole or majority shareholders to avoid including a ratable share of the value of the insurance proceeds received by the corporation notwithstanding that such a shareholder may exercise actual control over the insurance policy held by the corporation. This outcome is contrary to the Code's overall requirement that such proceeds must be included in gross estate where the decedent has incidents of ownership in the policy. The outcome is also contrary to the goal of the regulations to include at least a ratable share of the proceeds in a deceased shareholder's estate.

b. Incidents of Ownership Held by Partnerships and LLCs

The above described inconsistency between the *Blount* and *Cartwright* opinions and the regulations under sections 2031 and 2042 also affects the valuation of a deceased partner or member's equity interest in a partnership or LLC where a redemption agreement is in place. In general, partnerships may be treated both as an "entity" for federal income tax purposes and, under other circumstances, partnerships may be considered as an "aggregate" of individuals each treated as directly owning an interest in the partnership assets and operations.[185]

If a partnership or LLC is considered as an entity separate from its partners or members under the entity theory of partnership taxation, the partnership itself, instead of the partners individually, would have all the incidents of ownership.[186] For purposes of the following discussion, it will be assumed that an LLC is treated as a "partnership" for federal income tax purposes and that the members of an LLC are "partners" for purposes of analysis. In relation to valuation for purposes of estate taxation of a deceased partner's interest in a partnership, the entity theory of

185. *See* MCKEE ET AL., *supra* note 31, ¶ 1.02[1].
186. *Id.*; *see also* I.R.C. § 701; MCKEE ET AL., *supra* note 31, ¶ 1.02[2].

partnership taxation would appear to be analogous to the manner in which the shareholders of a corporation would be taxed, as previously discussed. Except in certain rare circumstances, subchapter K of the Code generally imposes tax on the individual partners rather than the partnership.[187]

A different outcome might be obtained under the aggregate theory of ownership. If a general partnership owns an insurance policy, incidents of ownership appear to be attributed to an aggregate of the general partners and, therefore, to each one individually.[188] This rule would attribute to each general partner without regard to the percentage of ownership in the partnership, incidents of ownership in policies of insurance held by the partnership on a partner's life. Under the aggregate theory, several minority general partners might each be treated as having incidents of ownership in an insurance policy. Thus, where a minority corporate shareholder is clearly not attributed incidents of ownership in an insurance policy held by the corporation, it is unclear whether a minority general partner should be attributed incidents of ownership. Further questions arise in relation to treatment of limited partners and members of an LLC.

i. Where Proceeds are Payable to Partnership or LLC

Under the aggregate and entity theories of partnership taxation, the question arises as to whether partners of a partnership or members of an LLC should be attributed incidents of ownership in a policy owned by the entity. Cases and administrative rulings have addressed this issue without complete agreement on the treatment of incidents of ownership in the partnership setting. Arguably, at least one similarity exists in the treatment of incidents of ownership in the corporate and partnership settings. If the proceeds of an insurance policy on the life of a deceased partner are payable to or for the benefit of the partnership, then such proceeds should not be included in the gross estate under section 2042.[189] Of course, this similarity immediately calls into question whether the *Blount* and *Cartwright* opinions impact valuation of a decedent partner's equity interest where a redemption agreement is in place between the decedent and the partnership. Beyond this similarity, several inconsistencies arise in the treatment of corporate shareholders as compared to partners.

In 1955, the Tax Court analyzed and discussed incidents of ownership in an insurance policy held by a general partnership in *Estate of*

187. *See* I.R.C. §§ 701-777.

188. Rev. Rul. 83-147, 1983-2 C.B. 158.

189. *See id.*; *see also* I.R.S. Gen. Couns. Mem. 39,034 (Sept. 21, 1983); *see also* ZARITSKY, *supra* note 43, ¶ 8.02[4][b] (indicating that the IRS has not indicated formally whether it will apply this rule when the insured is only a limited partner).

Knipp v. Commissioner.[190] In *Knipp*, the taxpayer was the estate of a deceased partner, Mr. Knipp.[191] At the time of his death on November 21, 1947, Mr. Knipp had eleven insurance policies outstanding on his life, ten of which were assigned to the partnership prior to his death.[192] The partnership owned the policies and Mr. Knipp had no right to change the beneficiary designation on the policies.[193] The value of each of the policies was entered on the books of the partnership as an asset and, thereafter, the partnership pledged the policies as collateral for loans.[194] "The partnership paid the premiums on the policies . . . [and] [t]he increase of cash surrender value of the policies each year was considered income of the partnership."[195]

The Commissioner argued that all of the proceeds of the insurance policy paid to the partnership should be directly included in Mr. Knipp's gross estate under section 2042 because, as general partner, Mr. Knipp possessed incidents of ownership in the policies at the time of his death.[196] The court disagreed, finding that because Mr. Knipp's interest in the partnership never exceeded fifty percent (50%), the premium payments were made by the partnership, the insurance proceeds were payable to the partnership, and the policies were assets of the partnership.[197] In support of its conclusion, the court found that the partnership had complete control over the policies and the decedent had no rights in the policies other than those flowing from his partnership interest.[198] The Commissioner conceded that the insurance policy was an asset of the partnership and that Mr. Knipp had no rights in the policies other than those flowing from his partnership interest.[199] Based upon its finding that the partnership controlled the policies, the court also found that the partnership held all incidents of ownership in the policies.[200] The court held that the partnership's incidents of ownership in the policies insuring the life of the decedent were not attributed to the decedent, and the insurance proceeds were not includible in the decedent's gross estate.[201] The aggregate and entity theories of partnership taxation were not explicitly applied for purposes of attributing incidents of ownership in an insurance policy owned by the partnership to either the entity or the partners as a group. However, the outcome of *Knipp* would appear to be more consis-

190. 25 T.C. 153, 154 (1955), *acq.*, 1956-2 C.B. 6, *nonacq.*, 156-2 C.B. 10, *aff'd on other grounds*, 244 F.2d 436 (4th Cir. 1957).
191. *See Knipp*, 25 T.C. at 154.
192. *Id.* at 157.
193. *See id.* at 168-69.
194. *Id.* at 157.
195. *Id.*
196. *Id.* at 166.
197. *See id.* at 167-68.
198. *See id.* at 167-69.
199. *Id.* at 168.
200. *Id.* at 167.
201. *Id.* at 169.

tent with the application of an entity theory of partnership taxation where incidents of ownership are concerned. Further, this outcome comports with the manner in which corporations and their shareholders are taxed in relation to the proceeds of an insurance policy on the life of a shareholder.

Like a majority shareholder in a corporation, the court's holding that Mr. Knipp was not attributed incidents of ownership for purposes of section 2042 allowed Mr. Knipp's estate to exclude the proceeds in calculating his gross estate. But the holding in *Knipp* does not allow the taxpayer's estate to completely escape estate taxation of the insurance proceeds. Under the circumstances in *Knipp*, the value of the decedent's partnership units would appear to increase proportionately due to the partnership's receipt of the insurance proceeds.[202] Thus, the holding in *Knipp* accords the same treatment to a 50% general partner as the regulations accord to a majority shareholder in a corporation.

Unfortunately, however, the holding in *Knipp* did not fully resolve the question of attribution of partnership incidents of ownership in an insurance policy owned by a partnership. In General Counsel Memorandum 39034 (the "Memorandum"), the IRS analyzed the facts of proposed Revenue Ruling 83-147 and provided a detailed comparison of the treatment of incidents of ownership between partnerships and corporations.[203] The Memorandum addressed the facts of the proposed rules whereby C, D, and E are equal minority general partners of the XYZ partnership. The XYZ partnership obtained a life insurance policy on the life of D and designated D's child A as the beneficiary.[204] Thereafter, the premium payments were made by the XYZ partnership.[205] The issue for consideration in the proposed ruling was whether a partner possesses incidents of ownership as an insured in a life insurance policy held by the partnership, where the proceeds are payable *not to the partnership but, rather, to a third unrelated party.*[206]

The IRS first reasoned in the Memorandum that although the facts involve partnership-owned life insurance, the treatment of corporate-owned life insurance is relevant.[207] Further, the IRS noted that "[p]artnership-owned life insurance 'could' be given treatment analogous

202. *See* I.R.C. § 2031(a).

203. *See* I.R.S. Gen. Couns. Mem. 39,034 (Sept. 21, 1983). By way of a memorandum, dated December 22, 1981, the Director of the Individual Tax Division of the requested that Chief Counsel consider two proposed revenue rulings, 83-147 and 83-148, referred to in the G.C.M. as Proposed Ruling A and Proposed Ruling B, respectively. *Id.* Ultimately, Proposed Ruling A was published as Revenue Ruling 83-147. *Compare* I.R.S. Gen. Couns. Mem. 39, 034 (Sept. 21, 1983), *with* Rev. Rul. 83-147, 1983-2 C.B. 158.

204. I.R.S. Gen. Couns. Mem. 39,034, at *2.

205. *Id.*

206. *Id.* at *1.

207. *Id.* at *4.

to corporate-owned life insurance."[208] The Memorandum again points out that a corporation's incidents of ownership are not attributed to a "controlling" shareholder where the proceeds are payable to the corporation.[209] The Service reasoned that this prevents the proceeds from being considered both as a factor in valuation of the decedent's stock and as a separate asset in the decedent's gross estate.[210] The IRS noted that this outcome was specifically intended by the drafters of Regulation section 20.2042-1(c)(6).[211] Citing *Knipp,* the IRS indicated that it acquiesced in and agreed with the holding in *Knipp* only where the insurance proceeds were payable to the partnership and where inclusion of the insurance proceeds would result in double inclusion of a substantial portion of the proceeds.

Based upon this reasoning, the Service indicated that it agreed with the conclusion in the proposed ruling that where:

> a partnership owns a life insurance policy on a partner's life and the proceeds are payable *other than to or for the benefit of the partnership*, [an] insured partner possesses incidents of ownership in the policy in conjunction with the other partners, that require direct inclusion of [all] the proceeds in the insured partner's gross estate under section 2042(2).[212]

In this regard, the Service noted that the proposed ruling treats "a partnership as an aggregate of the individual partners . . . where the insurance proceeds" are not payable to the partnership.[213] As such, the insurance proceeds are all directly includible in the decedent shareholder's gross estate under section 2042(2).

For purposes of section 2042(2), the Memorandum indicates that a partnership is treated "as an aggregate of individual partners."[214] Further, the Memorandum notes that "incidents of ownership will not be attributed to partners where the insurance proceeds are payable to or for the benefit of the partnership."[215] Under these circumstances, attribution of the partnership's incidents of ownership to the insured partner "would result in double taxation of the . . . proceeds."[216] The Service then stated that its position was to avoid such double inclusion by not including proceeds payable to the partnership directly in the decedent's estate under

---

208. *Id.* at *6 (citing Treas. Reg. § 20.2042-1(c)(6) (as amended in 1974)).
209. *See id.* at *5-6.
210. *Id.* at *7-8.
211. *Id.* at *5 (citing Memorandum from Commissioner on Amendment of Estate Tax Regulations (Feb. 28, 1974), 1974 TM LEXIS 84, *1-3).
212. *Id.* at *1-2 (emphasis added).
213. *See id.* at *7.
214. *Id.*
215. *Id.*
216. *Id.* at *7-8.

section 2042(2).[217] Although not specifically stated in the Service's reasoning, this conclusion would appear to indicate that where the proceeds are payable to the partnership, the partnership will be treated as an entity for federal tax purposes.

Post *Cartwright* and *Blount*, based upon *Knipp* and the Service's acquiescence to *Knipp* in the Memorandum, the same concern arises in the area of partnerships that exists in the corporate arena only to a much greater degree. The rulings in *Cartwright* and *Blount* allow insurance proceeds received by a corporation to be offset against a perceived obligation on the part of the corporation to redeem an equity interest from the decedent corporate equity holder. There would appear to be no logical reason why the Ninth and Eleventh Circuits, or any court following the decisions in *Cartwright* and *Blount*, would not apply the same reasoning to a redemption by a partnership or LLC of a decedent partner's or member's interest.

It is true that the *Cartwright* and *Blount* courts each based their holdings on regulations under sections 20.2031-2 which specifically relate to valuation of stocks of a corporation.[218] However, regulations under section 20.2031-3 address the valuation of interests in businesses and require taxpayers to determine the fair market value of a partnership in a manner consistent with the valuation of a corporation. Indeed, the net value of a partnership is determined on the basis of all relevant factors including, among other things, the factors set forth in paragraph (f) of Regulation section 20.2031-2 relating to the valuation of corporate stock.[219]

Thus, it would appear that the holding in *Blount* would equally apply to a partnership redemption causing a dollar-for-dollar offset of insurance proceeds received by a partnership or LLC on a policy insuring the life of a partner or member against the liability of the partnership or LLC to redeem a partner or member under a redemption agreement.

ii. Inconsistent Treatment of Minority Partners and Members of LLCs Where Insurance Proceeds Payable Outside the Partnership or LLC

Importantly, however, the Service went on in the Memorandum to make a distinction between a controlling shareholder and a partner under the circumstances of the proposed ruling. The Memorandum indicates that the proposed ruling takes a different approach than the regulations as they apply to corporations where the proceeds are payable outside the

217. *Id.* at *8.
218. *See* Treas. Reg. § 20.2031-2 (as amended in 1976).
219. *See* Treas. Reg. § 20.2031-3(c) (as amended in 1992); *see also id.* § 20.2031-2(f) (as amended in 1976).

partnership.[220] As previously discussed, the regulations under section 2042 provide that if a decedent corporate shareholder owns only a *minority* interest in a corporation, no incidents of ownership held by the corporation are attributable to the decedent shareholder and no proceeds are directly included in the decedent shareholder's estate.[221] However, unlike the case of a minority corporate shareholder, the Memorandum does not apply the same rule allowing the insurance proceeds to escape estate taxation where decedent is a minority general or limited partner or a minority member of an LLC.[222]

In addressing whether incidents of ownership in the policies would be attributed to the minority general partner, the Service conceded that partnership-owned life insurance could be given treatment analogous to corporate-owned life insurance in accordance with the regulations.[223] However, if treated the same as a shareholder-corporation arrangement, a partnership's incidents of ownership would only be attributed to a *majority* partner where the policy proceeds were payable to a third party. Whereas a *minority* partner, in theory, would not be attributed incidents of ownership.

However, the Memorandum indicates that a general partner insured under a policy held by a general partnership possesses incidents of ownership that are exercisable in conjunction with other general partners.[224] This broad statement attributes incidents of ownership in an insurance policy held by the partnership to any general partner regardless of the fact that such general partner is a majority or minority interest holder in the partnership. Such a conclusion would appear to be at odds with the Tax Court's holding in *Knipp*. Thus, the IRS appears to have taken a different approach in the Memorandum based upon the premise that "a partnership is generally regarded as an aggregate of its individual partners."[225] The Memorandum states that "[a]ny incidents of ownership in a life insurance policy held by the partnership are [in reality] held by the partners as individuals."[226] Based upon this reasoning, any policy proceeds payable to a third party are includible in the insured partner's gross estate under section 20.2042(2) regardless of whether the partner is a minority or majority owner of the partnership.[227] This conclusion appears to ignore the notion that "the term 'incidents of ownership' is not limited in its meaning to ownership of the policy in a technical legal

220. *See* I.R.S. Gen. Couns. Mem. 39,034, at *6-7.
221. *Id.*
222. *See id.* at *4-8.
223. *Id.* at *6.
224. *See id.* at *7 (citing STEPHENS ET AL., *supra* note 73, ¶ 4.14[5][b], *available at* 1999 WL 1031619).
225. *Id.* at *7.
226. *Id.* at *23.
227. *Id.* at *22-23.

sense."[228] "Generally speaking, the term has reference to the right of the insured or his estate to the economic benefits of the policy."[229] Further, as previously discussed, the conclusion also ignores the provisions under the regulations that attribute incidents of ownership in a corporate setting only to a controlling shareholder.[230] A minority shareholder of a corporation is not attributed incidents of ownership under any circumstances whereas a minority partner is attributed incidents of ownership where the policy proceeds are payable other than to the partnership.

In light of the above distinction, a question remains in relation to whether a limited partner or a member of an LLC would be attributed incidents of ownership in a policy held by the entity. In 2001, the IRS addressed a fact pattern which included a contribution of an insurance policy to a limited partnership. In Private Letter Ruling 200111038, two grantors (A and B) of two separate trusts (Trust A and Trust B) formed a new limited partnership (LP).[231] Trust A was to contribute insurance policies to LP in exchange for a limited partnership interest.[232] Trust B was to contribute cash in exchange for a general partnership interest in LP.[233] Grantors A and B also contributed cash in exchange for a limited partnership interest in LP.[234] Under the LP agreement, all taxable income and losses were to be allocated in accordance with the partners' interests in the partnership.[235] Further, the agreement provided that general partners had sole management authority over the partnership and limited partners had no right to participate in management or investment decisions including any decisions in relation to the LP's ownership of the insurance policies.[236] Importantly, however, the proceeds of the acquired insurance policies were to be paid to the LP.[237]

The ruling addressed the issue of whether, for purposes of section 2042, a limited partnership interest would be treated as being sufficient control such that the *limited partners* would be treated as possessing incidents of ownership by virtue of their interest in such an entity.[238] The IRS ruled that where the terms of a partnership agreement preclude the limited partners from exercising any control over the management and day-to-day affairs of the limited partnership or take part in the vote in respect to the limited partnership's management and operations, the limited partners will not possess incidents of ownership under section 2042

228. Treas. Reg. § 20.2042-1(c)(2) (as amended in 1974).
229. *Id.*
230. Treas. Reg. § 20.2042-1(c)(6).
231. I.R.S. Priv. Ltr. Rul. 200111038, at *8 (Dec. 15, 2000).
232. *Id.* at *9.
233. *Id.*
234. *Id.*
235. *Id.*
236. *See id.* at *9-11.
237. *Id.* at *10.
238. *See id.* at *1.

with respect to insurance policies held by the limited partnership.[239] Thus, in an initial analysis, it would appear that the question turned upon whether the limited partners (or presumably a non-managing member of an LLC) had "control" over the policy for purposes of incidents of ownership.

However, this line of reasoning loses much of its force due to the fact that the IRS then fell back on its analysis in Revenue Ruling 83-147. The IRS indicated that its conclusions were due to the fact that the policy proceeds were payable directly to the limited partnership and not to a third party.[240] Thus, because the proceeds are payable to the partnership, the proceeds will increase the value of the decedent limited partner's interest in the limited partnership and, therefore, no incidents of ownership held by the partnership should be attributed to the decedent limited partner.[241] Even though the reasons given in the ruling address the lack of the limited partners' control over the policies in relation to section 2042, the fact that the limited partnership received the proceeds makes the ruling less intriguing and less valuable. Again, the question of how insurance proceeds on a policy held by a limited partnership or LLC payable to a third party will be treated for purposes of section 2042 specifically remains unaddressed. Further, the question of whether a partner's status as a limited partner alone will result in a sufficient lack of control such that the limited partner will not be attributed incidents of ownership remains unanswered. If the analysis turned solely on "control," a limited partner (majority or minority) would not appear to have any incidents of ownership regardless of whether the proceeds were payable to a third party.

III. RECOMMENDATIONS FOR A SOLUTION TO THE PROBLEM

There are two basic problems in relation to application of the rules under section 2042 of the Code. First, the *Cartwright* and *Blount* decisions allow taxpayers to completely avoid inclusion of insurance proceeds payable to a corporation where there is a redemption agreement in place that requires the corporation to redeem a deceased shareholder's equity interest in the company. Second, the IRS has created an inconsistency in applying the regulations under section 2042 attributing incidents of ownership in an entity-owned life insurance policy where the policy proceeds are payable to a third party. The inconsistency results where a minority limited partner is attributed incidents of ownership while a minority corporate shareholder is not. Both of these inconsistencies should be addressed by Congress, the courts, and the Treasury Department.

---

239. *See id.* at *19-20.
240. *See id.* at *18-19.
241. *See* I.R.S. Gen. Couns. Mem. 39,034, at *7-8.

Of primary importance is the anomaly created by the *Cartwright* and *Blount* opinions which have created precedent that can be relied upon by estates of deceased shareholders to support a complete exclusion of the value of insurance proceeds received by the corporation. This loophole should be closed. The policy behind the exception to inclusion provided for in the regulations under section 2042 is premised upon ratable inclusion of the proceeds under section 2031. This policy is thwarted by the Eleventh Circuit's decision in *Blount* and by the Ninth Circuit's decision in *Cartwright.*

A. Amend Section 2031 Regulations to Require Inclusion of Life Insurance Proceeds in Valuation

One possible solution would be for the IRS to amend the regulations under section 2031 to clarify that life insurance proceeds payable to or for the benefit of a company shall, to the extent the proceeds are excluded by the regulations under section 2042, be included for purposes of determining the value of the company. Specifically, the fourth full sentence of Regulation section 20.2031-2(f) could be amended to include following additional language:

> In addition to the relevant factors described above, consideration shall also be given to non-operating assets, including proceeds of life insurance policies payable to or for the benefit of the company, to the extent that such non-operating assets have not been taken into account in determining the net worth, prospective earning power and dividend earning capacity. *Such non-operating assets shall include, among other things, proceeds of a life insurance policy payable to or for the benefit of the company which proceeds have not otherwise been included in the insured shareholder's estate under section 2042 or the applicable regulations thereunder.*

The above amended section of the regulations would clarify that insurance proceeds received on the life of an insured shareholder would be included for purposes of valuing the decedent's stock in the company. This amendment would require inclusion of the insurance proceeds in an attempt to prevent an interpretation such as, for example, the interpretations in *Cartwright* and *Blount*, that the insurance proceeds are not the kind of non-operating asset that should be included in the value of the company under the regulations.[242] However, the holding in *Blount* could continue to create ambiguity because the court's holding also requires that where a redemption agreement is in effect between the company and the shareholders, the company's obligation under the agreement should offset the insurance proceeds "dollar-for-dollar."[243] This second requirement, in effect, would undermine the effect of the proposed

242. *See* Estate of Blount v. Comm'r., 428 F.3d 1338, 1345-46 (11th Cir. 2005).
243. *See id.* at 1346.

amendment by leaving open the question of whether an estate must take the company's redemption obligation into consideration when valuing a decedent shareholder's stock.

The requirement that the company's redemption obligation be taken into consideration by the redeemed shareholder's estate in valuing the stock for purposes of determining gross estate is inappropriate for several reasons. First, such a requirement would result in undervaluing the redeemed shareholder's stock in virtually every instance due to the fact that a redeemed shareholder would never concede to selling his or her shares back for less than an arm's length consideration. Further, valuation based upon post-redemption assets is not arm's length for reasons explained by the Tax Court in its decision in *Blount*.[244] This portion of the holding by the Eleventh Circuit would not likely be addressed by the IRS in regulations but, rather, left to the courts to correct in a later case.

B. Amend Regulations Under Section 2042 to Require Attribution of Incidents of Ownership to Majority Shareholders, Partners, and Members of LLCs

Because of the inherent ambiguity in interpreting the court's holding in *Blount*, a larger, more expansive change in the Code and Regulations may be appropriate. Accepting the *Blount* court's holding that the insurance proceeds should not be included in the valuation calculation, it may be appropriate for the Treasury to reconsider the requirements of section 2042 and promulgate regulations that pertain to shareholders, partners, and members of LLCs that attribute incidents of ownership to a "controlling" equity holder.

Under current regulations, it appears that the Treasury envisioned that estates of deceased shareholders of corporations should be required to include only a ratable portion of the insurance proceeds received by the corporation. This policy is implemented via the regulations which prevent attribution of incidents of ownership to a sole or controlling shareholder where the economic benefits of a life insurance policy on the life of the shareholder are reserved to the corporation.[245] In turn, this allows the shareholder to exclude from the gross estate a portion of the value of the life insurance proceeds notwithstanding that such shareholder actually had control over the policy.

Historically, inclusion of the insurance proceeds via attribution of incidents of ownership to a controlling decedent shareholder resulted in a double inclusion of the proceeds. This outcome resulted because the value of the corporate shares in the decedent's estate was thought to increase proportionately due to the corporation's receipt of the insurance

244. *See id.* at 1343-45.
245. *See* Treas. Reg. § 20.2042-1(c)(6) (as amended in 1974).

proceeds. However, under the *Cartwright* and *Blount* decisions, the proceeds are no longer subject to double inclusion due to the fact that the proceeds are offset dollar-for-dollar by the company's obligation to redeem the deceased shareholder's stock.[246] Thus, it may be appropriate for the Treasury to consider whether to amend the current regulations under section 2042 to provide that a sole or controlling shareholder, partner or member of an LLC will be attributed incidents of ownership where the equity holder has sufficient ownership interest. This outcome would appear to be consistent with Congress' intent to impose estate tax on decedents who have the ability to, for example, change the beneficiary designation on the policy held by the entity.

C. Amend Regulations to Treat Shareholders, Partners and Members Consistently with Respect to Attribution Rules

The above problems stem from a set of circumstances wherein the company receives the proceeds from an insurance policy which insures the life of an equity holder. A separate and different inequity arises where the proceeds of a company-owned insurance policy are payable to a third party. It makes little sense that a minority corporate shareholder should be treated differently than a minority partner or minority member of an LLC where the proceeds of an insurance policy are held by the entity but payable to a third party.

The advent, evolution, and increased popularity of limited liability companies require that the regulations under section 2042 that apply to corporations be reassessed in an effort to provide similar treatment to partners of partnerships and members of LLCs. The general policy behind the section 2042 regulations, directing the proceeds of a corporate-owned insurance policy payable to a third unrelated party be excluded from a minority shareholder's gross estate should apply to minority partners, limited partners, and non-managing members of LLCs.

However, the Treasury rulings and agency memorandums indicate unwillingness on the part of the IRS to treat minority partners and minority members of LLCs in a similar fashion. Instead, the Treasury views the partners or LLC members that are treated as partners for tax purposes as being an aggregate of partners each of whom individually has control over incidents of ownership of a policy held by the partnership or LLC. While the general theory of aggregate versus an entity approach to partnerships may be relevant to the analysis, there is a very large unaddressed grey area between the rights of shareholders and the rights of limited partners and non-managing members of an LLC. More specifically, where a limited partnership or LLC agreement authorizes certain partners or members to manage the assets, including insurance policies, held by the company, attribution of incidents of ownership to such part-

246. *See Estate of Blount*, 428 F.3d at 1346.

ners or members would be appropriate in that they have actual control. However, there appear to be many instances in which a limited partner or non-managing member of an LLC would have little if any control over a policy owned by the entity. A minority limited partner and a minority member of an LLC who has no management authority over a company-held insurance policy would, like a minority shareholder of a corporation, have no control over such a policy and there would be little reason to attribute incidents of ownership to such partners or members of an LLC.

As a result, the Treasury's rulings and memorandums that require a deceased minority limited partner or deceased minority member of an LLC to include the whole value of an insurance policy in gross estate notwithstanding that the proceeds are paid to an unrelated third party are likely unsupportable and outdated. The Treasury should embark on a regulation project to analyze and propose amendments to or new regulations which focus on the core control requirement of section 2042. Attribution of incidents of ownership should be limited only to those equity holders of a company that have actual control over the policy or have the independent ability to control the ownership rights of an insurance policy held by the company.

IV. SUMMARY & CONCLUSION

Generally, section 2042 requires a decedent who possesses incidents of ownership in an insurance policy on his or her life to include in the decedent's gross estate all insurance proceeds receivable under such policy.[247] A decedent possesses incidents of ownership in a policy to the extent that the decedent can, for example, change the beneficiary designation or otherwise has rights to the economic benefits of the policy.[248] A sole or majority shareholder has the power to control the beneficiary designation of a policy owned by the corporation. Thus, a sole or majority shareholder generally would be attributed incidents of ownership in a corporate-owned insurance policy. However, current regulations under 2042 provide an exception whereby a decedent sole or majority shareholder will not be attributed incidents of ownership in such a policy where the proceeds are payable upon death directly to the corporation.[249] The purpose of this exception is to prevent the double inclusion of the value of the insurance proceeds once through section 2042 by attribution to the majority shareholder via incidents of ownership and again through an increase in the value of the shares of stock that are required to be included in gross estate under section 2031.

247. *See* I.R.C. § 2042.
248. *See* Treas. Reg. § 20.2042-1(c)(2) (as amended in 1974).
249. *See* Treas. Reg. § 20.2042-1(c)(6).

In relation to partnerships and limited liability companies, the cases and rulings have resulted in a similar exception with no attribution of incidents of ownership to partners or members where proceeds of a policy insuring the life of a decedent partner or member are payable to the entity. The Tax Court and the Treasury's rulings are supported by the same view that to require inclusion under section 2042 would result in unnecessary double inclusion of the value of the insurance proceeds in the decedent partner or member's estate.

Until recently, attribution of incidents of ownership and estate taxation of the receipt of policy proceeds by either an entity or the equity holder of such entity under the above described circumstances was largely undisputed under the cases, rulings and regulations. However, in 1999, the Ninth Circuit Court of Appeals in *Cartwright* found that proceeds received by a corporation from a policy insuring the life of the corporation's majority shareholder should not increase the value of the decedent's shares for purposes of determining the decedent's gross estate.[250] The court came to this conclusion based largely upon the fact that the value of the insurance proceeds was offset by an existing liability to the decedent shareholder's estate for services performed for the corporation prior to the decedent's death.[251] A small portion of the insurance proceeds received by the corporation were used to redeem the majority shareholder's stock pursuant to the terms of a stock redemption agreement that existed between the majority shareholder and the corporation.[252] The effect of the court's holding was to allow the shareholder's estate to completely avoid inclusion of the value of the insurance proceeds received by the corporation. Ignoring the small portion of the payment related to the redemption agreement, the outcome of the *Cartwright* case was consistent with the policy behind the estate tax provisions of the Code.

However, in October of 2005, the Eleventh Circuit Court of Appeals in *Blount* followed the decision of the Ninth Circuit in *Cartwright*. While the court's decision in *Blount* purports to follow the reasoning of the Ninth Circuit Court of Appeals in *Cartwright*, it appears to misapprehend the reasoning of the Tax Court which was affirmed by the Ninth Circuit in *Cartwright*. By holding that the liability to redeem the shares of the decedent shareholder operates to offset dollar-for-dollar the insurance proceeds received by the corporation for purposes arriving at the redemption value of a deceased shareholder's stock, the *Blount* court ignores the requirements of section 2042 and the regulations thereunder. In doing so, the holding of the Eleventh Circuit Court of Appeals in *Blount* completely removes the insurance proceeds in determining the

250. *See* Estate of Cartwright v. Comm'r., 183 F.3d 1034, 1038 (9th Cir. 1999).
251. *See Estate of Cartwright*, 183 F.3d at 1038.
252. *Id.* at 1035-37.

value of the shares of the corporation in the decedent shareholder's estate under section 2031. By removing the insurance proceeds from the calculation of the value of the decedent shareholder's stock under section 2031, no portion of the insurance proceeds are included in the decedent's estate.

The exception to inclusion of the proceeds in the regulations under section 2042 is based upon the premise that the increase in value of the stock due to the corporation's receipt of the insurance proceeds will result in a proportionate increase in the value of the corporate stock to be included in the shareholder decedent's gross estate under section 2031.[253] Consequently, the *Blount* opinion and the *Cartwright* opinion to a certain degree, appear to allow exclusion of the value of the proceeds from the section 2031 calculation. This, in turn, appears to create an apparent loophole that allows taxpayers to avoid estate tax payable on proceeds received from policies where such taxpayers clearly are attributed incidents of ownership. Further, there is no reason why the same anomalous result would not apply in the redemption of a partner or a member of an LLC where the entity received the policy proceeds.

In addition to the apparent loophole created by the *Cartwright* and *Blount* opinions, the IRS has created a distinction in the manner in which minority corporate shareholders and minority partners are treated when a third party, as opposed to the entity, receives insurance proceeds on the life of one of the equity holders. In a corporate setting under these circumstances, a decedent minority shareholder is not attributed incidents of ownership in the policy held by the corporation. By not attributing the corporate incidents of ownership to the decedent minority shareholder, no amount of the insurance proceeds received by the third party is required to be included in his or her estate. This same rule is not applied to minority partners or LLC members. Under these same circumstances, the minority member or partner is attributed incidents of ownership. As such, the draconian opposite outcome occurs requiring the estate of the decedent minority partner or member to include **all** of the value of the insurance proceeds in gross estate.

Given the widespread implementation of redemption agreements by shareholders and their closely held entities, there is a need to address the problems created by the holding of the Eleventh Circuit Court of Appeals in *Blount* and the Treasury's disparate treatment of minority corporate shareholders, minority limited partners and minority members of LLCs. Several possible solutions exist. First, the Treasury should amend the regulations to clarify that insurance proceeds received by a company on the life of a majority equity holder should be included in the equity holder's estate when valuing the equity interest. Although this outcome

253. *See* Treas. Reg. § 20.2042-1(c)(6) (as amended in 1974).

was intended by the Treasury and is explained in various Treasury pronouncements, the regulations are sufficiently ambiguous that two federal courts of appeal did not interpret the regulation as it was originally intended to operate.

In the alternative, the Treasury may wish to accept the treatment espoused by the courts of appeal and amend the regulations under section 2042 to require attribution of incidents of ownership to majority shareholders, partners, and members of LLCs. This alternative is in many ways more consistent with Congress' original intent when it enacted section 2042. Whereas, under current regulations, it was the intent of the Treasury to include only a ratable portion of the proceeds, this alternative would require majority equity holders in closely held entities to include in gross estate the whole amount of the insurance. Majority owners of entities that own insurance policies in fact have control of insurance policies held by their companies. It is for this reason that this alternative may be a viable solution.

Finally, the Treasury should equalize the treatment of minority shareholders with the treatment under the regulations of limited partners and non-managing members of LLCs. Under current regulations, where policy proceeds are payable to unrelated third parties outside of the company, minority shareholders are not attributed incidents of ownership under the rules. The Treasury's theory behind this treatment is premised upon the fact that a minority shareholder does not have actual control over the policy benefits. On the other hand, limited partners and non-managing members of LLCs are attributed incidents of ownership under the current rulings. This effectively requires such equity holders to include all of the policy proceeds in their estates upon their deaths. Such treatment is unwarranted and must be addressed by the Treasury in a manner that is consistent with the treatment accorded to corporate minority shareholders.

THE NEW "FETAL PROTECTION": THE WRONG ANSWER TO THE CRISIS OF INADEQUATE HEALTH CARE FOR WOMEN AND CHILDREN

LINDA C. FENTIMAN†

In 1999, Regina McKnight, a homeless, mentally retarded woman who was pregnant and addicted to cocaine, was charged with murder when her child was stillborn. The South Carolina Supreme Court affirmed her murder conviction and upheld the twenty-year sentence imposed.[1]

In 2002, a severely mentally disabled woman became pregnant after being raped by the owner of the group home where she lived. The wife of a Florida prosecutor sought to be appointed "guardian of the fetus" in order to prevent the woman from taking prescription drugs necessary to maintain her physical health and mental stability and to prevent the woman from having an abortion. Ultimately, the Florida courts rejected these efforts.[2]

In 2004, Melissa Rowland, a pregnant woman with a long history of mental illness, sought assistance at a hospital because she noticed a decrease in fetal movements. Doctors recommended a Caesarean delivery, but Rowland declined, and the hospital offered no other help. When one of the twins she was carrying was stillborn, Rowland was charged with murder, with prosecutors asserting that she had acted with depraved indifference to the value of human life.[3]

---

† Professor, Pace University School of Law; B.S. Cornell University, J.D. S.U.N.Y. Buffalo Law School, LL.M. Harvard University. I wish to thank all of those who made this article possible, including the students in my criminal law course, whose work on an appellate brief project helped me appreciate the significance of this problem. I also want to thank those colleagues who gave generously of their time and comments, including Adele Bernhard, Kathleen Boozang, Georganne Chapin, Elena Cohen, Deborah Denno, Jill Gross, Frances Miller, Margaret Moreland, Lynn Paltrow, Elizabeth Rapaport, Audrey Rogers, and Kelly Weisberg. I am very grateful for the energetic and unflagging research assistance of Malisa Chokshi, Briana Fedele, Nicole Giordano, Carly Grant, Heather Ingle, Jennifer Kim, Christine Love, Jennifer McAdam, Lauren Maier, Jennifer Ramme, Devon Towner, and Jennifer Turchetta, Pace University Law School alumni and students.

1. State v. McKnight, 576 S.E.2d 168, 171 (S.C. 2003).

2. *In re* Guardianship of J.D.S., 864 So. 2d 534, 536 (Fla. Dist. Ct. App. 2004); Bob Mahlburg, *Senate Chief Wary of Fetus Guardian Bill; Governor Finds House Support, Won't Back Down on Legislation*, ORLANDO SENTINEL, Jan. 28, 2004, at B1.

3. Pamela Manson, *Mother is Charged in Stillborn Son's Death; Criminal Homicide: Prosecutors Say the West Jordan Woman Ignored Numerous Warnings from Doctors and Refused a Surgery that Could Have Saved the Boy's Life; Prosecutors Say Mom Guilty in Baby's Death*, SALT LAKE TRIBUNE, Mar. 12, 2004, at A1; Linda Thomson, *Rowland Case Is Called 'Political,'*

In roughly two-thirds of the states, women who write advance directives to guide their medical care should they become incompetent may have their directives rendered unenforceable if they become pregnant.[4]

TABLE OF CONTENTS

DESERET MORNING NEWS March 13, 2004; Linda Thomson & Pat Reavy, *Rowland's Out of Jail, Heading to Indiana*, DESERET MORNING NEWS, Apr. 30, 2004.

4. *See* Amy Lynn Jerdee, Note, *Breaking Through the Silence: Minnesota's Pregnancy Presumption and the Right to Refuse Medical Treatment*, 84 MINN. L. REV. 971, 977 (2000).

INTRODUCTION

The last few years have witnessed an astonishing array of intrusive and punitive government actions against pregnant women. These government interventions, ranging from criminal prosecutions and fetal "guardianship" proceedings to statutes safeguarding "the unborn" and new "regulatory interpretations" of existing law, are touted as necessary to protect fetuses from harm, particularly harm from their own mothers, and are framed as a response to a new public health crisis.[5] While these government actions vary in the extent to which they threaten women's physical liberty and decision-making autonomy,[6] they share a common view of pregnant women as vessels for the developing fetus, with both the potential, and the obligation, to protect that fetus at all costs.

---

5. *See* Ziba Kashef, *The Fetal Position: Federal and State Dollars Are Subsidizing a Boom in Antiabortion 'Crisis Pregnancy Centers,'* MOTHER JONES, Jan./Feb. 2003, *available at* http://www.motherjones.com/news/outfront/2003/01/ma_218_01.html. Other government activities support the position that embryos and fetuses are full human beings, as the Bush Administration has funded so-called "Snowflake Adoptions" (the directed donation of embryos to infertile couples) and fetal imaging technology for "pregnancy crisis centers," whose raison d'être is to discourage women from choosing abortions. *See* Elissa K. Zirinsky, *Adoption's New Frontier*, CBS News, July 28, 2005, http://www.cbsnews.com/stories/2005/07/28/national/printable712541.shtml; Anna Mulrine, *A Home for Frozen Embryos,* USNEWS.COM, Sept. 27, 2004, http://www.usnews.com/usnews/health/articles/040927/27babies.b1.htm (discussing the $1 million federal grant to two "embryo adoption" organizations to promote public awareness of these programs). The Department of Health and Human Services has funded so-called "crisis pregnancy centers" since 1996. *See* Kashef *supra*; *see also* The Abortion Access Project, Impeding the Right to Choose: Crisis Pregnancy Centers, and sources cited therein (on file with Author); Informed Choice Act, S.755.1S, 109th Cong. § 2(a) (providing for additional funding for ultrasound equipment to be used to provide pregnant women with a visual image of the fetus).

6. One might distinguish, for example, between prosecuting a woman for homicide because she used drugs while pregnant and a law that requires that pregnant women be told of the possibility of fetal pain before having an abortion. My point is not that all government "fetal protection" initiatives are equivalent, but that they each diminish the ability of a competent adult to make choices about her life and her body, and does so based upon the actor's status as a *pregnant woman.*

Current "fetal protection"[7] efforts pack a triple whammy: they undermine women's health, limit women's ability to fully participate in the economic life of the nation, and disproportionately affect the indigent and racial minorities. First, the new "fetal protection" threatens to limit women's ability to participate in the workforce and control their reproductive capability by raising the specter of civil or criminal liability if they engage in potentially risky activities before or during pregnancy. Second, many "fetal protection" initiatives seek to redefine the fetus as a person, with rights fully equal to those of a born human being, in a thinly disguised effort to limit abortion access.[8] Finally, efforts to constrain women's actions for the benefit of their fetuses frequently reflect racial, gender, and class stereotypes about how women in general, or certain groups of women, do or should behave.[9] It does not appear coincidental that poor women and women of color are the main targets of "fetal protection" efforts.[10]

---

7. The term "fetal protection" was apparently first used by legal commentators in the early 1980's, referring to employers' policies that excluded fertile women from the workplace, or at least better-paying jobs within the workplace. The ostensible purpose of these "fetal protection" policies was to ensure that children born to their female employees would not be injured by their mothers' on-the-job exposure to toxic chemicals, but the goal of protecting employers against tort liability was also important. *See, e.g.*, Wendy Williams, *Firing the Woman to Protect the Fetus: The Reconciliation of Fetal Protection with Employment Opportunity Goals Under Title VII*, 69 GEO. L.J. 641 (1981). In the mid-1980s the use of the term was broadened to include state laws prohibiting the experimentation on, and transfer of, embryos and fetuses, the fore-runner of today's controversy over stem cell research. *See, e.g.*, Note, *Reproductive Technology and the Procreation Rights of the Unmarried*, 98 HARV. L. REV. 669 (1985). In the late 1980s, courts and commentators began to use the term "fetal protection" to encompass tort actions and criminal prosecutions of women based on their conduct during pregnancy, as well as broader questions about how to consider the interests of women and their fetuses in the abortion context. *See, e.g.*, George Annas, *The Impact of Medical Technology on the Pregnant Woman's Right to Privacy*, 13 AM. J.L. & MED. 213, 229 (1987); John A. Robertson, *Gestational Burdens and Fetal Status: Justifying* Roe v. Wade, 13 AM. J.L. & MED. 189, 202 (1987); Dawn Johnsen, *From Driving to Drugs: Government Regulation of Pregnant Women's Lives After* Webster, 138 U. PA. L. REV. 179, 187-89 (1989). What I call the "new 'fetal protection'" is the increased range of government actions, beginning in the late 1990s and continuing through the present, taken against, or about, pregnant women, encompassing health care access and decisionmaking, civil commitment, and criminal and tort actions.

8. Julia L. Ernst, Laura Katzive, & Erica Smock, *The Global Pattern of U.S. Initiative Curtailing Women's Reproductive Rights: A Perspective on the Increasingly Anti-Choice Mosaic*, 6 U. PA. J. CONST. L. 752, 781 (2004); *see also infra* text accompanying notes 41-47, (discussing in detail the debate surrounding Unborn Victims of Violence Act, Pub. L. 108-212, 118 Stat. 568 (2004)).

9. Indeed, some of the most aggressive criminal prosecutions of pregnant women brought in the name of fetal protection have been brought in the former slave states of Florida, Missouri, South Carolina, and Texas; although these states are not the only locus of prosecution. *See infra* note 66; *see, e.g.*, Dorothy E. Roberts, *Punishing Drug Addicts Who Have Babies: Women of Color, Equality, and the Right of Privacy*, *in* CRITICAL RACE FEMINISM, A READER 127, 128-31 (Adrien Katherine Wing ed., New York University Press 2d ed. 1997) [hereinafter Roberts, *Punishing Drug Addicts Who Have Babies*] (suggesting that the "devaluation of [black women] as mothers . . . has its roots in the unique experience of slavery"); Dorothy E. Roberts, *Unshackling Black Motherhood*, 95 MICH. L. REV. 938, 939 (1997) [hereinafter Roberts, *Unshackling Black Motherhood*].

10. *See* Ira J. Chasnoff, *The Prevalence of Illicit-Drug or Alcohol Use During Pregnancy and Discrepancies in Mandatory Reporting in Pinellas County, Florida*, 322 NEW ENG. J. MED. 1202, 1206 (1990) (observing that black women were ten times as likely as white women to be reported by their physicians for using drugs, despite equal rates of drug use); LAURA E. GÓMEZ, MISCONCEIVING MOTHERS: LEGISLATORS, PROSECUTORS AND THE POLITICS OF PRENATAL DRUG EXPOSURE 118 (1997). Of course, one could observe that the poor and people of color are disproportionately repre-

Strikingly, the new "fetal protection" crusades have failed utterly to deliver more health care to poor children or women or to improve the health status of at-risk children. Rather, they are potent symbolic gestures, offering a quick fix to complex social, medical, and economic problems. By blaming individual women for conduct which is often not freely chosen,[11] government avoids taking responsibility for its continuing failure to meaningfully address the reality that many poor and low-income Americans lack access to health care or acknowledge the special problems faced by women who are victims of domestic violence, suffering from mental illness, and/or addicted to drugs and alcohol.[12] Effective health policy requires the provision of adequate heath care services for all, including reproductive health care across the life span and targeted services addressing our most vulnerable women and children.

This article will expand upon the feminist critique by focusing on children's health as well as the health and liberty interests of their mothers. In the first part of this article, I examine the legal and cultural underpinnings of "fetal protection" and explore its current manifestations. In the second part, I place "fetal protection" in a broader context, documenting the ways in which American law currently promotes fetal life, while simultaneously neglecting the lives and health of born children. The third part of the article offers concrete recommendations about how government, both state and federal, can actually achieve the goal of bringing healthy children into the world and enabling them to live healthy lives, paying particular attention to the problems of children who are born into domestic violence and/or poverty and are therefore at high risk for poor educational and health care outcomes.[13] If we are to truly

---

sented in the criminal justice system, in both the courts and prisons. *See, e.g.*, William H. Edmonson, Note, *A "New" No-Contact Rule: Proposing an Addition to the No-Contact Rule to Address Questioning of Suspects After Unreasonable Charging Delays*, 80 N.Y.U. L. REV. 1773, 1785 (2005). However, this does not explain the extraordinary frequency and ferocity of criminal prosecutions against women of color, particularly when one considers that it is alcohol, not crack cocaine (stereotypically connected with African-Americans), that is the drug most clearly shown to cause long-term developmental harm. *See infra* notes 14-22 and discussion in accompanying text.

11. The theme of "choice" is frequently raised by proponents of "fetal protection," ignoring the reality that for many poor women, the systemic lack of health care, education, and employment denies them the ability to make optimal choices for themselves or their children. *See, e.g.*, Erin Nelson, *Reconceiving Pregnancy: Expressive Choice and Legal Reasoning*, 49 MCGILL L.J. 593, 623, 624 (2004).

12. There is a significant link between a woman's experiencing domestic violence (physical or sexual abuse) as a child or an adult and her subsequent development of mental illness and/or substance abuse problems. WOMEN'S LAW PROJECT, RESPONDING TO THE NEEDS OF PREGNANT AND PARENTING WOMEN WITH SUBSTANCE USE DISORDERS IN PHILADELPHIA 4, 6 (Sept. 2002), *available at* http://www.womenslawproject.org/reports/Pregnant_parenting_PVS.pdf; Lynn M. Paltrow, *Pregnancy, Domestic Violence, and the Law: The Interface of Medicine, Public Health, and the Law: Governmental Responses to Pregnant Women Who Use Alcohol or Other Drugs*, 8 DEPAUL J. HEALTH CARE L. 461, 477 (2005).

13. Poverty is a major contributor to poor birth outcomes and later childhood health problems. *See* Charles Oberg, *The Impact of Childhood Poverty on Health and Development*, HEALTHY GENERATIONS, May 2003, at 2-3, *available at* http://www.epi.umn.edu/mch/resources/hg/hg_childpoverty.pdf; Jane D. McLeod & Michael J. Shanahan, *Trajectories of Poverty and Children's Mental Health*, 37 J. HEALTH & SOCIAL BEHAVIOR 207, 207 (1996).

become a society in which "no child [is] left behind," we must implement a comprehensive public health strategy to promote women's and children's health across the lifespan, not just during the few months in which women are pregnant.

What's "New" About the "New Fetal Protection?"

At the outset, one might be tempted to ask, "What's all the fuss about? Are these government actions so different from those taken before?" Women's use of alcohol, tobacco, and other drugs during pregnancy has long been controversial,[14] and there has been significant debate over whether criminal prosecutions and involuntary civil commitment are an appropriate or effective way to ensure that children are born healthy and drug-free.[15]

Virtually all observers agree that drug use, broadly defined, during pregnancy is harmful to the newborn child, although there is disagreement about the extent, and permanence, of the harm.[16] Research shows that 5-6% of women use illegal drugs during pregnancy, while 25% used alcohol, and maternal alcohol use is the leading cause of mental retardation.[17] Some researchers have concluded that maternal cocaine use may lead to subtle, long-lasting neurological deficits, including "the ability to habituate or self-regulate" and small but statistically significant deficits

14. Currently, there is increasing attention paid to methamphetamine, which, like cocaine in its day, is giving rise to media stories about the grave risks of in utero drug exposure for the long-term development of children. *See, e.g.*, Katie Zernike, *A Drug Scourge Creates its Own Form of Orphan*, N.Y. TIMES, July 11, 2005, at A1; *U.S. Warns of 'Global Meth Threat,'* BBC News, http://news.bbc.co.uk/1/hi/world/americas/4757179.stm. Others have criticized this media coverage as sensational and poorly informed. *See, e.g., Meth and Myth: Top Doctors, Scientists and Specialist Warn Mass Media on "Meth Baby" Stories*, July 29, 2005, http://stopthedrugwar.org/chronicle/397/methandmyth.shtml; *see also* RYAN S. KING, THE SENTENCING PROJECT, THE NEXT BIG THING? METHAMPHETAMINE IN THE UNITED STATES 16 (June 2006) (asserting that the media have failed utterly to accurately predict the science and epidemiological data surrounding methamphetamine addiction).

15. *See, e.g.*, Sarah Childress, *Justice: A New Controversy in the Fetal-Rights Wars*, NEWSWEEK, March 29, 2004, at 7; Lynn Paltrow, *Pregnant Drug Abusers, Fetal Persons, and the Threat to* Roe v. Wade, 62 ALB. L. REV. 999, 1008, 1009 (1999); Brian Maffly, *'Fetal Abuse' Charges Give Rise to Debate; Mothers-to-Be Need Help, Not Fear, Critics Say*, SALT LAKE TRIBUNE, Dec. 1, 1997, at D1; Wendy Chavkin, Vicki Breitbart, & Paul H. Wise, *Finding Common Ground: The Necessity of an Integrated Agenda for Women's and Children's Health*, 22 J.L. MED. & ETHICS 262, 263 (1994) (arguing that choosing to reduce infant mortality, HIV transmission, and drug exposure only by reducing harm to the fetus (and thus intervening through the body of the pregnant woman) ignores data showing that provision of quality health care across a woman's life is the best guarantee of ensuring healthy babies).

16. *See, e.g.*, DAN STEINBERG & SHELLY GEHSHAN, NATIONAL CONFERENCE OF STATE LEGISLATURES, STATES RESPONSES TO MATERNAL DRUG AND ALCOHOL USE: AN UPDATE (Jan. 2000), www.ncsl.org/programs/health/forum/maternalabuse.htm; JANET R. HANKIN, NATIONAL INSTITUTE ON ALCOHOL ABUSE AND ALCOHOLISM, FETAL ALCOHOL SYNDROME PREVENTION RESEARCH (Aug. 2002), http://pubs.niaaa.nih.gov/publications/arh26-1/58-65.htm.

17. *Addiction Medicine: Psychopathology of Pregnant Women with Alcohol and Drug Dependencies Examined*, WOMEN'S HEALTH WEEKLY, August 23, 2001, at 8 [hereinafter *Addiction Medicine*].

in IQ and language ability,[18] but others have found that most infants exposed in utero to cocaine "catch up to their peers in physical size and health status by age 2."[19] In contrast, maternal alcohol use during pregnancy is known to cause serious harm to children with significant in utero exposure, and infants born to mothers who drank moderately while pregnant may still experience deficits in IQ, learning, and attention.[20] Using tobacco during pregnancy poses risks similar in type to those of cocaine.[21] Most recent research emphasizes the multiple factors leading to poor birth outcomes, including maternal poverty, homelessness, a history of domestic violence, and lack of prenatal care, undermining the argument that drug use, whether legal or illegal, is the primary cause of children being born with deficits.[22]

Recent "fetal protection" efforts have been most aggressive in the criminal arena. In an unprecedented use of criminal law's heaviest artillery, prosecutors in two instances filed murder charges against women who delivered stillborn infants, based, respectively, on the woman's drug use while pregnant[23] or her refusal to have a Caesarean section.[24] The result was a murder conviction in the first instance and a conviction for felony child endangerment in the second.[25] Women have also been charged with other types of homicide[26] and with child abuse or reckless

---

18. Steven J. Ondersma et al., *Prenatal Drug Exposure and Social Policy: The Search for an Appropriate Response*, 5 CHILD MALTREATMENT 93, 95 (2000), *available at* http://cmx.sagepub.com/cgi/reprint/5/2/93.

19. STEINBERG & GEHSHAN, *supra* note 16.

20. Ondersma et al., *supra* note 18, at 96.

21. N. Kistin, A. Handler, F. Davis, & C. Ferre, *Cocaine and Cigarettes: a Comparison of Risks*, 10 PEDIATRIC PERINATAL EPIDEMIOLOGY 269 (1996) (noting that while children exposed to cocaine in utero were more likely to have adverse birth outcomes than children whose pregnant mothers consumed no drugs, children whose mothers used tobacco products while pregnant were at risk for the same adverse outcomes as children whose mothers used cocaine, although the magnitude of the risk was lower. "[G]iven the greater number of cigarette smokers than cocaine users in the population . . . [there are likely to be more children harmed by their mothers' smoking than by their mothers' cocaine use during pregnancy]").

22. *See, e.g.*, Ondersma et al., *supra* note 18, at 95; Deborah A. Frank et al., *Growth, Development, and Behavior in Early Childhood Following Prenatal Cocaine Exposure*, 285 JAMA 1613, 1615 (2001). Because some women who use illegal drugs also abuse alcohol, researchers recognize the need for comprehensive and intensive drug treatment programs that take into account the complex needs of this population, which has high "[r]ates of homelessness, poverty, unemployment, and prostitution . . . [and] histories of emotional, physical, and sexual abuse." *See Addiction Medicine*, *supra* note 17.

23. *McKnight*, 576 S.E.2d at 171.

24. *Mom in Caesarean Case Gets Probation*, CHI. TRIB., April 30, 2004, at 18; *see also* Linda Thomson, *Mother is Charged in Stillbirth of a Twin*, DESERET MORNING NEWS, March 12, 2004; Linda Thomson, *Rowland Airs Her Case on Cable TV*, DESERET MORNING NEWS, May 13, 2004.

25. *See McKnight*, 576 S.E.2d at 171; Jacob Santini, *Stillborn Twin Case Fades, Issues Stay*, THE SALT LAKE TRIBUNE, April 16, 2004, at B4. Criminal prosecutions against women who used alcohol or other drugs while pregnant began in the late 1980's. *See infra* note 66. However, in only one state, South Carolina, were these prosecutions and convictions ultimately sustained by the courts. *See* Whitner v. State, 492 S.E.2d 777, 778-79, 786 (S.C. 1997) (upholding conviction under child endangerment statute for drug use during pregnancy because viable fetus is a "child" under the statute); *McKnight*, 576 S.E.2d at 171.

26. State v. Deborah J.Z., 596 N.W.2d 490, 491 (Wis. Ct. App. 1999); State v. Aiwohi, 123 P.3d 1210, 1210-11 (Haw. 2005).

endangerment based on their alcohol and drug use while pregnant,[27] relying on reports from physicians or newborn toxicology testing. Yet the unanimous judgment of medical and public heath groups is that such prosecutions will only drive a wedge between pregnant women and their physicians, and render it less, not more, likely that the women will seek appropriate pre- and post-natal care, including substance abuse treatment.[28]

In addition to criminal prosecutions, in the last several years, a breathtaking array of civil suits and statutory and regulatory initiatives has sought to treat fetuses as entirely separate from the pregnant women whose bodies sustain them. In 2002, the federal Department of Health and Human Services (HHS) issued regulations "clarify[ing] and expand[ing]" the statutory definition of "child" in the State Children's Health Insurance Program (SCHIP), a program which provides health care to low-income children.[29] The regulations redefined "child," from "an individual under 19 years of age"[30] to "an individual under the age of 19 including the period from conception to birth."[31] Critics asserted that this recasting of fetuses as "children" was both unnecessary and ineffective if, as HHS claimed, its goal was to provide pregnant women with

27. *See infra* discussion in text accompanying notes 112-15 (discussing recent prosecutions initiated in Maryland, Missouri, Texas, and Wyoming).

28. *See* Ferguson v. City of Charleston, 532 U.S. 67, 85-86 (2001) (holding that public hospital's policy of testing pregnant women for drug use, developed in conjunction with local prosecutors and police, and turning drug results over to authorities for criminal prosecution, did not come within the "special needs" exception to the Fourth Amendment). The Court observed that "an intrusion on . . . [a patient's expectation of privacy in regard to diagnostic medical tests] may have adverse consequences because it may deter patients from receiving needed medical care." *Id.* at 78 n.14. In a separate article, I will explore in greater depth the anti-deterrent impact of criminal prosecutions on women seeking prenatal care and substance abuse treatment.

29. SCHIP was established in 1997 under Title XXI of the Social Security Act, 42 U.S.C. §§ 1397aa-1397jj (2000), and gives states the opportunity to provide additional health insurance coverage to children whose parents are too "wealthy" to qualify for Medicaid. Medicaid, which was enacted in 1965 and is authorized by Title XIX of the Social Security Act, 42 U.S.C. §§ 1396v et seq., provides health care insurance for the very poorest of American children. BARRY R. FURROW ET AL., THE LAW OF HEALTH CARE ORGANIZATION AND FINANCE 420-21 (4th ed. 2001). Both Medicaid and SCHIP are federal/state partnerships, with the federal and state governments sharing in both the financing and administration of the two programs. However, there are important differences. Medicaid is an entitlement program, in which all eligible persons must receive the same benefits. SCHIP gives states much more flexibility in terms of the services that a particular state may choose to provide. FURROW ET AL., *supra*, at 418-21, 438-39; *see also* Sara Rosenbaum, Anne Markus, & Colleen Sonosky, *Public Health Insurance Design for Children: The Evolution from Medicaid to SCHIP*, 1 J. HEALTH & BIOMED. L. 1, 3-12 (2004) (arguing that Medicaid, because it provides a more comprehensive set of benefits, including Early and Periodic Screening, Diagnostic, and Treatment (EPSDT) is a superior program). Children who receive Medicaid "are more likely than uninsured children and as likely as privately insured children to receive well-child visits and to visit the doctor in a given year. KAISER COMMISSION ON MEDICAID AND THE UNINSURED, MEDICAID FACTS, EARLY AND PERIODIC SCREENING, DIAGNOSTIC, AND TREATMENT SERVICES (Oct. 2005) [hereinafter MEDICAID FACTS], *available at* http://www.kff.org/medicaid/upload/Early-and-Periodic-Screening-Diagnostic-and-Treatment-Services-Fact-Sheet.pdf.

30. 42 U.S.C. § 1397jj(c)(1) (2000).

31. 42 C.F.R. § 457.10 (2006).

prenatal care.[32] Instead, it appears that the real purpose of the regulation was to establish the legal principle that fetuses are children, with all the rights that accrue to that status.[33]

Government lawyers have also sought the involuntary civil commitment of pregnant women, in order to impose "treatment" on the women and their fetuses,[34] as well as court orders mandating Caesarean sections.[35] While the avowed goal of these actions is to ensure the birth of healthy children, here too the consensus among medical professionals is that such interventions are unjustified.[36] More than thirty states' laws permit civil commitment based on the use of alcohol and other drugs,[37]

---

32. ROBERT WOOD JOHNSON FOUNDATION, THE MEDICAID EXPANSIONS FOR PREGNANT WOMEN AND CHILDREN (1995).

33. For example, in the Omnibus Budget Reconciliation Act (OBRA) of 1986 Congress expanded Medicaid to permit states to enroll all pregnant women with incomes up to 100% of the Federal Poverty Level and adopted procedural changes that made enrollment easier, thus significantly increasing the number of pregnant women eligible to receive pre- and post-natal care as a means of ensuring better birth outcomes. *See, e.g.*, Omnibus Budget Reconciliation Act of 1986, Pub. L. No. 99-509, § 9401(b)(2), 100 Stat. 1874 (1986); GENERAL ACCOUNTING OFFICE, PRENATAL CARE: EARLY SUCCESS IN ENROLLING WOMEN MADE ELIGIBLE BY MEDICAID EXPANSIONS 7 (February 1991), *available at* http://archive.gao.gov/d21t9/143346.pdf. While these changes in the Medicaid program were not totally effective in achieving the birth of healthier children, ROBERT WOOD JOHNSON FOUNDATION, *supra* note 32, at 1-2, no one had ever suggested that the result would be better if the fetuses were enrolled rather than the women in whose bodies they were developing.

34. These include the case of *State ex rel. Angela M.W. v Kruzicki*, 561 N.W.2d 729, 732 (Wis. 1997) and Rebecca Corneau, a pregnant woman who belonged to a religious sect that did not believe in Western medicine, who was confined in a "secure hospital facility for pregnant prison inmates" by a Massachusetts juvenile court judge until she agreed to medical examination and treatment. *See* Marilyn L. Miller, Note, *Fetal Neglect and State Intervention: Preventing Another Attleboro Cult Baby Death*, 8 CARDOZO WOMEN'S L. J. 71, 71 (2001). These cases will be discussed in more detail in Part I. C., *infra.*

35. *News . . . Husband to Challenge Court Order in Lawsuit over Wife's Refusal of Caesarean Section*, PENN. LAW WEEKLY, Jan. 26, 2004, at 9; Associated Press, *New Questions about Childbirth Rights*, May 19, 2004, http://keyetv.com/health/health_story_140110423.html (discussing the case of Amber Marlowe, who was the subject of an ex parte order to have a Caesarean section because her fetus weighed 11 pounds, despite her having delivered 6 very large children previously).

36. *See, e.g.*, AMERICAN MEDICAL ASSOCIATION, POLICY H-420.969: LEGAL INTERVENTIONS DURING PREGNANCY, *available at* http://www.ama-assn.org/ama/noindex/category/11760.html, (click "accept," search "420.969") (propounding a general rule that "[j]udicial intervention is inappropriate when a woman has made an informed refusal of a medical treatment designed to benefit her fetus" and specifically recognizing the need for rehabilitative treatment for pregnant substance abusers); AMERICAN COLLEGE OF OBSTETRICS AND GYNECOLOGY, *Patient Choice in the Maternal-Fetal Relationship*, *in* ETHICS IN OBSTETRICS AND GYNECOLOGY (2d ed. 2004), *available at* http://www.acog.org/from_home/publications/ethics/ethics034.pdf (stating that "court-ordered intervention against the wishes of a pregnant woman is rarely if ever acceptable"); American Academy of Pediatrics, Committee on Bioethics, *Fetal Therapy – Ethical Considerations*, 103 PEDIATRICS 1061, 1062 (May 1999) (after discussing the range of medical interventions to promote fetal health and the legal-ethical issues involved, concluding that "Under no circumstances should a physician physically intervene [to insist on medical treatment] without the explicit consent of the pregnant woman without judicial review . . . .").

37. ALA. CODE. § 22-52-1.2 (LexisNexis 2006) ; ALASKA STAT. § 47.37.190 (2006); ARK. CODE ANN. § 20-64-815 (2006); CAL. WEL & INST CODE § 3050 (Deering 2006); COLO. REV. STAT. § 25-1-1107 (2006); CONN. GEN. STAT. § 17a-685 (2006); DEL. CODE ANN. tit. 16, § 2212 (2006); D.C. CODE ANN. §7-1303.04 (LexisNexis 2006); FLA. STAT. ANN. § 397.675 (LexisNexis 2006); GA. CODE ANN. § 37-7-41 (2006); HAW. REV. STAT. ANN. § 334.60.2 (LexisNexis 2006); IDAHO CODE ANN. § 66-329 (2006); IND. CODE ANN. § 12-23-11-1 (LexisNexis 2006); IOWA CODE §

and several states have recently enacted laws specifically authorizing the civil commitment of pregnant women based on substance abuse.[38] Further, a majority of states which authorize the use of advance medical directives to govern the medical care of mentally incompetent individuals suspend the operation of these directives if the patient is pregnant.[39]

In June 2003, the wife of a Florida prosecutor sought to be appointed "guardian" of the fetus of a mentally disabled patient who lived in a group home in order to prevent the woman from having an abortion.[40] Although the Florida courts ultimately rejected the suit, the case became a cause célèbre in Florida.

In March 2004, Congress enacted the Unborn Victims of Violence Act (the UVVA or Act),[41] which made it a crime to injure or cause the death of a fetus while committing another federal offense.[42] Both supporters and opponents of the Act acknowledged the significant problem of violence against pregnant women;[43] however, opponents objected to the Act's solution. Rather than focusing on the injury suffered by the pregnant woman herself and providing that a person who harms a pregnant woman who in the process injures or kills the fetus should receive

---

125.75 (2006); KAN. STAT. ANN. § 59-29b54 (2006); LA. REV. STAT. ANN. § 28:54 (2006); MASS. ANN. LAWS ch. 123, § 35 (LexisNexis 2006); MISS. CODE. ANN. § 41-30-27 (2004); NEB. REV. STAT. ANN. § 71-919 (LexisNexis 2006), see also NEB. REV. STAT. ANN. § 71-908 (LexisNexis 2006); N.H. REV. STAT. ANN. § 135-C:27 (2006); N.M. STAT. ANN. § 43-2-8 (LexisNexis 2006); N.D. CENT. CODE § 12.1-04.1-22 (2006); R.I. GEN. LAWS § 21-28.2-3 (2006); S.C. CODE ANN. § 44-52-50 (2005); S.D. CODIFIED LAWS § 29A-5-311 (2006); TEXAS HEALTH & SAFETY CODE ANN. § 574.034 (Vernon 2006); VA. CODE ANN § 37.2-809 (2006); WASH. REV. CODE ANN. § 70.96A.140 (LexisNexis 2006); W. VA. CODE ANN. § 27-5-2 (LexisNexis 2006); WIS. STAT. § 51.15 (2006); WYO. STAT. ANN. § 25-10-110 (2006). Minnesota, Oklahoma, and South Dakota have involuntary commitment laws specifically for pregnant women who use drugs. MINN. STAT. § 626.5561 (2006); OKLA. STAT. tit. 43A, § 5-410 (2005); see also OKL. STAT. tit. 63 § 1-546.5 (2005); S.D. CODIFIED LAWS § 34-20A-70 (2006).

38. *See, e.g.*, WIS. STAT. §§ 48.205 (2006) (permitting the civil commitment of pregnant girls and women, dubbed "The Cocaine Mom law"); *see also* Tom Kertscher, *'Cocaine Mom' Law Invoked in Attempt to Detain Woman, Racine Case Thought to Be First Time Law is Used Without Other Crime*, MILWAUKEE JOURNAL SENTINEL, Nov. 5, 1999, at 1.

39. *See* discussion *infra* Part I.D.

40. This attempt was rejected by the Florida District Court of Appeal in *In re Guardianship of J.D.S.*, 864 So. 2d 534, 535 (Fla. Dist. Ct. App. 2004), which held that under the Florida guardianship statute, a guardian can be appointed only for a "person," and that fetuses were not "persons" under Florida law. *Id.* at 538.

41. Pub. L. 108-212, 118 Stat. 568 (2004).

42. 18 U.S.C.A. § 1841 (West 2006). The Act enumerated a lengthy list of federal offenses, including drive-by shootings in connection with drug offenses (18 U.S.C.A. § 36), violence at international airports (18 U.S.C.S. § 37), and assault on a federal officer or employee (18 U.S.C.A. § 111). *See* 18 U.S.C.A. § 1841(b)(1).

43. H.R. REP. NO. 108-420, Pt. 1, at 4 n.2 (2004) (Conf. Rep.) (citing Victoria Frey, *Examining Homicide's Contribution to Pregnancy-Associated Deaths*, 285 JAMA 1510 (2001) (summarizing the various studies)); Isabelle L. Horon & Diana Cheng, *Enhanced Surveillance for Pregnancy-Associated Mortality--Maryland, 1993-1998*, 285 JAMA 1455 (2001); Linn H. Parsons & Margaret A. Harper, *Violent Maternal Deaths in North Carolina*, 94 OBSTET. GYNECOL. 990, 991 (1999); Dannenberg et al., *Homicide and Other Injuries as Causes of Maternal Death in New York City, 1987 through 1991*, 172 AM. J. OBSTET. GYNECOL. 1557 (1995); Fildes et al., *Trauma: The Leading Cause of Maternal Mortality*, 32 J. TRAUMA 643-45 (1992).

an enhanced penalty for that harm,[44] the UVVA makes such an attack or injury a separate crime.[45] To do so, the UVVA defines "unborn child" broadly, as "a member of the species homo sapiens, *at any stage of development* . . . ."[46] Like the SCHIP regulation, this language raises concern that the statute's real goal is to limit women's ability to obtain an abortion.[47]

Most recently, laws have been proposed which emphasize fetal "personhood" in new ways. These include laws requiring women seeking abortion to be told about fetal pain,[48] to be informed of the need to prepare a fetal death certificate, or to be given the opportunity to view a sonogram or listen to the heartbeat of their fetus prior to deciding to have an abortion.[49] Supporters of these statutes justify them as providing "informed consent," but the statutes are unusual in mandating the substantive details of what patients contemplating a medical procedure must be told. In contrast, most American informed consent[50] law focuses on the *process* of ensuring full communication between patients and their health care providers rather than on the *content* of the physician-patient dialogue,[51] relying on the health care professional to determine what information to convey to a particular patient based on her individual needs.

---

44. Senator Dianne Feinstein proposed an amendment to the Senate bill to accomplish this, which was defeated by a vote of 50-49, largely along party lines. A similar amendment offered by Representative Zoe Lofgren was also defeated in the House of Representatives, by a 229-186 vote. Edward Epstein, *Bill to Make Harming Fetus a Crime is Passed by Senate; Assailant of a Pregnant Woman Could be Charged with 2 Separate Federal Offenses*, S. F. CHRON., March 26, 2004, at A1; *see also* H.R. REP. NO. 108-420, Pt. 1, at 86.

45. 18 U.S.C.A. § 1841 (West 2006).

46. 18 U.S.C.A. § 1841(d). Under the law, "the term unborn child means a child in utero, and the term "child in utero" or "child, who is in utero" means a member of the species homo sapiens, at any stage of development, who is carried in the womb." *See also* 10 U.S.C.A. § 919a(d) (West 2006).

47. Senator Feinstein argued that the UVVA was a deliberate effort to undermine abortion rights, by "'set[ting] . . . the stage for a jurist to rule that a human being an any stage of development deserves . . . rights under the law' . . . ." Epstein, *supra* note 46.

48. *See* S.51, 109th Cong. § 2902 (2005) (proposed by Senator Sam Brownback of Kansas, a fierce abortion opponent); H.B. 238, 59th Leg., Reg. Sess. (Mont. 2005). Both bills are discussed *infra* in text accompanying notes 244-252.

49. *See* discussion *infra* in Section I.D. The way for these laws has been paved by federal funding of fetal imaging machinery, through federal and state grants that are given to organizations that promote "abstinence only" sex education. The so-called "pregnancy crisis centers" have been a major beneficiary of such grants. Kashef, *supra* note 5; The Abortion Access Project, *supra* note 5.

50. Informed consent doctrine has roots in both the common law tort of battery and in negligence. It protects a patient's interest in choosing when to be touched (a battery is an unconsented touching and includes medical treatment which the patient did not agree to). *See, e.g.*, Schloendorff v. The Society of New York Hospital, 105 N.E. 92, 93 (N.Y. 1914). It also ensures that a patient receives medical treatment from a physician who has explained to the patient those risks and benefits of treatment that a reasonable patient would wish to know. *See, e.g.*, Canterbury v. Spence, 464 F.2d 772, 787 (D.C. Cir. 1972); N.Y. PUB. HEALTH LAW § 2805-d (McKinney 2001).

51. For example, some abortion statutes require that the pregnant woman be told certain details about the fetus, such as its gestational age and its potential to survive outside the womb, and be informed of the availability of medical assistance for prenatal care, childbirth, and neonatal care, as well as options for child support and adoption. *See, e.g.*, LA. REV. STAT. ANN § 40:1299.35.6 (2006); TEX. HEALTH & SAFETY CODE ANN §171.012 (Vernon 2003); Planned Parenthood v. Casey, 505 U.S. 833, 881 (1992). In addition, there are other areas of health care in which state laws

In sum, the mounting numbers of civil and criminal actions against pregnant women, along with statutes and regulations equating fetuses with born children, mean that the new "fetal protection" crusade can no longer be ignored. These government initiatives are particularly disturbing because their focus on the harm that could be caused by a woman's behavior during pregnancy ignores the widespread failings of the American health care system, which does not promote women's and children's health.

The United States falls far short of other developed countries in objective indicators of health status,[52] and indeed, American infant mortality rates have risen in recent years.[53] Two-thirds of American infants who die in their first year of life suffer from low birthweight, attributable in part to their mothers' lack of prenatal care and long-standing health problems, as well as to multiple births.[54] One-eighth of American children are born pre-term, at an estimated cost of $26 billion per year.[55] There are significant racial disparities in birth outcomes and other measures of children's health, which reflect major problems of health care access, including the lack of a primary care physician and the lack of health insurance.[56] More than ten million American children have no

---

mandate that patients (usually women) be told of alternative medical or surgical options. *See, e.g.*, CAL. HEALTH & SAFETY CODE § 1690 (West 2006) (sterilization); OR. REV. STAT. ANN. § 436.225 (West 2003) (sterilization); S.G. Nayfield et al., *Statutory Requirements for Disclosure of Breast Cancer Treatment Alternatives*, 86 J. NAT'L CANCER INST. 1202 (1994); ARK. CODE ANN. § 6-18-702 (West 2006) (childhood vaccination). Medical procedures that are less politically charged rarely have such "informed consent" requirements.

52. *See, e.g.*, ASSOCIATION OF MATERNAL AND CHILD HEALTH PROGRAMS, PLANNING FOR HEALTHY FAMILIES (June 2004), *available at* http://amchp.org/aboutamchp/publications/familyplanning2004.pdf (noting that the United States ranks 29th in the world in infant mortality and arguing that more attention should be devoted to encouraging family planning as a way of ensuring good birth outcomes, as data show that when pregnancies are intended children are less likely to be born premature and with health problems); Margaret A. Harper et al, *Pregnancy-Related Death and Health Care Services*, 102 OBSTETRICS & GYNECOLOGY 273, 273, 275, 276 (Aug. 2003), *available at* http://www.greenjournal.org/cgi/reprint/102/2/273 (noting that "[m]aternal mortality statistics for the United States have shown little improvement for 2 decades, and 20 countries have lower rates," and concluding in a study of North Carolina maternal pregnancy deaths that lacking access to prenatal care made maternal death slightly more likely and having a Caesarean section made maternal death nearly four times as likely). *See also* detailed discussion in Part II.B, *infra*.

53. T.J. Matthews, *Racial/Ethnic Disparities in Infant Mortality- United States, 1995-2002*, MORBIDITY AND MORTALITY WEEKLY REPORT, June 10, 2005.

54. *Id. See also* ANNIE E. CASEY FOUNDATION, 2004 KIDS COUNT DATA BOOK, 34 (2005), *available at* www.kidscount.org [hereinafter CASEY FOUNDATION, KIDS COUNT]. Low birthweight is also linked significantly to being born as a twin or other multiple births. *Id.* at 34; Tarun Jain et al., *Trends in Embryo-Transfer Practice and in Outcomes of the Use of Assisted Reproductive Technology in the United States*, 350 N. E. J. MED. 1639, 1640 (2004).

55. Press Release, Institute of Medicine, Preterm Births Cost U.S. $26 Billion a Year; Multidisciplinary Research Effort Needed to Prevent Early Births (July 13, 2006), *available at* http://www8.nationalacademies.org/onpinews/newsitem.aspx?RecordID=11622 [hereinafter IOM Report]. The Report defines "preterm" as any birth that occurs at less than 37 weeks of pregnancy (a full-term pregnancy is 38-42 weeks post-conception) and notes that the rate of pre-term births has risen 30% since 1981. *Id.*

56. IOM Report, *supra* note 55; Matthews, *supra* note 53 (noting significant racial disparities in infant mortality rates within and across states); Kenneth E. Thorpe, Jennifer Flome & Peter Joski, The Distribution of Health Insurance Coverage Among Pregnant Women, 1999 (Emory University

health insurance at all,[57] even though at least 70% live in families where at least one parent works full time.[58]

If the goal of government policymakers and prosecutors were actually to ensure that more children are born healthy and have the opportunity to stay that way, the United States would adopt radically different policies. In addition to the lack of health care access, two notable omissions from the rhetoric of fetal protection are the harms posed to children by assisted reproductive technology (ART), used largely by the middle and upper classes,[59] and the risk to all children posed by environmental

---

April 2001) (identifying disparities in health insurance coverage along racial, employment status, and income lines) (paper prepared for the March of Dimes, on file with the author).

57. KAISER COMMISSION ON MEDICAID AND THE UNINSURED, MEDICAID FACTS, ENROLLING UNINSURED LOW-INCOME CHILDREN IN MEDICAID AND SCHIP (March 2005) (summarizing 2002 data), *available at* http://www.kff.org/medicaid/2177-04.cfm [hereinafter KAISER, ENROLLING CHILDREN]. In 2005, more than 12% of children under age 18 lacked health insurance for at least part of the previous year. ROBIN A. COHEN & MICHAEL E. MARTINEZ, CENTERS FOR DISEASE CONTROL, HEALTH INSURANCE COVERAGE: ESTIMATES FROM THE NATIONAL HEALTH INTERVIEW SURVEY, JANUARY–SEPTEMBER 2005, 3 (Mar. 20, 2006), *available at* http://www.cdc.gov/nchs/data/nhis/earlyrelease/insur200603.pdf. "[U]ninsured but Medicaid-eligible children are twice as likely as those enrolled in Medicaid to have an unmet medical need, to have not seen a doctor, and to have substantial family out-of-pocket spending on health care." KAISER, ENROLLING CHILDREN, *supra.*

58. Associated Press, *Most Uninsured Children Have Parents with Jobs*, Sept. 28, 2006, http://www.foxnews.com/printer_friendly_story/0,3566,216338,00.html.

59. What has been absent from the government initiatives described above are any efforts to regulate the new assisted reproductive technologies (ARTs) designed to address problems of infertility, which increase the risk of adverse birth outcomes, but which are used primarily by middle- and upper- income Americans. *See, e.g.*, Jain, *supra* note 54, at 1640 (noting the continuing high rate of multiple births in the United States and their adverse consequences, but observing that the United States, in contrast to many other countries, has not regulated ART practices, "in part because of the basic belief that such decisions should be left to couples and their physicians"); Liza Mundy, *A Special Kind of Poverty: The Poor Get Used to Going Without, but Going Without a Baby is Hard to Get Used to*, WASH. POST, Apr. 20, 2003, at W08 (describing costs of infertility treatments and how poor men and women seek subsidized or alternative access to fertility treatment). People who use ART are much more likely than the rest of the population to have twins or other multiple births, which in turn dramatically increases the chances of having a low birthweight infant (one who weighs less than 2500 grams, or about 5.5. pounds), from 6% to 57%. See CASEY FOUNDATION, KIDS COUNT, *supra* note 54, at 34. Low birth weight is a major contributor to infant mortality and developmental defects. *Id.* Yet even singleton births achieved through ART are at risk for harm. Jennifer L. Rosato, *The Children of ART (Assisted Reproductive Technology): Should the Law Protect Them From Harm?*, 2004 UTAH L. REV. 57, 60, 62-66, 69-70, 77-80 (summarizing the data showing that up to 10% of children born using ART suffer some adverse consequences and criticizing the regulatory hands-off position of states and the federal government); *see also* John A. Robertson, *Procreative Liberty and Harm to Offspring in Assisted Reproduction*, 30 AM. J. L. & MED. 7, 9 (2004) (noting that intracytoplasmic sperm injection (ICSI), which is used in nearly half of American in vitro fertilization (IVF) treatments, may cause a higher incidence of rare birth defects as well as low birth weight). *But see* Anja Pinborg et al., *Neurological Sequelae in Twins Born After Assisted Conception: Controlled National Cohort Study*, 329 BRIT. MED. J. 311 (July 15 2004) (finding no difference in adverse birth outcomes between infants conceived through ICSI and IVF, but noting that children born through ART methods have higher rates of stillbirths and neurological problems). *See also* PHILIP G. PETERS, JR., HOW SAFE IS SAFE ENOUGH? OBLIGATIONS TO THE CHILDREN OF REPRODUCTIVE TECHNOLOGY (2004). Similarly, the bioethics issues raised by particular uses of IVF are generally ignored. Arlene Judith Klotzko, *Medical Miracle or Medical Mischief? The Saga of the McCaughley Septuplets*, Hastings Center Report 5 (May 1998); Susan M. Wolf, Jeffrey P. Kahn, & John E. Wagner, *Using Preimplantation Genetic Diagnosis to Create a Stem Cell Donor: Issues, Guidelines & Limits*, 31 J.L. MED. & ETHICS 327, 331 (2003) (arguing that the combination of IVF technology with preimplantation genetic diagnosis in order to produce a child who is a poten-

hazards, including mercury in fish[60] and the forests[61] to pesticides[62] to lead from older buildings and manufacturing.[63] The United States has also failed to promote fetal and children's health and development by providing paid parenting leaves.[64] When compared to other developed nations where universal health care and subsidized parenting leave are the norm,[65] the approach of the United States is both seriously out of step and actively unhelpful in promoting childhood health.

---

tial organ donor for a sibling with a rare genetic disorder is so ethically questionable that it should only take place under human subject research protocols which have been thoroughly reviewed by institutional review boards). Only recently have suggestions been made that ART should be monitored and regulated. THE PRESIDENT'S COUNCIL ON BIOETHICS, REPRODUCTION AND RESPONSIBILITY: THE REGULATION OF NEW BIOTECHNOLOGIES, EXECUTIVE SUMMARY 2-9 (March 2004), *available at* www.bioethics.gov. Due to the volatile politics evoked by discussions of fetal and embryonic life, little regulation is likely to occur anytime soon. Rosato, *supra*, at 74, 75. In July 2006 the Institute of Medicine issued a report noting the 30% increase in pre-term labor over the last 25 years, and urging further study of the contribution of ART to this growth. IOM Report, *supra* note 55.

60. Many species of fish pose risks to adults, children, and fetuses, primarily through exposure to mercury and polychlorinated biphenyls (PCB). "Children born to women exposed to high levels of methylmercury [the organic form of mercury found naturally in the environment] during or before pregnancy may face numerous health problems, including brain damage, mental retardation, blindness, and seizures. Lower levels of methylmercury exposure in the womb have caused subtle but irreversible deficits in learning ability." Jennifer Fisher Wilson, *Balancing the Risks and Benefits of Fish Consumption*, 141 ANNALS INT. MED. 977, 978 (2004). PCBs are a probable carcinogen. In addition, "[i]n children, PCB exposure in utero and from breast milk consumption has been linked with neurodevelopmental delays, impaired cognition, immune problems, and alterations in male reproductive organs." *Id.* at 979.

61. Anthony DePalma, *Study of Songbirds Finds High Levels of Mercury*, N.Y. TIMES, Jul. 25, 2006, at B1.

62. A number of widely-used pesticides are suspected of being endocrine disrupters, which affect both male and female reproductive systems and increase the chances of infertility and other reproductive harms. SHARON L. DROZDOWSKI & STEPHEN G. WHITTAKER, WASHINGTON STATE DEPARTMENT OF LABOR & INDUSTRIES' SAFETY & HEALTH ASSESSMENT & RESEARCH FOR PREVENTION PROGRAM, WORKPLACE HAZARDS TO REPRODUCTION AND DEVELOPMENT: A RESOURCE FOR WORKERS, EMPLOYERS, HEALTH CARE PROVIDERS, AND HEALTH & SAFETY PERSONNEL 48, 49 (Aug. 1999).

63. Lead poses risks to male and female workers, as well as their children. In men, lead exposure leads to lowered sperm counts, abnormal sperm shapes, altered sperm transfer, and altered hormone levels. The results can be sterility and infertility. In women, lead can cause miscarriages, stillbirths, and infertility, as well as developmental disorders in children exposed in utero. NATIONAL INSTITUTE FOR OCCUPATIONAL SAFETY AND HEALTH, THE EFFECTS OF WORKPLACE HAZARDS ON FEMALE REPRODUCTIVE HEALTH 2-3 (Feb. 1999), *available at* http://www.cdc.gov/niosh/99-104.html. Lead that workers bring home on their skin, hair, clothes, tool box or car can cause severe lead poisoning for everyone who comes into contact with it, and can lead to neurobehavioral and growth effects in a fetus. NATIONAL INSTITUTE FOR OCCUPATIONAL SAFETY AND HEALTH, THE EFFECTS OF WORKPLACE HAZARDS ON MALE REPRODUCTIVE HEALTH (1997), http://www.cdc.gov/niosh/ malrepro.html.

64. The Family and Medical Leave Act of 1993, Pub. L. No. 103-3, 107 Stat. 6 (1993), discussed *infra* in text accompanying notes 329-44, requires employers of more than fifty employees to permit employees to take an unpaid leave for their own illness or a family member's birth, adoption, or illness. However, in contrast to almost all developed countries, the United States does not mandate paid leave. KURT H. DECKER, FAMILY AND MEDICAL LEAVE IN A NUTSHELL 9-14 (2000).

65. Sakiko Tanaka, *Parental Leave and Child Health Across OECD Countries*, 115 THE ECON. J. F7, F8, F9 (Royal Economic Society 2005).

I. THE NEW "FETAL PROTECTION"

A. Criminal Prosecutions

Criminal actions against pregnant women have risen sharply in the last decade. Although women have been prosecuted for drug and alcohol use during pregnancy in more than half the states since the 1970s, courts have quashed prosecutions or overturned convictions in all but one state, South Carolina.[66] There the courts have upheld women's convictions for homicide and child abuse based on their conduct during pregnancy.[67] Virtually all the women who faced these criminal charges were living at the very margins of society, suffering from poverty, substance abuse, and often, mental disability.[68] Frequently they were sexually abused as children, and often they are current victims of domestic violence.[69]

In 1996, Deborah J.Z. was charged with attempted first-degree intentional homicide[70] and first-degree reckless injury,[71] after she went into labor while at a tavern and said that she would drink herself and her fetus to death.[72] Her child was born with a high blood alcohol level and physical features showing fetal alcohol effects.[73] Although the Wisconsin Supreme Court condemned her conduct, it barred criminal prosecution because under Wisconsin's "born alive" rule, a fetus was not a hu-

66. The first reported effort at prosecution was in 1977, when a California prosecutor indicted Margaret Reyes on two counts of felony child endangering based on her heroin use while pregnant, which allegedly caused her twin sons to be born addicted to heroin. The California Court of Appeal issued a writ or prohibition, enjoining further prosecution of the case. Reyes v. Superior Court, 141 Cal. Rptr. 912 (Cal. Ct. App. 1977). A rash of prosecutions began in the late 1980s, beginning with the 1987 prosecution of Pamela Rae Stewart, who had intercourse with her husband and used amphetamines, both against medical advice. *See* George Annas, *The Impact of Medical Technology on the Pregnant Woman's Right to Privacy*, 13 AM J. L. & MED. 213, 229 (1987). Since then, at least 30 states have prosecuted women for manslaughter, child abuse or endangerment, or drug delivery to a minor. *See* CENTER FOR REPRODUCTIVE RIGHTS, PUNISHING PREGNANT WOMEN FOR THEIR BEHAVIOR DURING PREGNANCY: AN APPROACH THAT UNDERMINES WOMEN'S HEALTH AND CHILDREN'S INTERESTS 2 [hereinafter CTR. FOR REPRODUCTIVE RIGHTS, PUNISHING PREGNANT WOMEN], *available at* http://www.reproductiverights.org/pdf/pub_bp_punishingwomen.pdf (listing cases wherein women have been prosecuted for behavior during pregnancy).

67. *See* Whitner v. State, 492 S.E.2d 777, 778 (S.C. 1997) (upholding conviction under child endangerment statute for drug use during pregnancy because viable fetus is a "child" under the statute); *see also* State v. McKnight, 576 S.E.2d 168, 171 (S.C. 2003) (upholding conviction for homicide by child abuse).

68. *See* CTR. FOR REPRODUCTIVE RIGHTS, PUNISHING PREGNANT WOMEN, *supra* note 66, at 2.

69. *See* WOMEN'S LAW PROJECT, *supra* note 12, at 7; *see also* Paltrow, *supra* note 12, at 477.

70. WIS. STAT. ANN. § 940.01 (West 2005) ("[F]irst-degree intentional homicide" [provides that:] "(a) . . . whoever causes the death of another human being with intent to kill that person or another is guilty of a Class A felony.").

71. § 940.23 ("[R]eckless injury" [provides that:] "(a) Whoever recklessly causes great bodily harm to another human being under circumstances which show utter disregard for human life is guilty of a Class D felony.").

72. Deborah J.Z. "allegedly told a nurse that 'if you don't keep me here, I'm just going to go home and keep drinking and drink myself to death and I'm going to kill this thing because I don't want it anyways.'" Deborah J.Z. also expressed fear about the pain of giving birth and the baby's race. State v. Deborah J.Z., 596 N.W.2d 490, 491 (Wis. Ct. App. 1999).

73. *Deborah J.Z.*, 569 N.W.2d at 491. The baby's blood alcohol level at birth was 0.199%.

man being.[74] The court cited the ongoing debate over whether substance abuse should be addressed through treatment or punishment, noting the concern that threatening criminal prosecution could deter women from seeking prenatal care.[75] The court held that to permit prosecution of Deborah Z. would mean that "a woman could risk criminal charges for any perceived self-destructive behavior during her pregnancy that may result in injuries to her unborn child . . . [including] smoking or abusing legal medications . . . [or] 'the failure to secure adequate prenatal medical care and overzealous behavior, such as excessive exercising or dieting.'"[76]

In 1999, Regina McKnight, a homeless African-American woman with an IQ of 72 and an addiction to crack cocaine, delivered a stillborn child.[77] When she and the child tested positive for cocaine metabolites, she was charged with homicide by child abuse.[78] McKnight was convicted and sentenced to twenty years in prison.[79] The South Carolina Supreme Court upheld her conviction, rejecting McKnight's arguments that there was insufficient evidence to show causation or mens rea.[80] The court also rebuffed her argument that she was denied due process by being prosecuted for homicide when the South Carolina legislature had not enacted a statute declaring that a fetus was a child. The court relied on its prior decisions upholding convictions for felony child abuse based

74. *Id.* at 496. Wisconsin law defines a "human being" as "one who has been born alive." WIS. STAT. ANN. § 939.22 (16). The court explained its decision as required by the rule of strict construction of penal laws and by deference to the legislature in a complex public policy area. *Deborah J.Z.*, 596 N.W.2d at 494-95.

75. *Deborah J.Z.*, 596 N.W.2d at 495. The court's concern is supported by a study of low-income women who delivered their babies at an inner city hospital in Detroit, who stated their belief that if Michigan adopted a law mandating that women whose babies tested positive for drugs would be sent to jail, substance-abusing women would be less likely to seek prenatal care, drug testing, or drug treatment. When the study's authors attempted to interview women in a state with a law that threatened incarceration, all known drug users refused to participate in the study out of fear of self-incrimination. *See* Marilyn L. Poland et al., *Punishing Pregnant Drug Users: Enhancing the Flight from Care*, 31 DRUG & ALCOHOL DEPENDENCE 199, 201-02 (1993).

76. *Deborah J.Z.*, 596 N.W.2d at 494-95 (citing Hillman v. Georgia, 503 S.E.2d 610, 613 (Ga. Ct. App. 1998)).

77. Robyn E. Blummer, *Moralists' New Target: Pregnant Women*, ST. PETERSBURG TIMES, Aug. 10, 2003, at 7D; *McKnight*, 576 S.E.2d at 171, 173 (S.C. 2003).

78. *McKnight*, 576 S.E.2d at 171, 173. Under title 16, article 3, section 85 of the South Carolina Code, a person may be found guilty of "homicide by child abuse" if he "causes the death of a child under the age of eleven while committing child abuse or neglect, and the death occurs under circumstances manifesting an extreme indifference to human life." S.C. CODE ANN. § 16-3-85 (2005).

79. *See McKnight*, 576 S.E.2d at 171. The court suspended the sentence upon service of twelve years in prison.

80. The court rejected Ms. McKnight's argument that the evidence was insufficient to survive her motion for a directed verdict of acquittal, finding that there was evidence of her "extreme indifference to human life" based on her use of cocaine during pregnancy in light of South Carolina precedents which upheld felony child abuse convictions based on a woman's drug use while pregnant, holding that both she and women in South Carolina generally were on ample notice that the use of cocaine while pregnant causes fetal harm. The court also found sufficient evidence to send the case to the jury on the causation question, despite evidence that in approximately 40% of stillbirths it is impossible to make a medical judgment about the cause of death. *Id.* at 172, 73.

on a woman's drug use while pregnant as providing sufficient notice,[81] effectively ignoring her argument that her crime could not be homicide, because a fetus cannot be treated as a child under criminal law unless the legislature expressly declares it to be so.[82] The court also spurned McKnight's argument that the homicide prosecution violated her right to privacy and autonomy.[83] Further, the court rejected her argument that a twenty-year prison term for the stillbirth of a child was unconstitutional under the Eight Amendment, insisting that the proper comparison was to criminal abortion,[84] rather than to other murders.[85]

The *McKnight* decision was condemned around the nation, as unfairly singling out a poor, African-American woman for unprecedented criminal punishment, while failing to address the underlying problems of addiction and lack of health care access. Critics charged that threatening drug-abusing pregnant women with criminal prosecution, rather than providing them with social and economic support and effective drug rehabilitation, would drive women away from treatment, out of fear that they would lose their babies or be imprisoned.[86]

Similar concern was expressed when, in 2003, a Hawaii prosecutor charged Tayshea Aiwohi with manslaughter based on her use of methamphetamine while pregnant, which caused the death of her infant two days after birth.[87] Ms. Aiwohi was not charged until two years after her child's death, when she had successfully completed a drug treatment program.[88] Both the prosecutor and the trial judge spoke of the need to hold her accountable and to send a message to prevent other mothers from using drugs while pregnant.[89] The trial judge also rejected any suggestion that Aiwohi's addiction might be a mitigating factor, declaring that,

---

81. *Id.* at 176 (citing Whitner v. State, 492 S.E.2d 777 (S.C. 1997)).

82. The maximum sentence for a woman who procures an abortion in South Carolina is two years and the crime is a misdemeanor. S.C. CODE ANN. § 44-41-80(b) (2002).

83. *McKnight*, 576 S.E.2d at 176-77.

84. *Id.* at 174, 177. The court declined to address McKnight's contention that the abortion statute was applicable, saying that she had not preserved the issue for appellate review. *Id.* at 174.

85. *Id.* at 174, 177. The court compared McKnight's sentence to the sentence received by other convicted murderers in South Carolina, and murderers of children in other states. *Id.*

86. *See, e.g.*, Kirsten Scharnberg, *Prosecutors Targeting Pregnant Drug Users; Some Fear Women Will Shun Treatment*, CHI. TRIB., Nov. 23, 2003, at C1; Patrik Jonsson, *South Carolina Tests the Bounds of a Fetus's Rights*, CHRISTIAN SCIENCE MONITOR, Jun. 28, 2001, USA Section at 1.

87. *See* State v. Aiwohi, 123 P.3d 1210, 1210-13 (Haw. 2005).

88. *See id.*; *see also* State v. Aiwohi, FCCR03-1-0036 (Haw. Aug. 25, 2004), *available at* http://www.courts.state.hi.us (use search function with the name "Aiwohi" and then use the hyperlink for the case file FCCR03-1-0036).

89. *See id.* The trial judge ruled that "the State, with good reason, has served clear notice that such conduct can and will result in serious felony charges brought where the child is born alive and later dies or suffers injury due to knowing, intentional or reckless drug use." The prosecutor "hailed the judge's remarks" finding the indictment was necessary "to get justice for the baby" and to hold the mother accountable. Ken Kobayashi, *Mother Gets Probation in Ice Death*, HONOLULU ADVERTISER, Aug. 26, 2004, at 1B.

> [D]rug usage, including the use of crystal methamphetamine is a matter of choice and not an illness. Certainly it is a conscious choice to obtain and use the drug initially and worse yet, while pregnant . . . . If drug usage were an illness from the get go, we would today be in [a] medical center with a physician present in a diagnosis, treatment mode.[90]

Ms. Aiwohi pleaded no contest in order to appeal the denial of her motion to dismiss, and received a twenty-year prison sentence, which was suspended on condition that she comply with the terms of probation for the next ten years.[91] The Supreme Court of Hawaii overturned the conviction, holding that one of the elements of manslaughter was the attendant circumstance that the victim be a person at the time of the defendant's conduct.[92]

Yet even after *McKnight* and *Aiwohi*, few observers were prepared for the 2004 prosecution of Melissa Rowland for capital murder in Utah after she declined to have a recommended Caesarean section and her son was stillborn.[93] Like Ms. McKnight, Ms. Rowland was a vulnerable woman with few resources. Her own mother died shortly after birth, and Melissa Rowland had a history of serious mental illness dating from childhood, as well as substance abuse problems.[94] Ms. Rowland moved to Utah to deliver her children at the request of the adoption agency that was handling their adoption because Utah's loose adoption laws made adoption easier.[95] In Utah, she lived on Social Security disability payments and a $100 weekly stipend from the adoption agency; she also used cocaine and tobacco.[96] Ms. Rowland sought help at three hospitals because she could not feel fetal movements, but rejected their advice to have a Caesarean section (C-section).[97] None of the hospitals sought a court order requiring a C-section or made any other effort to provide Ms. Rowland with medical treatment.[98] However, after Ms. Rowland delivered a stillborn son and a living daughter, she was arrested and charged with murder.[99] After spending more than three months in jail, Rowland

---

90. *See* State v. Aiwohi, FCCR03-1-0036 (Haw. Aug. 25, 2004), *available at* http://www.courts.state.hi.us (use search function with the name "Aiwohi" and then use the hyperlink for the case file FCCR03-1-0036).

91. *See id.*

92. *See* State v. Aiwohi, 123 P.3d 1210, 1223 (Haw. 2005).

93. Linda Thomson & Pat Reavy, *Rowland's Out of Jail, Heading to Indiana*, DESERET MORNING NEWS, Apr. 30, 2004.

94. *See id.*

95. *See* Katha Pollitt, *Pregnant and Dangerous*, 278 (#16) THE NATION 9, Apr. 26, 2004.

96. *See id.*; *see also* Pamela Manson, *Mother is Charged in Stillborn Son's Death . . .*, SALT LAKE TRIBUNE, Mar. 12, 2004, at A1.

97. Thomson & Reavy, *supra* note 93; *see also* Manson, *supra* note 96. Prosecutors charged that she refused to have a Caesarean section because of cosmetic concerns that the operation would disfigure her. *See id.* But Rowland stated that she never would have said that because she had already delivered two children by Caesarean. *See* Pollitt, *supra* note 95.

98. *See, e.g.*, Thompson & Reavy, *supra* note 93; *see also* Manson, *supra* note 96.

99. *See* Manson, *supra* note 96.

pleaded guilty to two counts of felony child endangerment based on her drug use during pregnancy pursuant to a plea bargain. [100]

The Rowland case aroused a storm of controversy. [101] Prosecutors argued that Rowland's failure to undergo a C-section when she was warned that the fetuses might be harmed by a delay in their birth was a culpable omission, demonstrating the "depraved indifference to human life" necessary for a murder charge. [102] The indictment was expressly predicated on a theory of maternal "selfishness," as prosecutors argued that Ms. Rowland had refused the surgical procedure solely out of vanity. [103] The prosecutors suggested, contrary to established tort law principles of informed consent, [104] that Ms. Rowland did not have a right to decline medical treatment, but was required to "choose among alternative treatments available," rather than electing the option of no treatment. [105]

Virtually all other observers condemned the prosecution as unwarranted and legally unsound. [106] In part, they argued that it was improper and dangerous to subject anyone to the risk of criminal prosecution based on a decision to forego a potentially dangerous surgical procedure, particularly when there had been no effort to seek a court order mandating medical treatment. [107] Commentators also asserted that criminal prosecution was not the way to handle potentially risky pregnancies, as it would drive vulnerable women away from medical treatment due to fear that they would face criminal charges if they admitted to having a drug or mental health problem. [108] Similarly, vulnerable women may fear criminal charges if they underwent drug testing while receiving prenatal care or delivering the baby. [109] Further, some writers cautioned that prosecuting pregnant women under these circumstances would lead to a slippery

---

100. *See* Santini, *supra* note 25; *see also* Doug Smith & Linda Thomson, *Rowland in New Trouble*, DESERET MORNING NEWS May 27, 2004. Ms. Rowland was sentenced to two concurrent five year prison terms, with sentence suspended while on "good behavior" probation for eighteen months, requiring her to complete mental health and substance abuse treatment as well as a "parenting skills" course. *See id.* Ms. Rowland went directly from jail to a rehabilitation facility in Indiana, but left it after a month. *See id.*

101. *See* Thomson & Reavy, *supra* note 93.

102. Thomson, *Rowland Case Is Called 'Political,'* DESERET MORNING NEWS Mar. 13, 2004.

103. Associated Press, *Mother Accused of Murder of Unborn Child Pleads Not Guilty*, Mar. 15, 2004, *available at* http://www.signonsandiego.com/news/nation/20040315-1529-wst-mothercharged.html.

104. See discussion *infra* in Part II. Utah Code Ann. § 78-14-5 codifies the common law of informed consent, although it presumes that "when a person submits to health care rendered by a health care provider ... that what the health care provider did was [ ]expressly or impliedly authorized" by the patient. UTAH CODE ANN. § 78-14-5(1) (2006). However, patients may still have a cause of action for battery without meeting the requirements of § 78-14-5 if they allege that they did not consent at all to medical treatment. Lounsbury v. Capel, 836 P.2d 188 (Utah Ct. App. 1992).

105. Linda Thomson, *Mother Is Charged in Stillbirth of a Twin*, DESERET MORNING NEWS (Salt Lake City, Utah), Mar. 12, 2004.

106. *See, e.g.*, Thomson & Reavy, *supra* note 93; Matt Canham, *Proposed Law Targets Pregnant Drug Users . . .* , SALT LAKE TRIBUNE, Apr. 10, 2004, at A1.

107. *See* Thomson & Reavy, *supra* note 93; *see also* Manson, *supra* note 96.

108. *See, e.g.*, Thomson & Reavy, *supra* note 93; *see also* Canham, *supra* note 106.

109. *See, e.g.*, Thomson & Reavy, *supra* note 93; *see also* Canham, *supra* note 106.

slope: there was no principled way to distinguish the prosecution in that case from prosecution of pregnant women who smoke or who do not follow their physicians' recommendations about healthy eating.[110] Finally, observers again noted the problem of selective prosecution, since almost all those facing such criminal charges are poor and women of color.[111]

Nonetheless, prosecutors continue to bring criminal charges against women who have used drugs while pregnant, most recently in Maryland,[112] Missouri,[113] Texas,[114] and Wyoming.[115] It appears that prosecutors are more interested in scoring points with the public or in pushing the legislature to expand criminal sanctions against pregnant women who use drugs than in addressing the underlying causes of substance abuse. For example, a Wyoming prosecutor who lost his case declared, "We

---

110. Associated Press, *Arrest in C-Section Case Alarms Women's Groups*, THE HOLLAND SENTINEL, *available at* http://hollandsentinel.com (use search function and type in name of article).

111. *See, e.g.*, Roberts, *Unshackling Black Motherhood*, *supra* note 9, at 938; Chasnoff et al., *supra* note 10, at 1206.

112. *See generally* Kilmon v. State, 905 A.2d 306 (Md. 2006). The Maryland Court of Appeals has recently invalidated the prosecution of two women for reckless endangerment based on their use of cocaine while pregnant. *Id.*

113. A Missouri prosecutor charged Keila Lewis with first degree felony child endangerment, based on her newborn baby's positive test for marijuana and Lewis' admission that she smoked marijuana once while pregnant. Brief of Amici Curiae in State v. Lewis, Case 03CR113048, Chariton County, Missouri Circuit Ct. (on file with author) Ms. Lewis was charged with violating section 568.045, which provides *inter alia* that "A person commits the crime of endangering the welfare of a child in the first degree if . . . [t]he person knowingly acts in a manner that creates a substantial risk to the life, body, or health of a child less than seventeen years old . . . ." MO. ANN. STAT. § 568.045 (West 2006). The case was dismissed in 2005 because the infant's toxicology test was inadmissible under Missouri law. Personal communications with Jane Aiken, Professor, Washington University School of Law, March 7, 2005, and Jenean Thompson, Counsel for Keila Lewis, June 21, 2005.

114. In September 2003, an Amarillo, Texas prosecutor invoked a newly enacted state law when she asked local physicians to report all women who used illegal drugs while pregnant, so that they could be prosecuted for child abuse. The new law, redefined the term "individual" in certain statutes to mean "a human being who is alive, including an unborn child at very stage of gestation from fertilization until birth." S. 319, 78th Leg., Reg. Sess. (Tex. 2003). The law also redefined death to "include... for an individual who is an unborn child, the failure to be born alive." Tex. Atty. Gen. Opinion No. GA-0291, January 5, 2005, *available at* http://http://www.oag.state.tx.us/opinions/GA/GA0291.pdf. *See* Letter from Rebecca King, 47th District Attorney to all Physicians Practicing in Potter County, Texas (Sept. 22, 2003) (on file with author). The prosecutor charged at least eighteen women with crimes before the Texas Attorney General issued an Opinion concluding that the new law neither authorized prosecution for maternal drug use under the Controlled Substances Act nor required physicians to report such drug use. News from Lynn Paltrow, Executive Director of National Advocates for Pregnant Women, http://realcostofprisons.org/blog. The convictions of two women for delivery of a controlled substance to a fetus were overturned by the Texas Court of Appeals, on the ground that the prosecution did not show that the fetus possessed the drug as required by Texas law. Ward v. State, 188 S.W.3d 874 (Tex. Ct. App. 2006) and Rhonda Tulane Smith v. State, No. 07-04-0490, 2006 Tex. App. LEXIS 2370 (Tex. Ct. App. Mar. 29, 2006) (unpublished opinion).

115. *See* Associated Press, *Judge Drops 'Meth Baby' Charge*, CASPER STAR-TRIBUNE, Sept. 29, 2005, *available at* http://www.casperstartribune.com (use search function). In 2004, a Wyoming prosecutor charged Michelle Foust with causing a child to ingest methamphetamine based on blood tests of Foust and her newborn child. *See id.* The case was dismissed on the ground that the law encompassed conduct taken in regard to a "'child,' ... not a 'fetus' or 'unborn child.'" *Id.*; *see also* Associated Press, *Woman Charged with Using Meth While Pregnant Arrested Again*, May 2, 2005, *available at* http://www.billingsgazette.com (use search function).

stuck our toe in the water on this thing . . . . People need to understand there's a big hole in the law that needs to be filled." [116] It has been left to the judiciary to restrain overzealous prosecutors, through the application of the fundamental principle that legislative intent to criminalize certain behavior must be clear, and in recognition of the multiple policy considerations that argue against punishing pregnant women rather than offering them treatment.[117]

B. New Criminal Statutes: Changing the Born-Alive Rule

While the prosecutions of pregnant women for murder in *McKnight* and *Rowland* made headlines, the trend toward third party criminal liability for causing the death of a fetus has been underway for more than thirty years.[118] This change has been accomplished primarily by legislative action, as courts have been reluctant to overturn the common law "born-alive rule" without explicit legislative authorization. [119] Under the "born-alive rule," the fetus was not seen as a legal person, separate from its mother, and a homicide prosecution could not be brought.[120] Criminal

---

116. Associated Press, *Judge Drops 'Meth Baby' Charge*, *supra* note 115.

117. *See* Kilmon v. State, 905 A.2d 306 (Md. 2006). As the Maryland Court of Appeals explained in *State v. Kilmon*, to accept the prosecutor's argument to construe the reckless endangerment statute to apply to pregnant women who used drugs could mean the criminalization of:

> [N]ot just the ingestion of unlawful controlled substances but a whole host of intentional and conceivably reckless activity . . . [including] everything from becoming (or remaining) pregnant with knowledge that the child likely will have a genetic disorder that may cause serious disability or death . . . to smoking, to not maintaining a proper and sufficient diet . . . to exercising too much or too little . . . .

Id. at 311-12. The court also noted that the Maryland Legislature had considered but a penal approach to pregnant women's drug use, choosing instead to provide drug treatment for pregnant women and to consider a woman's drug use while pregnant, if she subsequently refused to enter a treatment program, as evidence supporting the termination of her parental rights. *Id.* at 312.

118. *See generally* Keeler v. Superior Court of Amador County, 470 P.2d 617 (Cal. 1970).

119. *See, e.g., id.* The "born alive" rule is of very long standing. *See id.* Lord Coke is frequently cited for his articulation of the rule:

> If a woman be quick with childe, and by a potion or otherwise killeth it in her wombe, or if a man beat her, whereby the child dyeth in her body, and she is delivered of a dead childe, this is great misprision, and no murder; but if he childe be born alive and dyeth of the potion, battery, or other cause, this is murder; for in law it is accounted a reasonable creature, *in rerum natura*, when it is born alive.

State v. Dickinson, 275 N.E.2d 599, 601 (Ohio 1971) (citing 3 Coke, Institutes 58 [1648]). Coke was followed by other British jurists, including Blackstone, Hale, Hawkins, and Sir James Stephen, and the rule passed into American common law. *Keeler*, 470 P.2d at 620.

120. *See, e.g.*, Commonwealth v. Morris, 142 S.W. 3d 654, 657 (Ky. 2004). In part this was due to difficulties in proof, because in the case of a stillbirth it could not be established beyond a reasonable doubt that the fetus had been alive when injured or that the defendant's conduct was the proximate cause of death. *See id.* (citing Clarke D. Forsythe, *Homicide of the Unborn Child: The Born Alive Rule and Other Legal Anachronisms*, 21 VAL. L. REV. 563, 575 (1987)). At the same time, the born alive rule reflected the essential unity of the pregnant woman and her fetus, and the latter's absolute dependence on her for existence. *Cf.* Dobson v. Dobson, [1999] 2 S.C.R. 753, ¶¶ 95-96 (Can.) (explaining, in the context of deciding not to impose maternal tort liability on the basis of prenatal harm, that "a pregnant woman cannot have a duty of care to her own foetus, which is at law but a part of herself . . . . [T]he physical unity of pregnant woman and foetus means that the imposition of a duty of care would amount to a profound compromise of her privacy and autonomy.").

statutes were interpreted in light of that rule, unless the statute specifically included fetuses within the class of victims.[121]

In the last few years, there has been an expanded push to characterize the harm caused by battering a pregnant woman solely as harm to the fetus, effectively erasing the woman herself.[122] New criminal laws directed at fetal harm ignore the psychic injuries imposed on the woman by such attacks. These include the fear of future domestic violence and subjugation by an intimate partner, the loss of self-determination, and the harm to the woman's interest in carrying her pregnancy to term.[123] At least nineteen states have enacted statutes authorizing a homicide prosecution for causing the death of a fetus,[124] and Congress accomplished this via the Unborn Victims of Violence Act.[125] Several states have enacted separate feticide statutes or other statutes focusing on fetal harm.[126] Some state courts have achieved the same result through judicial interpretation, rejecting the common law "born alive" rule as outmoded.[127] However, with the exception of South Carolina, each court has been

---

121. *See, e.g.*, *Dickinson*, 275 N.E.2d at 600-02. The same approach was taken with tort suits for prenatal injury and wrongful death, and inheritance proceedings. *See id.* (citing Robbins v. State, 8 Ohio St. 131 (1857)); *see also* Keeler, 470 P.2d at 627 (citing State v. McKee, 1 Add. 1 (Pa. 1797)); *see also* Tucker v. Carmichael & Sons, Inc., 65 S.E.2d 909, 910 (Ga. 1951) (citing Blackstone's Commentary on the Laws of England and the common law rule that an infant "in the mother's womb . . . is capable of having a legacy . . . made to it."). *See generally* Remy v. MacDonald, 801 N.E.2d 260 (Mass. 2004) (discussing common law cases permitting a born child to sue a third party for causing prenatal injuries and the Massachusetts wrongful death statute's applicability to a viable stillborn fetus).

122. *See* Deborah Tuerkheimer, *Conceptualizing Violence Against Pregnant Women*, 81 IND. L. J. 667, 694-97 (2006).

123. *Id.* at 669, 677-85.

124. *See* ARIZ. REV. STAT. § 13-1103(A)(5) (2006); CAL. PENAL CODE § 187(a) (West 2006); FLA. STAT. ANN. §§ 782.071, 782.09 (West 2006); 720 ILL. COMP. STAT. ANN. 5/9-1.2, 5/9-2.1, 5/9-3.2 (West 2006); MINN. STAT. ANN. §§ 609.266, 609.2661-65 (West 2006); MISS. CODE ANN. § 97-3-37 (West 2006); MO. ANN. STAT. §§ 1.205, 565.024, 565.020 (West 2006); NEV. REV. STAT. ANN. § 200.210 (West 2006); N.Y. PENAL LAW § 125.00 (McKinney 2006); N.D. CENT. CODE §§12.1-01-04, 17.1-01-04 (2006); OHIO REV. CODE ANN. §§ 2903.01-2903.07, 2903.09 (West 2006); 18 PA. CONS. STAT. ANN. § 2601-09 (West 2006); S.D. CODIFIED LAWS §§ 22-16-1, 22-16-15, 22-16-20, 22-16-41 (2006) (including definitions: 22-1-2(31), 22-1-(50A)); TENN. CODE ANN. §§ 39-13-201, 39-13-202, 39-13-210 (amended 2006), 39-13-211, 39-13-213 to 215 (West 2006); UTAH CODE ANN. § 76-5-201 (West 2006); VA. CODE ANN. § 18.2-31 (West 2006); WASH. REV. CODE ANN. § 9A.32.060(1)(b) (West 2006); WIS. STAT. ANN. §§ 939.75, 940.01-.02, 940.05-.06, 940.08-.10 (West 2006).

125. Pub. L. 108-212, 118 Stat. 568 (2004).

126. GA. CODE ANN. §§ 16-5-80, 40-6-393.1, 52-7-12.3 (West 2006); IND. CODE ANN. § 35-42-1-6 (West 2006); IOWA CODE ANN. § 707.8 (West 2006); KAN. STAT. ANN. § 21-3440 (West 2006); H. 108, 2004 Gen. Assem., Reg. Sess. (Ky. 2004); LA. REV. STAT. ANN. §§ 14:32.5-32.8 (2006) ; N.H. REV. STAT. ANN. §§ 631:1-631:2 (2006); N.M. STAT. ANN. § 30-3-7 (West 2006); N.C. GEN. STAT. ANN. § 14-18.2 (West 2006). In Virginia, killing a pregnant woman with the intent to terminate her pregnancy is capital murder. VA. CODE ANN. § 18.2-31 (West 2006).

127. *See, e.g.*, Commonwealth v. Morris, 142 S.W.3d 654 (Ky. 2004) (holding that born alive rule should be eliminated through a reinterpretation of the term "human being"); Hughes v. State, 868 P.2d 730 (Okla. Crim. App. 1994) (holding in a vehicular manslaughter case that born alive rule should be abandoned, but only prospectively); Commonwealth v. Cass, 467 N.E.2d 1324, 1326-27 (Mass. 1984) (holding that a viable fetus is a "person" within the meaning of the vehicular homicide statute, but applying it prospectively only "in order to ensure fairness to the defendant and . . . [others] who did not have the benefit of the warning provided by our construction").

careful to apply its rule prospectively only, recognizing the due process "legality" problem that would arise if a court were to change the common law and apply it to the case before it.[128]

C. Regulatory Redefinition: SCHIP–Turning the Fetus into a Child

At the same time that criminal initiatives which treat the fetus as a legally separate person have increased, the Bush Administration has adopted regulatory policies with the same goal. In 2002, the Department of Health and Services (HHS) promulgated a regulation purporting to "clarify and expand"[129] the definition of "child" under the State Child Health Insurance Program[130] by redefining "child" from "an individual under 19 years of age"[131] to "an individual under the age of 19 including the period from conception to birth."[132] The Bush Administration offered two justifications for this change. The first was to promote the birth of healthy children by expanding government coverage of prenatal care, fetal surgery, and other medical interventions which it asserted would provide "continuity of care," benefit children after birth, and ultimately save the SCHIP program money.[133] The second goal was to maximize states' "regulatory flexibility,"[134] permitting them to cover "unborn children," including those of immigrant women, who would otherwise be excluded from federal health care programs under the provisions of the

---

128. *See, e.g., Morris*, 142 S.W.3d at 654; *see also Hughes*, 868 P.2d at 730; *see also Cass*, 467 N.E.2d at 1326-27.

129. State Children's Health Insurance Program; Eligibility for Prenatal Care for Unborn Children, 67 Fed. Reg. 9936, 9937 (proposed Mar. 5, 2002) (to be codified at 42 C.F.R. 457) [hereinafter SCHIP Proposed Rule].

130. The State Children's Health Insurance Program, or SCHIP, as it is popularly known, is a joint federal state program which, for ten years beginning in 1997, provides coverage for many poor children whose parents earn too much money to qualify for Medicaid. *See* 42 U.S.C. §§ 1397aa-jj. SCHIP is authorized under Title XXI of the Social Security Act, 42 U.S.C. 1397aa-jj. Many applaud SCHIP because it gives as an expansion of health care services for poor children and does so by providing states with 70% of its costs. *See* HHS News, *States May Provide SCHIP Coverage for Prenatal Care, New Rule to Expand Health Care Coverage for Babies, Mothers*, Sept. 27, 2002, *available at* www.hhs.gov/news [hereinafter HHS News]. However, SCHIP is also frequently criticized as providing a much more meager package of health benefits than Medicaid, particularly those comprehensive Early and Periodic Screening Diagnostic and Treatment (EPSTD) benefits that are especially important for children with disabilities. *See, e.g.*, Rosenbaum, Markus & Sonosky, *supra* note 29, at 1, 3-12. Medicaid and SCHIP are both faulted, for having administrative barriers that make it difficult for enrollees to maintain eligibility. *See id.* at 10; Wendy Chavkin & Paul H. Wise, *The Data Are In: Health Matters in Welfare Policy*, 92 AM J. PUB. HEALTH 1392, 1393-94 (2002).

131. Section 2110 of SCHIP; 42 U.S.C. 1397jj(c)(1) (2006) (defining "child" as "an individual under 19 years of age").

132. State Children's Health Insurance Program; Eligibility for Prenatal Care and Other Health Services for Unborn Children, 67 Fed. Reg. 61956 (Oct. 2, 2002) (to be codified at 42 C.F.R. 457) (discussing the procedural history of the regulation, which was subsequently codified at 42 C.F.R. § 457.10) [hereinafter SCHIP Final Rule].

133. SCHIP Proposed Rule, *supra* note 129, at 9937.

134. This "state choice" mantra was evident throughout the Federal Register notice promulgating the final regulations, with the Department of Health and Services reiterating its view "that States should have the option to include unborn children as eligible targeted low income children. We are therefore retaining a revised definition [of child] . . . that permits States maximum flexibility in extending SCHIP eligibility." SCHIP Final Rule, *supra* note 132, at 61960.

Personal Responsibility and Work Opportunity Reconciliation Act of 1996 (PRWORA).[135]

Critics of the proposed regulation responded with several arguments. First, the most fundamental objection to the new SCHIP regulations was that they constituted an *ultra vires* action by HHS, since nothing in the SCHIP statute or its legislative history suggested that Congress intended "child" to have anything other than its common, everyday meaning, as one who had been born. Critics noted that Congress knew how to use the term "unborn child," as it had in the Medicare and Medicaid statutes, and that all the services covered under SCHIP, such as "well-baby and well-child care," were manifestly applicable only to born children.[136] Critics asserted that Congressional silence about "unborn children" in enacting SCHIP meant that Congress had not intended them to be covered by the SCHIP program.[137] HHS responded that the silence meant only that Congress had not "directly spoke[n] to . . . whether the term 'child' could include unborn children."[138]

Second, critics noted that when the regulations were proposed, a bipartisan coalition in Congress was already working to amend SCHIP to provide prenatal and postpartum care to pregnant women. The critics asserted that the regulatory change to include "unborn children" was unnecessary if in truth HHS' goal was simply to provide the needed care.[139]

Third, many objected to the regulation's blatant politicization of maternal and children's health.[140] By adopting a definition of "child" which was *not* in the statute and *was* inconsistent with long-standing Supreme Court precedent that a fetus was not a child,[141] the HHS regula-

---

135. Personal Responsibility and Work Opportunity Reconciliation Act of 1996, Pub. Law No. 104-193 (codified at 42 U.S.C. §§ 601-19).

136. SCHIP Final Rule, *supra* note 132, at 61961-62.

137. *See id.* at 61962.

138. *Id.*

139. Cynthia Dailard, *New SCHIP Prenatal Care Rule Advances Fetal Rights At Low-Income Women's Expense*, THE GUTTMACHER REPORT ON PUBLIC POLICY, Dec. 2002, at 3. Among the bills pending were the Mothers and Newborns Health Insurance Act of 2001, S. 724, 107th Cong. (2002); the Start Healthy/Stay Healthy Act of 2001, S. 1016, 107th Cong. (2001) and the Start Healthy/Stay Healthy Act of 2001, H.R. 3729, 107th Cong. (2002); the Immigrant Children's Health Improvement Act of 2001, S. 582, 107th Cong. (2001) and the Legal Immigrant Children's Health Improvement Act of 2001, H.R. 1143, 107th Cong. (2001), all of which proposed to amend Medicaid and SCHIP to permit states to offer health care to more infants and pregnant women, including immigrant women who were excluded from eligibility under the Personal Responsibility and Work Opportunity Reconciliation Act of 1996, 42 U.S.C. §§601 *et seq.*

140. SCHIP Final Rule, *supra* note 132, at 61957; Editorial, *A Cynical Political Act*, WASH. POST, Feb. 3, 2002, at B06; AMERICAN CIVIL LIBERTIES UNION, COMMENTS ON STATE CHILDREN'S HEALTH PROGRAM: ELIGIBILITY FOR PRENATAL CARE FOR UNBORN CHILDREN, May 6, 2002, http://www.aclu.org/reproductiverights/fetalrights/16400leg20020506.html [hereinafter ACLU COMMENTS]; Dailard, *supra* note 139, at 5.

141. In *Roe v. Wade*, the Court held that the fetus was not a person within the meaning of the Fourteenth Amendment. 410 U.S. 113, 158 (1973). In *Burns v. Alcala*, the Court held, in a case challenging the denial of welfare benefits to pregnant women under the Aid to Families with De-

tions were part of a strategy to undermine access to abortion.[142] It surely was not coincidental that HHS Secretary Tommy Thompson announced the new regulations at the Conservative Political Action Committee Conference, touting the Bush Administration's "commitment to the unborn."[143]

Fourth, critics charged that the regulations violated the First Amendment in imposing a particular theological viewpoint on the American public, i.e., that life begins at conception.[144] The critics asserted that poor pregnant women who did not believe that a fetus was a child would be forced to choose between acting in conformity with their beliefs and accepting a government benefit enshrining a different religious belief, and that this violated the Supreme Court's Free Exercise of Religion cases.[145] HHS rejected this concern, stating that "[i]f a woman has a religious objection, she simply would not accept SCHIP benefits,"[146] ignoring the reality that many poor women need these health care services.

Fifth, critics contended that the regulation was bad health policy, as it provided continuity of care for the fetus but not for the woman, who would not be entitled to post-partum care because it was her "child," and not she, who was the patient under SCHIP.[147] Not only did this devalue women by treating them as mere "vessels" for the fetus, but the regula-

---

pendent Children (AFDC) program, that pregnant women were not entitled to receive AFDC benefits because "dependent child" does not include an unborn child. 420 U.S. 575, 580 (1975). The *Burns* court stated:

> Following the axiom that words used in a statute are to be given their ordinary meaning in the absence of persuasive reasons to the contrary . . . and reading the definition "dependent child" in its statutory context, we conclude that Congress used the word "child" to refer to an individual already born, with an existence separate from its mother.

Id. at 580-81.

142. *See* Elisabeth H. Sperow, *Redefining Child Under the State Children's Health Insurance Program: Capable of Repetition, Yet Evading Results*, 12 AM. U. J. GENDER SOC. POL'Y & L. 137, 143-44 (2003).

143. *Id.* at 156-57 (citing Tommy G. Thompson, Sec'y Health & Human Servs., Compassionate Conservatism and Health Care Policy, Remarks at the Conservative Political Action Committee Conference (Jan. 31, 2002) (transcript available at http://www.hhs.gov/news/speech/2002).

144. SCHIP Final Rule, *supra* note 132, at 61963.

145. *Id.* In *Sherbert v. Verner*, 374 U.S. 398 (1963), the Supreme Court held that a state could not deny unemployment insurance benefits to a person fired for refusing to work on her Sabbath without running afoul of the First Amendment's guarantee of the free exercise of religion. *Id.* at 410. This ruling has been undercut to some extent by the Court's decision in *Employment Div. of Oregon v. Smith*, 494 U.S. 872, 878-79 890, which held that Oregon could deny unemployment benefits on the ground of work-related misconduct, when Smith, a drug counselor, used peyote (a controlled substance whose use was prohibited by Oregon criminal law) during a Native American Church ceremony.

146. SCHIP Final Rule, *supra* note 132, at 61963. However, in *Sherbert*, the Supreme Court emphasized that the exercise of religious freedom could not be predicated on a right/privilege distinction. 374 U.S. at 404-05 (citing American Commc'ns Ass'n v. Douds, 339 U.S. 382, 390 (1950); Speiser v. Randall, 357 U.S. 513 (1958)).

147. SCHIP Final Rule, *supra* note 132, at 61960, 61967-70. Further, as HHS conceded, although covering the "unborn child" meant continuity of coverage after birth, this did not change the state's normal "redetermination of eligibility" period, which could, depending on state rules, fall shortly after the child was born. *See id.* at 61964.

tion was also contrary to accepted medical practice and counterproductive for the newly born child, who depended on care from a healthy mother.[148] These critics noted that a wide range of medical conditions could affect the mother without directly affecting her fetus or newborn,[149] and that these might not be covered under the HHS regulation.[150] HHS conceded the point, declaring that "care after delivery, such as postpartum services could not be covered as part of the Title XXI [SCHIP] State Plan, . . . because they are not services for an eligible child."[151] Further, HHS informed state SCHIP administrators that states could cover "at least one postnatal visit" only if they used a "bundled fee payment" or "global fee method" in paying for pregnancy and delivery services.[152] Since only twenty-eight states use this billing method, there was concern that many new mothers would have serious gaps in their health care.[153] There were also fears that some health care providers would choose not to treat pregnant women under SCHIP at all, to avoid the ethical and malpractice issues raised by having the fetus, and not the woman, as their patient.[154]

The sixth concern was that the SCHIP change would not deliver more care to pregnant immigrant women.[155] Commentators asserted that because the new regulation did not exempt the states from their reporting obligations to the Immigration and Naturalization Service, many immigrant women would still be afraid to request services.[156]

---

148. *Id.* at 61968-70; *see also Impact on Infant and Maternal Mortality: Hearings on Uninsured Pregnant Women Before the S. Comm. on Health, Education, Labor & Pensions*, 107th Cong. 747 (2002) (statement of Senator Jeff Bingaman) ; *see also* ACLU COMMENTS, *supra* note 140.

149. These included "breast masses, influenza . . . vaccination, . . . peptic ulcer disease, . . . postpartum treatment of hemorrhage, infection, episiotomy repair, and postpartum depression." MARCH OF DIMES COMMENTS ON PROPOSED RULE TO REDEFINE CHILD UNDER SCHIP, May 6, 2002, http://www.marchofdimes.com (use search function).

150. *See, e.g., id.*

151. SCHIP Final Rule, *supra* note 132, at 61969.

152. Letter from Dennis G. Smith, Dir., Center for Medicaid and State Operations of Centers for Medicare and Medicaid Services, to State Health Officials (Nov. 12, 2002) (on file with author).

153. Dailard, *supra* note 139, at 5.

154. *Id.* at 4.

155. HHS asserted that "the new regulation makes sure that all low-income immigrants have access to important prenatal care for their babies." HHS News, *supra* note 130.

156. Dailard, *supra* note 139, at 5; SCHIP Final Rule, *supra* note 132, at 61965-66. Many immigrants continue to be deterred from seeking government supported health care for which they and/or their children are eligible, due to confusion about eligibility requirements and fear that health care providers or government insurance programs will report illegal immigrants to the federal enforcement authorities. *See* URBAN INSTITUTE, IMMIGRANT FAMILIES AND WORKERS, THE HEALTH AND WELL-BEING OF YOUNG CHILDREN OF IMMIGRANTS, Brief No. 5, Feb. 2005, at 3, *available at* http://www.urban.org/publications/310584.html; KAISER COMMISSION ON MEDICAID AND THE UNINSURED, IMMIGRANTS' HEALTH CARE COVERAGE AND ACCESS, Aug. 2003, at 2, *available at* http://www.kff.org/medicaid/2241-index.cfm; NATIONAL IMMIGRATION LAW CENTER, IMMIGRATION-FRIENDLY HEALTH COVERAGE OUTREACH AND ENROLLMENT, June 2002, at 1, *available at* http://www.nilc.org/immspbs/health/Issue_Briefs/Immigrant-Friendly_App_Enrllmnt.PDF. Indeed, some legal immigrants are denied care; *see also* Julia Preston, *Texas Hospitals' Separate Paths Reflect the Debate on Immigration*, N.Y. TIMES, July 18, 2006, at A1, A18. When the SCHIP rule was proposed, and the final rule promulgated, enforcement authority was held in the Immigration and Naturalization Service. On November 25, 2002 President Bush signed the Home-

Finally, critics argued that justifying the regulation as a means of expanding health care access[157] was disingenuous and contrary to the fiscal structure of SCHIP. Because the regulation did not (and could not) authorize additional funding for SCHIP, it could not lead to any more health care being delivered.[158] Further, because SCHIP is optional both with regard to states' decision to participate and the range of services they offer,[159] states had no incentive to add "unborn children" to their programs. In practice, these predictions have been fulfilled. Only nine states decided to cover "unborn children" under SCHIP,[160] and one of these—Rhode Island—had already included pregnant women under an HHS-approved waiver.[161] Because total SCHIP funds are capped, and many states are struggling to handle mounting Medicaid and SCHIP costs in a time of budget shortfalls,[162] they are unlikely to expand SCHIP programs.[163] Indeed, many states have been cutting or redesigning SCHIP to limit costs.

D. Compelled Medical Treatment, Civil Commitment, and Fetal Guardianships

In recent years, other efforts to "protect" fetuses from their mothers' actions during pregnancy have increased,[164] limiting women's autonomy to make medical decisions for themselves. Two-thirds of states either preclude or limit the enforcement of a woman's previously expressed wishes about foregoing medical treatment from being implemented while she is pregnant.[165] In addition, many states authorize the civil commit-

---

land Security Act of 2002, Pub. L. No. 107-296, 116 Stat. 2135, which transfered immigration enforcement authority to the Directorate of Border and Transportation Security, within the Department of Homeland Security.

157. HHS News, *supra* note 130.

158. ACLU COMMENTS, *supra* note 140; *see also* KAISER COMMISSION ON MEDICAID AND THE UNINSURED, ASSESSING THE ROLE OF RECENT WAIVERS IN PROVIDING NEW COVERAGE, Dec. 2003, at iii, *available at* http://www.kff.org/medicaid/4158.cfm [hereinafter KAISER, ASSESSING WAIVERS] (noting that SCHIP funds are capped nationally, and that "[n]ot all states have available funds to redirect toward coverage").

159. Rosenbaum, Markus & Sonosky, *supra* note 29, at 1, 17-21.

160. KAISER COMMISSION ON MEDICAID AND THE UNINSURED, IN A TIME OF GROWING NEED: STATE CHOICES INFLUENCE HEALTH COVERAGE ACCESS FOR CHILDREN AND FAMILIES, Oct. 2005, at 19, *available at* http://www.kff.org/medicaid/7393.cfm [hereinafter KAISER, IN A TIME OF GROWING NEED].

161. HHS News, *supra* note 130 (noting that Rhode Island, Colorado and New Jersey had previously obtained waivers to include pregnant women in their SCHIP programs).

162. KAISER, ASSESSING WAIVERS, *supra* note 158, at iv (observing that Medicaid and SCHIP enrollment grew by about 3.2 million people in 2002, largely "because more people became eligible for Medicaid due to the downturn in the economy").

163. Many states have cut or redesigned their SCHIP programs to limit costs. JOHN HOLAHAN ET AL., STATE RESPONSES TO 2004 BUDGET CRISES: A LOOK AT TEN STATES, URBAN INSTITUTE, Feb. 2004, at 1-2, 7-8, *available at* http://www.urban.org/UploadedPDF/410946_StateBudgetCrises.pdf; IAN HILL, HOLLY STOCKDALE & BRIGETTE COURTOT, SQUEEZING SCHIP: STATES USE FLEXIBILITY TO RESPOND TO THE ONGOING BUDGET CRISIS, URBAN INSTITUTE, June 2004, at 1-2, *available at* http://www.urban.org/url.cfm?ID=311015.

164. *See generally* Jerdee, *supra* note 4.

165. *See generally id.*

ment of pregnant women for substance abuse treatment,[166] as well as court orders that compel women to accept medical attendants at labor and delivery, or mandate particular forms of care such as Caesarean sections.

1. The Pregnancy Exception from Advanced Medical Directives

Women have regularly been denied the right to self-determination and bodily integrity by state laws that, in the name of "fetal protection," automatically invalidate advance health care directives when a woman is pregnant. Advance directives are widely seen as an important and binding vehicle that enable competent individuals to indicate the kinds of health care treatment they want should they become terminally ill, suffer a stroke or other neurological injury, or are in a persistent vegetative state, and, due to mental incapacity, can no longer make treatment decisions.[167] Advance directives are particularly important for women, because of judicial gender bias in many cases in which men, but not women, have been deemed strong, self-reliant, and courageous enough to choose death rather than a lifetime spent in a persistent vegetative state or other mentally compromised condition.[168]

Yet two-thirds of states either limit absolutely or make it more difficult to enforce women's advance directives when they become pregnant. Seventeen states provide statutory exceptions to their "living will" or health care proxy statutes which render advance directions automatically ineffective if the patient is pregnant.[169] Another sixteen states render the living will or health care proxy inapplicable in a variety of circumstances, ranging from a possibility to a probability that the fetus will "develop to a live birth."[170] Minnesota gives a slight bow to women's autonomy by establishing a rebuttable presumption that a pregnant woman would want health care to be provided if there is a "real possibility [that] . . . the fetus could survive to the point of life birth," even if "the withholding or withdrawal of such health care would be authorized were she not pregnant."[171] The presumption can be rebutted by an ex-

166. *See, e.g.*, WIS. STAT. §§ 48.205 (2006) (permitting the civil commitment of pregnant girls and women, dubbed "The Cocaine Mom law"); *see also* Tom Kertscher, *'Cocaine Mom' Law Invoked in Attempt to Detain Woman, Racine Case Thought to Be First Time Law is Used Without Other Crime*, MILWAUKEE JOURNAL SENTINEL, Nov. 5, 1999, at 1.

167. *See* Linda C. Fentiman, *Privacy and Personhood Revisited: Substitute Decisionmaking for the Incompetent Incurably Ill Adult*, 57 GEO. WASH. L. REV. 801, 818-828 (1989) (discussing living wills and documents designating a medical treatment agent); *see also* BARRY R. FURROW ET AL., *supra* note 29, at 842.

168. *See generally* Steven H. Miles & Allison August, *Courts, Gender, and "The Right to Die,"* 18 L. MED. & HEALTH CARE 85 (1990).

169. Jerdee, *supra* note 4, at 978 n.35.

170. *Id.* at 978-79 nn.36-44. The Alaska statute cited in n.37, ALASKA STAT § 18.12.040, was repealed in 2004.

171. The Minnesota law states in pertinent part:

> When a patient lacks decision-making capacity and is pregnant, and in reasonable Medical judgment there is a real possibility that if health care to sustain her life and the life of

plicit statement to the contrary in the advance directive itself, or by clear and convincing evidence presented at a hearing.[172] While Minnesota's law endeavors to strike a balance between the woman's interest in autonomy and the provision of a living maternal body in which the fetus can continue to develop, it still enshrines a normative view of women—that any "reasonable" woman would choose to continue on life-support if it meant that her fetus would survive until birth.

The potential impact of these statutes is substantial. In *University Health Services v. Piazzi*,[173] a Georgia trial court granted a hospital's petition to continue life support for a brain-dead pregnant woman over the objections of the woman's husband and other family members,[174] so that a fetus could be delivered. Although Ms. Piazzi had not executed an advance directive, the court reached out to decide the case based on the Georgia Natural Death Act, which rendered a pregnant woman's advance directive inoperable if the fetus was viable.[175] The court declared that because Ms. Piazzi was brain dead, she no longer had a constitutionally protected right to privacy,[176] but implied that even if she did, any interest she had would be rendered irrelevant by the Georgia living will statute.[177]

In contrast to the *Piazzi* court's expansive interpretation, two other courts have rebuffed constitutional challenges to state advance directive statutes, dismissing the cases for lack of justiciability. In *DiNino v. Gorton*,[178] a woman alleged that the Washington Model Health Care Directive Act was unconstitutional because it made advance directives ineffective during pregnancy.[179] DiNino drafted an advance directive contrary to Washington law, and her physician refused to place it in her medical file, asserting fears of potential liability.[180] DeNino sought a declaratory judgment that her advance directive was valid and enforce-

---

> the fetus is provided the fetus could survive to the point of live birth, the health care provider shall presume that the patient would have wanted such health care to be provided, even if the withholding or withdrawal of such health care would be authorized were she not pregnant. This presumption is negated by health care directive provisions described in section 145C.05, subdivision 2, paragraph (a), clause (10), that are to the contrary, or, in the absence of such provisions, by clear and convincing evidence that the patient's wishes, while competent, were to the contrary.

MINN. STAT. § 145C.10 (g) (2006).

172. *Id.*

173. This unreported case, No. CV86-RCCV-464, was cited in Daniel Sperling, *Maternal Brain Death*, 30 AM. J. L. & MED. 453, 494-95 (2004) and Molly C. Dyke, *A Matter of Life and Death: Pregnancy Clauses in Living Will Statutes*, 70 B.U. L. REV. 867, 871-72 (1990).

174. Daniel Sperling, *Panel Discussion American Association of Law Schools Panel: Panel on the Use of Patients for Teaching Purposes Without Their Knowledge or Consent: Article: Do Pregnant Women Have (Living) Will?*, 8 J. HEALTH CARE L. & POL'Y 331, 336 (2005).

175. GA. CODE ANN. § 31-32-3(b) (2006).

176. Sperling, *supra* note 173, at 336-37.

177. *Id.* at 337.

178. 684 P.2d 1297 (Wash. 1984) (en banc).

179. *DiNino*, 684 P.2d at 1299.

180. *Id.* at 1298-99.

able, and that her physician would not be held liable if he acted in accordance with its provisions. In the alternative, she argued the statute violated her Fourteenth Amendment rights to privacy, which encompassed her right to seek abortion and to forego medical treatment.[181] Her suit was dismissed as presenting a "purely hypothetical and speculative controversy."[182] In *Gabrynowicz v. Heitkamp*,[183] the plaintiffs challenged the constitutionality of two North Dakota statutes which rendered a woman's advance directive inoperable while she was pregnant, and further mandated that pregnant women receive medical treatment to permit the "continuing development and live birth of the unborn child."[184] While the United States District Court for North Dakota acknowledged the statutes' potential constitutional problems, it ruled that because the plaintiff was neither pregnant nor in a terminal condition, her case was non-justiciable.[185] What both *DiNino* and *Gabrynowicz* overlook, of course, is that the entire point of a living will or advance directive statute is to permit competent adults to announce their wishes for treatment prior to becoming incapacitated, when they are able to think through their choices. Denying a woman the opportunity to bring a constitutional challenge while healthy and non-pregnant will mean, in practical terms, that she will never be able to challenge the law.

2. Civil Commitment for Substance Abuse and Other Treatment

Thirty-four states and the District of Columbia currently permit pregnant women to be civilly committed for alcohol and other drug abuse.[186] Most of these state statutes address substance abuse in the same way that statutes authorize the civil commitment of the mentally ill when the mental illness poses danger to "self or others."[187] Three states' statutes specifically target pregnant women.[188] For example, the Minnesota statute requires physicians to report pregnant patients' use of alcohol and controlled substances during pregnancy and mandates toxicology

181. *Id.* at 1299.

182. *Id.* at 1300.

183. 904 F. Supp. 1061 (N.D. 1995).

184. *Gabrynowicz*, 904 F. Supp. at 1062. One of the statutes being challenged stated:

> Notwithstanding a declaration executed under this chapter, medical treatment must be provided to a pregnant patient with a terminal condition unless . . . such medical treatment will not maintain the patient in such a way as to permit the continuing development and live birth of the unborn child or will be physically harmful or unreasonably painful to the patient or will prolong severe pain that cannot be alleviated by medication.

N. D. CENT. CODE § 23-06.4-07(3) (repealed 2005)

185. *Gabrynowicz*, 904 F. Supp. at 1063-64.

186. *See supra* note 37 and statutes cited therein.

187. *See, e.g.*, IOWA CODE § 125.81 (2005) (authorizing the immediate custody of a chronic substance abuser, defined in § 125.2 (2005) as a person who "[h]abitually lacks self control as to the use of chemical substances to the extent that the person is likely to endanger the person's health, or to physically injure the person's self or others, if allowed to remain at liberty without treatment" and "[l]acks sufficient judgment to make responsible decisions with respect to the person's hospitalization or treatment").

188. MINN. STAT. § 626.5561-63 (2005); OKL. STAT. tit. 63 § 1-546.5 (2005); S.D. CODIFIED LAWS § 34-20A-70 (2006).

testing of mothers and newborns shortly after delivery if there is reason to believe that a mother has used controlled substances.[189] At least one other state, Wisconsin, has amended its child abuse law to permit the involuntary commitment of pregnant women who are abusing alcohol or controlled substances throughout pregnancy to the extent constitutionally permissible.[190]

The Wisconsin law, popularly dubbed the "Cocaine Mom Bill," was a reaction to the Wisconsin Supreme Court's decision in *State ex rel. Angela M.W. v. Kruicki*.[191] When Angela M.W. was pregnant, her physician reported her to child welfare authorities after she tested positive for drugs and failed to show up for an appointment.[192] The court held that Ms. M.W. could not be compelled to participate in in-patient drug treatment through a "child custody" proceeding, concluding that the legislature had not intended the term "child" in the child neglect statute to include children not yet born.[193]

The Arkansas Supreme Court relied on *Angela M.W.* in *Arkansas Dep't of Human Services v. Collier*,[194] a case which challenged a family court judge's authority to declare that the fetus of a pregnant woman who was using methamphetamines was a dependent-neglected child.[195] Acting *sua sponte*, the trial judge ordered the "child" to be taken into custody, by holding the woman at a county detention center until she went into labor.[196] The Arkansas Supreme Court held that the judge's actions exceeded her statutory authority.[197]

In sharp contrast to these two decisions, in 2000 a Massachusetts family court judge ordered the involuntary medical treatment of Rebecca Corneau, a pregnant woman who rejected all medical care and was suspected of membership in a religious cult that denied children adequate nutrition and medical care.[198] The court ordered Ms. Corneau to be taken to a prison hospital and compelled to submit to a medical examination to determine her health, pregnancy status, and the anticipated birth date of her "unborn child," so that the court could determine what prena-

189. MINN. STAT. § 626.5561 (2005).
190. WIS. STAT. § 48.01(1) (2006).
191. 561 N.W.2d 729 (Wis. 1997).
192. *State ex rel. Angela M.W.*, 561 N.W.2d at 732-33.
193. *Id.* at 737.
194. 95 S.W.3d 772 (2003).
195. *Collier*, 95 S.W.3d at 773.
196. *Id.* at 773; *see also* Bennett v. Collier, 95 S.W.3d 782, 785 (2003) (holding in a related civil contempt proceeding that the trial judge lacked authority to hold the pregnant woman in contempt of court once the judge terminated her parental rights in a different case).
197. *Collier*, 95 S.W.3d at 781 (rejecting arguments that the term "child" in the child neglect statute should be construed to include a fetus even though other areas of Arkansas law authorized actions to be brought on behalf of a fetus).
198. Dave Wedge & Edwin Molina, *Experts: Ruling May Just Open Up Floodgates*, BOSTON HERALD, Sept. 1, 2000, at 4.

tal treatment should be ordered.[199] Ms. Corneau did not appeal the judge's ruling and she was imprisoned until after her child's birth, when the child was declared neglected and Corneau's parental rights were terminated.[200]

3. Compelled Caesarean Sections

Women continue to be forced to have Caesarean sections over their objections.[201] In *Jefferson v. Griffin Spaulding County Hospital*,[202] the Supreme Court of Georgia upheld a trial judge's decision to order a woman who was close to her delivery date to undergo a Caesarean section, based on hospital physicians' concern that because the fetus was in a breech position, it was likely to die in childbirth.[203] The court rejected Ms. Jefferson's religious objections to the procedure, but before the order could be enforced, she gave birth to a healthy child.[204] In the landmark case of *In re A.C.*,[205] the District of Columbia Court of Appeals took a different view, declaring that pregnant women had the right to control their medical treatment, even when it could affect the health of the fetus.[206] The court stated, "[s]urely, . . . a fetus cannot have rights in this respect superior to those of a person who has already been born."[207]

In 1996, a Florida trial court held an *ex parte* hearing concerning Laura Pemberton, a pregnant woman who refused medical advice to deliver by Caesarean section because it was feared that her previous cae-

199. *See* Barbara F. v. Bristol Div. of Juvenile Court, 735 N.E. 2d 357 (Mass. 2000) (discussing the juvenile court's order in the course of rejecting a challenge to that court's decision brought by another pregnant woman, who asserted that the juvenile court's order had a chilling effect on her own conduct). *See also* Wedge & Molina, *supra* note 198, at 4; Dave Wedge, *Appeal Cites Abortion Ruling in Bid to Free Pregnant Cultist*, BOSTON HERALD, Sept. 6, 2000, at 10; Michael Paulson, *Fetus Dispute Brings Wider Issues to Fore*, BOSTON GLOBE, Sept. 10, 2000, at B5; Miller, *supra* note 34, at 71-74.

200. Miller, *supra* note 34, at 73-74 nn. 31 & 34.

201. For illuminating analyses of the issues raised by these cases, see Nancy K. Rhoden, *The Judge in the Delivery Room: The Emergence of Court-Ordered Cesareans*, 74 CAL. L. REV. 1951 (1986); Roberts, *Punishing Addicts Who Have Babies*, *supra* note 9, at 964. Women of color have borne the brunt of these interventions. *Id.* at 939.

202. 274 S.E.2d 457 (Ga. 1981).

203. *Jefferson*, 274 S.E.2d at 458.

204. *Id.* at 459. *See also* Alicia Ouellette, *New Medical Technology: A Chance to Reexamine Court-Ordered Medical Procedures During Pregnancy*, 57 ALB. L. REV. 927, 940 (1994).

205. 573 A.2d 1235 (D.C. 1990).

206. *See In re A.C.*, 573 A.2d at 1252.

207. *Id.* at 1244. The case arose out of tragic circumstances. In 1987, Angela Carder, a twenty-seven year old woman who had survived cancer as a teenager, became pregnant. At twenty-five weeks of pregnancy, she was found to be suffering from cancer again. *Id.* Her condition deteriorated rapidly, and at twenty-six weeks, she was hospitalized and expected to die shortly. When the hospital sought a court order to perform a Caesarean section, a trial court held a hearing, and ordered a Caesarean section to be performed. The baby died within two hours, and Carder died two days later. *Id.* at 1238. On appeal, the District of Columbia Court of Appeal declared first that the pregnant woman had the right to self-determination, which included the right to refuse a Caesarean section. The court further held that in cases such as this one, where it was difficult to ascertain the pregnant woman's wishes because she was close to death, under the influence of medication, and suffering extreme emotional distress, the trial court should undertake a substituted judgment analysis, so that it could decide what the patient would have chosen if she were competent. *Id.* at 1247-49.

sarean section put her at higher risk for uterine rupture.[208] The court ordered Ms. Pemberton to be transported to the hospital via ambulance against her will, and then continued the hearing in her presence, ultimately ordering the Caesarean section to be performed.[209] When Ms. Pemberton subsequently brought a section 1983 civil rights action an action in federal court, the court held that Ms. Pemberton's constitutional rights "did not outweigh Florida's interest in preserving the life of the unborn child."[210]

In a less dramatic case, in 2004 a Pennsylvania woman, Amber Marlowe, who had previously given birth to six children, each weighing more than eleven pounds, went to deliver her seventh child, only to be told that the fetus was so large that it was not safe to deliver vaginally.[211] When Marlowe declined the hospital's advice to have a Caesarean section, the hospital obtained an *ex parte* order compelling her to have the surgery.[212] Marlowe and her husband did not learn of the order until after she had safely delivered her baby at another hospital.[213]

These court-ordered detentions and medical interventions are contrary to the prevailing view of medical professionals that medical treatment against the pregnant woman's wishes is rarely, if ever, appropriate.[214] The American College of Obstetrics and Gynecology (ACOG), for example, emphasizes that medical judgment is limited and fallible, that pregnant women do not lose their rights to autonomy merely by becoming pregnant, and, most importantly, that doctors who believe that a medical intervention will benefit the fetus and/or the mother should exhaust all possible avenues of explanation and persuasion before seeking a court order, including consultation with an institutional ethics committee.[215]

Significantly, medical groups recognize the need to view pregnant women's decisionmaking in its full social and economic context. ACOG alerts physicians to the risk that "clinicians' conclusions reinforce exist-

208. Pemberton v. Tallahassee Mem'l Reg'l Med. Ctr., Inc., 66 F. Supp. 2d 1247 (N.D. Fla. 1999). At the hearing, five physicians testified that the risk of uterine rupture during vaginal delivery was too high (from "four to six percent"), and that if there was rupture, the fetus was likely to die during delivery. *Pemberton*, 66 F. Supp. at 1252-53. Neither the state nor federal court addressed the question of informed consent, that is, who was entitled to make the fetal risk assessment. *See generally id.* at 1251-57 (discussing only the issues of substantive due process, procedural due process, professional negligence, and false imprisonment).

209. *Id.* at 1250.

210. *Id.* at 1251.

211. David Weiss, *Court Delivers Controversy, Mom Rejects C-Sections; Gives Birth on Own Terms*, TIMES LEADER Jan. 16, 2004; Associated Press, *New Questions about Childbirth Rights*, May 19, 2004, *available at* http://keyetv.com/health/health_story_140110423.html.

212. Weiss, *supra* note 211.

213. *Id.*

214. *See* AMERICAN MEDICAL ASSOCIATION, *supra* note 36; AMERICAN COLLEGE OF OBSTETRICS & GYNECOLOGY, *supra* note 36, at 35; American Academy of Pediatrics, *supra* note 36.

215. AMERICAN COLLEGE OF OBSTETRICS & GYNECOLOGY, *supra* note 36, at 34-35.

ing gender, class, or racial inequality."[216] The American Medical Association (AMA) agrees that legal interventions during pregnancy are seldom, if ever, proper,[217] and emphatically rejects "criminal sanctions or civil liability" for pregnant women based on their conduct "toward [their] . . . fetus."[218] Instead, the AMA endorses a comprehensive, long-term approach to substance abuse, to begin in adolescence and continue through pregnancy and beyond, in recognition of the fact that addiction to alcohol and other drugs is a disease.[219]

Many physicians believe that seeking judicial intervention to compel women to accept treatment during pregnancy is counter-productive, not only leading to a loss of trust in the health care system on the part of the particular woman who is the object of the intervention, but also deterring other women from seeking care.[220] This problem is compounded when women are forcibly restrained while an unwanted medical procedure is performed,[221] or when pregnant women are confined in prison hospitals, which often meet only the most minimal standards of health care and, at the same time, make it possible for a woman to have access to drugs.[222]

4. Fetal Guardians

Another way that courts have separated pregnant women from their fetuses, undermining women's right to self-determination, is through the appointment of fetal guardians. Sometimes such appointments are the means used to compel women to accept medical treatment to benefit the

---

216. *Id.* at 35.

217. AMERICAN MEDICAL ASSOCIATION, *supra* note 36 (stating that a physician might seek judicial action to override a woman's informed treatment refusal only "[i]f an exceptional circumstance could be found in which a medical treatment poses an insignificant or no health risk to the woman, entails a minimal invasion of her bodily integrity, and would clearly prevent substantial and irreversible harm to her fetus").

218. *Id.*

219. AMERICAN MEDICAL ASSOCIATION, POLICY H-420.976: ALCOHOL AND OTHER SUBSTANCE ABUSE DURING PREGNANCY, *available at* http://www.ama-assn.org/ama/noindex/category/11760.html (click "accept"; search "420.976"): AMERICAN MEDICAL ASSOCIATION, POLICY H-95.976: DRUG ABUSE IN THE UNITED STATES – THE NEXT GENERATION, *available at* http://www.ama-assn.org/ama/noindex/category/11760.html (click "accept"; search "95.976").

220. Helene Cole, *Legal Interventions During Pregnancy*, 264 J.A.M.A. 2663, 2667 (1990); AMERICAN COLLEGE OF OBSTETRICS & GYNECOLOGY, *supra* note 36, at 36. Similar concerns have been raised about mandatory maternal HIV testing in order to reduce maternal-fetal transmission of the virus. *See, e.g.*, Comm. on HIV Prenatal/Newborn Testing, *Prenatal/Newborn HIV Testing*, 49 THE RECORD OF THE ASS'N OF THE BAR OF THE CITY OF NEW YORK 420 (1994).

221. *See, e.g.*, *In re* Fetus Brown, 689 N.E.2d 397, 400 (Ill. App. 1997) (reversing a trial court's decision to appoint a hospital the "temporary custodian of Fetus Brown, with the right to consent to one or more blood transfusions for [his mother] when advised of such necessity by any attending physician").

222. Cole, *supra* note 220.

fetus.[223] More frequently, they are a direct challenge to a pregnant woman's right to have an abortion.[224]

In Alabama, trial judges have appointed fetal "guardians" as part of their procedures for determining whether pregnant teenagers can obtain an abortion without parental consent, a practice whose legality the Alabama Supreme Court has declined to address.[225] Such fetal guardianships turn what may already be an intimidating but parentalistic proceeding into an adversarial one, effectively changing the burden of proof established by the legislature in authorizing a judicial bypass proceeding.[226] For example, in one case, the trial court appointed a guardian ad litem for the fetus, whom the court denominated "Baby Teresa."[227] The guardian ad litem called three witnesses who worked at organizations opposed to abortion, "Sav-A-Life" and the "COPE Crisis Pregnancy Center," and cross-examined the minor about whether she had consulted these or other anti-abortion groups in making her decision.[228] In another case, a judge (perhaps the same one) declared,

> I have . . . as has been my practice for five years now, appointed a lawyer to represent *your unborn child*, because I do not feel that the court should be placed in the position of being a cross-examiner, an advocate for one side or the other, so I've appointed someone to represent *the silent voice* in this case.[229]

In 2003, the wife of a Florida prosecutor petitioned to be appointed guardian of the fetus in the case of a severely retarded woman who had been raped at the group home where she resided, in order to prevent the woman from having an abortion.[230] The rape victim suffered from autism, cerebral palsy, and a seizure disorder, in addition to her retardation,

---

223. *See, e.g.*, *In re* Fetus Brown, 689 N.E.2d at 400, 406 (reversing a trial court's decision to appoint a hospital the "temporary custodian of Fetus Brown" and appointing the Cook County, Illinois Public Guardian as the "guardian ad litem" for Fetus Brown in order to override the pregnant woman's refusal of blood transfusions, based on her religious beliefs).

224. *See* Benten v. Kessler, 1192 U.S. Dist. LEXIS 14747 (E.D.N.Y. 1992) (holding in a pregnant woman's class action suit to enjoin the Food and Drug Administration's ban on the importation of RU-486, an abortifacient drug, that an attorney could not intervene on behalf of the woman's non-viable fetus, declaring that "[t]he Federal Rules of Civil Procedure make no provision for appointing a guardian ad litem for a fetus"). *See also In re* Nancy Klein, 1989 N.Y. App. Div. LEXIS 1613 (N.Y. App. Div. 1989) (holding, in the case of a comatose woman injured in a accident, that her husband should be appointed guardian with authority to terminate her pregnancy, and that a stranger could not be appointed guardian of her fetus, since a non-viable fetus was not "a legally recognized 'person,'" (citing Roe v. Wade, 410 U.S. 113 (1973))).

225. *In re* Anonymous, 810 So. 2d 786, 795 (Ala. 2001).

226. As one trial judge explained, "it has been my practice for three years now when I'm faced with these cases to not only have a lawyer for you but to have a lawyer to represent the interest of the unborn child." *Id.* at 789. *See also In re* Anonymous, 889 So. 2d 525, 525 (Ala. 2003) (Johnstone, J., dissenting, in discussing evidence of the trial court's bias against abortion).

227. *In re* Anonymous, 733 So. 2d 429, 429-30 (Ala. Civ. App. 1999).

228. *In re Anonymous*, 733 So. 2d at 429-30.

229. *In re Anonymous*, 889 So. 2d at 527 (Johnstone, J., dissenting, in discussing evidence of the trial court's bias against abortion) (emphasis added).

230. *In re* J.D.S., 864 So. 2d 534, 536 (Fla. Dist. Ct. App. 2004).

but the putative guardian only expressed concern that the woman was taking medications which could injure the fetus.[231] Both the trial judge and the mid-level appellate court rebuffed the request for guardianship, holding that Florida law did not authorize such an appointment.[232] They noted that Florida law instead authorized guardians to be appointed for mentally incompetent women to protect *their* interests, with the guardian to consider whether an abortion was appropriate as "necessary to save the life or preserve the health of the pregnant woman," subject to court approval.[233] In response to the decision in *J.D.S.*, Governor Jeb Bush announced his intention to seek a change in Florida law, but these efforts have been unsuccessful so far.[234]

E. New Abortion Laws

New laws seeking to limit abortions reflect new fetal protection in two different ways. First, in addition to the federal Partial Birth Abortion Act,[235] a number of states, as well as Congress, have enacted laws which seek to limit women's access to necessary health care by prohibiting so-called "partial birth" abortions, a term which suggests that the fetus in such cases is fully formed or capable of being born.[236] Second, laws have been proposed in Congress and state legislatures which, in the name of requiring "informed consent," make the fetus more vivid and alive, with the goal of discouraging women from choosing abortion.

So-called "partial birth" abortion laws are criticized on four major grounds.[237] First, and foremost, these statutes do not acknowledge that most "partial birth" abortions take place under urgent or even emergent medical circumstances. These include the discovery that the fetus is anencephalic, hydrocephalic, or suffers from a fatal genetic defect, or the determination that an alternative abortion procedure would put the mother's health at risk.[238] If "partial birth" abortions are prohibited,

231. *In re J.D.S.*, 864 So. 2d at 536.

232. *Id.*

233. *Id.* at 539 (discussing the requirements of FLA. STAT. ANN. 390.0111(1) and (3) (West 2006)).

234. Bob Mahlburg, *Senate Chief Wary of Fetus Guardian Bill [-] Governor Finds House Support, Won't Back Down on Legislation*, ORLANDO SENTINEL, Jan. 28, 2004, at B1.

235. 18 U.S.C.A. § 1531 (West 2006).

236. The Virginia law, for example, refers to the types of abortion prohibited as "partial-birth infanticide." VA. CODE. ANN. § 18.2-71.1 (West 2006).

237. A full discussion of "partial-birth" abortion bans is beyond the scope of this article; in any case, the Supreme Court will consider the constitutionality of the federal law next term. Carhart v. Gonzales, 413 F.3d 791 (8th Cir. 2005), *aff'g* Carhart v. Ashcroft, 331 F. Supp. 2d 805 (D. Neb. 2004), *cert. granted sub nom.* Gonzales v. Carhart, 126 S. Ct. 1314 (2006).

238. PLANNED PARENTHOOD, ABORTION AFTER THE FIRST TRIMESTER, *available at* http://www.plannedparenthood.org/news-articles-press/politics-policy-issues/abortion-access/trimester-abortion-6140.htm (citing SHELDON CHERRY & IRWIN MERKATZ, COMPLICATIONS OF PREGNANCY: MEDICAL, SURGICAL, GYNECOLOGIC, PSYCHOSOCIAL, AND PERINATAL (4th ed. 1991) and MAUREEN PAUL, A CLINICIAN'S GUIDE TO MEDICAL AND SURGICAL ABORTION (1999)). Medical complications for the woman include infections, heart failure, malignant hypertension, uncontrolled diabetes, renal disease, depression and suicidal tendencies. *Id.* In addition, many

women may be rendered sterile or otherwise unable to have other children, suggesting that "partial birth" abortion bans are not only anti-female, but also "anti-life" in practice.[239] Second, these "partial birth" laws describe abortion procedures using language inconsistent with medical parlance, making it impossible for a physician to know whether the technique used in a particular case is proscribed and thus rendering the law unconstitutionally vague.[240] Third, the laws' lack of an exemption for the mother's health is incompatible with prior decisions of the Supreme Court, which has held repeatedly that the state's interest in protecting the potential life of a fetus cannot supersede its interest in protecting the health and life of the pregnant woman.[241] Finally, these laws are not limited to "late" or third trimester abortions, but could apply as early as twelve weeks, when a fetus would clearly not be viable, and thus are incompatible with *Roe v. Wade*[242] and *Planned Parenthood of Southeastern Pennsylvania v. Casey*.[243]

Recently, proposals have been offered which, while nominally permitting abortion, conceptualize the fetus as a child in order to encourage women not to have abortions. The Unborn Child Pain Awareness Act of 2005 would require that abortion providers inform all women seeking abortion that Congress has determined that an "unborn child" may experience pain at twenty weeks or more after fertilization,[244] give each

---

women face barriers that delay their access to abortion until the second trimester, including poverty, partner abuse, geographic difficulty in finding an abortion provider, and teenage status. *See* Brief for Seventy-Five Organizations Committed to Women's Equality as Amici Curiae Supporting Respondents, Stenberg v. Carhart, 530 U.S. 914 (2000) (No. 99-830), 2000 WL 340122.

239. Forcing physicians to use other, older abortion techniques increases the risk of infection and laceration as an option increases the chance of uterine perforations, cervical lacerations, hemorrhaging, and infection, all of which can lead to sterility. *See* Women's Med. Prof'l Corp. v. Voinovich, 911 F. Supp. 1051, 1069-70, *aff'd* 130 F.3d 187 (1997).

240. Stenberg v. Carhart, 530 U.S. 914, 939-46 (2000); *see also* Carhart v. Gonzales, 413 F.3d 791 (8th Cir. 2005), *aff'g* Carhart v. Ashcroft, 331 F. Supp. 2d 805 (D. Neb. 2004), *cert. granted sub nom.* Gonzales v. Carhart, 126 S. Ct. 1314 (2006).

241. *Stenberg*, 530 U.S. at 930-31; Thornburgh v. Am. Coll. of Obstetricians & Gynecologists, 476 U.S. 747, 768-69 (1986); Richmond Med. Ctr. for Women v. Hicks, 409 F.3d 619 (4th Cir. 2005).

242. 410 U.S. 113, 165-66 (1973).

243. 505 U.S. 833, 879 (1992).

244. Before an abortion could be performed, the woman would have to be told the following:

> You are considering having an abortion of an unborn child who will have developed, at the time of the abortion, approximately **XX** weeks after fertilization. The Congress of the United States has determined that at this stage of development, an unborn child has the physical structures necessary to experience pain. There is substantial evidence that by this point, unborn children draw away from surgical instruments in a manner which in an infant or an adult would be interpreted as a response to pain. Congress finds that there is substantial evidence that the process of being killed in an abortion will cause the unborn child pain, even though you receive a pain-reducing drug or drugs. Under the Federal Unborn Child Pain Awareness Act of 2005, you have the option of choosing to have anesthesia or other pain-reducing drug or drugs administered directly to the pain-capable unborn child if you so desire. The purpose of administering such drug or drugs would be to reduce or eliminate the capacity of the unborn child to experience pain during the abortion procedure. In some cases, there may be some additional risk to you associated with administering such a drug.

S. 46, 109th Cong. §2902 (2005).

woman an Unborn Child Pain Awareness Brochure,[245] and require her to sign an Unborn Child Pain Awareness Decision Form, indicating whether or not she wishes to have her "pain-capable unborn child" to receive anesthesia.[246] Other proposals would set an even earlier date for a woman contemplating abortion to be told about the fetus' potential to experience pain.[247] At least one law would impose a twenty-four hour waiting period in which the pregnant woman is to consider the information regarding fetal pain prior to having the abortion,[248] exacerbating the pressure on women not to have an abortion.[249]

Although legislative mandating of explicit informed consent requirements is not unprecedented,[250] the Unborn Child Pain Awareness Act of 2005 is unusual in its detailed explanation of the procedure contemplated.[251] All other laws mandating the specifics of the informed consent dialogue require informing the patient of the consequences of the contemplated medical procedure *to her*, not a third party. Imagine, for example, a statute requiring a prospective kidney donor be told about the impact of the decision to donate on the donor's child, because of the risk that the donor might die after donating the kidney. Courts have increasingly recognized that as competent adults, parents are able to accept or reject medical treatment based on their personal views of what is best for them, and have not required them to take their children's interests into account.[252]

---

245. This Brochure would be developed by the Department of Health and Human Services, and include: "the same information as required under the statement under subsection (b)(2)(A)(i), including greater detail on her option of having a pain-reducing drug or drugs administered to the unborn child to reduce the experience of pain by the unborn child during the abortion." S. 46, 109th Cong. §2902 (2005).

246. The law would require that the woman sign an Unborn Child Pain Awareness Decision Form, which shall:

> (A) with respect to the pregnant woman—
> (i) contain a statement that affirms that the woman has received or been offered all of the information required in subsection (b);
> (ii) require the woman to explicitly either request or refuse the administration of pain-reducing drugs to the unborn child; and
> (iii) be signed by a pregnant woman prior to the performance of an abortion involving a pain-capable unborn child.

S. 46, 109th Cong. §2902 (2005).

247. Under the Montana Unborn Child Pain Prevention Act, all women contemplating abortion at 16 weeks or later would have to be informed that the fetus could feel pain. H.B. 238, 59th Leg., Reg. Sess. (Mont. 2005).

248. H.B. 238, 59th Leg., Reg. Sess. (Mont. 2005).

249. Imposing a waiting period is believed to add a significant barrier to abortion access. PLANNED PARENTHOOD, *supra* note 228.

250. *See, e.g.*, S.G. Nayfield et al., *Statutory Requirements for Disclosure of Breast Cancer Treatment Alternatives*, 86 J. NAT'L CANCER INST. 1202 (1994) (discussing state laws that require that women contemplating mastectomy be told of the range of treatment options available); and similar laws governing hysterectomy (CAL. HEALTH & SAFETY CODE § 1690 (West 2006)) and sterilization (OR. REV. STAT. ANN. § 436.225 (West 2006)).

251. *Cf.* OR. REV. STAT. §§ 127.800–.897 (1998) (popularly known as Oregon's Death with Dignity Law).

252. *See, e.g.*, Norwood Hospital v. Munoz, 564 N.E.2d 1017, 1024 (Mass. 1991) (holding that a mother who was a Jehovah's Witness could decline to receive a medically recommended blood

Several legislators have introduced laws which compel women to contemplate their fetuses as "unborn" children in other ways. These include laws that require that women be given the opportunity to visualize their fetus in a sonogram and to hear the fetus' heartbeat, as a precondition to the "informed consent" necessary to receive an abortion,[253] as well as laws that provide federal funding to purchase sonogram equipment for this purpose.[254] These laws increase government involvement in a heretofore private trend: the use of sonogram technology by anti-abortion groups,[255] which have found that women are less likely to choose abortion if they see a sonographic image of their fetus.[256]

II. THE FETUS AND CHILD IN AMERICAN LAW AND SOCIETY

A. The Status of the Fetus in American Law

To put current "fetal protection" law in context, it is necessary to trace briefly the common law and statutory trend toward recognition of the fetus as a legal entity, although in most instances this recognition arises after a child has been born.[257] At common law the fetus was not considered a legal person, and it was only after birth that a child had legal rights. The "born alive" rule governed criminal,[258] tort,[259] and inheritance[260] law.

---

transfusion, even though doctors testified that without it she was likely to die, leaving her young son with only one parent).

253. *See, e.g.*, S.B. 76, 2005 Reg. Sess. (Ind. 2005) (amending title 28, article 34, chapter 2, section 1.1 of the Indiana Code to require that at least eighteen hours prior to abortion, a woman be told of the "availability of fetal ultrasound imaging and auscultation of fetal heart tone services to enable the pregnant woman to view the image and hear the heartbeat of the fetus and how to obtain access to those services").

254. The proposed Informed Choice Act would authorize the Department of Health and Human Services to make grants to tax-exempt "community based pregnancy help medical clinic[s]" to provide ultrasound equipment to be used to give the woman a "visual image of the fetus," information about its probable gestational age, and information on abortion and its alternatives. H.R. Con. 216, 109th Cong. (2005).

255. Since 2001, the Department of Health and Human Services has been funding crisis pregnancy centers because they support President Bush's "abstinence-only" position on premarital sex education. *See* Kashef, *supra* note 5.

256. Neela Banerjee, *Church Groups Turn to Sonogram to Turn Women from Abortions*, N.Y. TIMES, Feb. 2, 2005, at A1.

257. *See infra* notes 260-273. *But see* Mone v. Greyhound Lines, Inc., 331 N.E.2d 916, at 920 (Mass. 1975) (permitting parents to recover under a wrongful death theory due to bus company's negligence that caused the stillbirth of a viable fetus).

258. *See supra* text accompanying notes 113-121 (noting the born-alive rule precluded prosecution for homicide for causing the death of a fetus); *See also* Commonwealth v. Morris, 142 S.W.3d 654, 655-57 (Ky. 2004) (discussing the history of the rule).

259. *See, e.g.*, Sheldon R. Shapiro, Annotation, *Right to Maintain Action or to Recover Damages for Death of Unborn Child*, 84 A.L.R. 3d 411 (2004).

260. *See, e.g.*, Tucker v. Carmichael & Sons, Inc., 65 S.E.2d 909, 910 (Ga. 1951) (discussing common law rules on inheritance, permitting children to sue after birth for interests that came into being while they were in utero); *see also* UNIF. PROBATE CODE § 2-109 (amended 2005) (permitting children conceived before an individual's death but born thereafter to inherit). *See also* David E. Koropp, Note, *Setting the Standard: A Mother's Duty During the Prenatal Period*, 1989 ILL. L. REV. 493, 495, n.13.

1. Tort Liability

The first case to apply the "born alive" rule in American tort law was *Dietrich v. Inhabitants of Northampton.*[261] In *Dietrich*, a pregnant woman slipped and fell and suffered a miscarriage, with her four or five month old fetus living only for a few minutes. Mr. Justice Holmes, then sitting on the Massachusetts Supreme Judicial Court, observed that "no case . . . has ever decided that, if the infant survived, it could maintain an action for injuries received by it while in its mother's womb." In his view, this was wholly appropriate, since "*the unborn child was a part of the mother at the time of the injury.*"[262]

Dietrich was followed for seventy-five years, until the 1946 decision in *Bonbrest v. Kotz.*[263] In *Bonbrest,* the United States District Court for the District of Columbia held that a viable fetus, injured through medical malpractice, had a cause of action against the physician who negligently delivered him. The decision spawned a rapid retreat from the born alive rule, accomplished by both statutory and case law. Today, every state allows a suit for prenatal injuries if the infant is born alive,[264] and most states permit a wrongful death suit to be brought on behalf of a viable fetus who succumbs prior to birth due to prenatal injury.[265] A minority of states also permits suit on behalf of a non-viable fetus.[266] Other states have recognized causes of action for loss of consortium on the part of parents whose fetus has been killed due to the tortious acts by others.[267]

---

261. 138 Mass. 14 (1884).

262. *Dietrich*, 138 Mass. at 17 (emphasis added).

263. 65 F. Supp. 138 (D.D.C. 1946).

264. Farley v. Sartin, 466 S.E.2d 522, 528 (W.Va. 1995).

265. *See* Michael P. Penick, *Wrongful Death of a Fetus*, 19 AM. JUR. 3D *Proof of Facts* 107 (2004); Jill D. Washburn Helbling, Symposium, *To Recover or Not to Recover: A State by State Survey of Fetal Wrongful Death Law*, 99 W. VA. L. REV. 363 (1996). *See also* Meyer v. Burger King Corp., 26 P.3d 925, 928-30 (Wash. 2001) (holding that the Washington worker's compensation statute did not bar a suit brought by a child allegedly deprived of oxygen in utero due to his mother's employer negligence). Indeed, employers' fear of tort liability for causing harm to the fetuses of their female employees is a major rationale of fetal protection policies in the workplace, which exclude some women from high-paying but hazardous positions. *See* Elaine Draper, *Reproductive Hazards and Fetal Exclusion Policies after* Johnson Controls, 12 STAN. L. & POL'Y REV. 117, 118, 121 (2001). For a fuller discussion of the gendered nature of the construction of workplace risks, *see id.* and discussion *infra* at text accompanying notes 323-344. Nine states (California, Florida, Iowa, Maine, Nebraska, New Jersey, New York, Texas, and Virginia) require that a child be born alive before a suit for prenatal injuries can be brought. *See* Penick, *supra*; Helbling, *supra.*

266. *See, e.g.*, Wiersma v. Maple Leaf Farms, 543 N.W.2d. 787 (S.D. 1996) (holding that the South Dakota wrongful death statute should be interpreted to authorize a cause of action for wrongful death for non-viable fetuses, and surveying the law in other jurisdictions in the process); Connor v. Monkem Co., Inc., 898 S.W.2d 89, 91-93 (Mo. 1995) (construing Missouri's wrongful death statute to permit a cause of action for a non-viable fetus in light of a Missouri statute of general applicability that declares that "[t]he life of each human being begins at conception....").

267. *See, e.g.*, Broadnax v. Gonzales, 809 N.E.2d 645 (N.Y. 2004) (permitting a woman to recover for emotional injury for a miscarriage or stillbirth due to medical malpractice, even if she herself does not suffer physical injury); Dunn v. Rose Way, Inc., 333 N.W.2d 830, 832 (Iowa 1983) (holding that a parent could recover for loss of consortium even though Iowa does not recognize a cause of action for wrongful fetal death, because a loss of consortium action is based on parental,

Only a handful of American courts have addressed the question of whether tort liability can be imposed on pregnant women based on their conduct during pregnancy. Three courts have upheld such liability, while three courts have rejected it. These cases are important in addressing the normative question of who *is* the reasonable pregnant woman, as well as the related question of who gets to evaluate her conduct.

Grodin v. Grodin[268] was the first case to permit a woman to be sued for her actions while pregnant.[269] The Michigan Court of Appeals allowed the father of a child born with discolored teeth to sue the child's mother (and his wife) for these injuries, allegedly caused by the woman's taking Tetracycline® while pregnant. Without analyzing the consequences of its decision for pregnant women, the court framed the question as a simple one of fact: did the woman's use of Tetracycline® constitute a "reasonable exercise of parental discretion?" If it did, this conduct would fall within an exception to the general abrogation of parental-child tort immunity under Michigan law, and the woman could not be sued.[270]

Two other cases, *Bonte v. Bonte*,[271] and *National Casualty Co. v. Northern Trust Bank of Florida, N.A.*,[272] also permitted suit to be brought on behalf of children who were injured due to their mothers' alleged negligence while pregnant. In *Bonte,* the mother was struck by a car while crossing the street and her child was born with severe brain damage and cerebral palsy.[273] The New Hampshire Supreme Court held that a suit could go forward, relying on the abrogation of parent-child tort immunity (in part in recognition of the availability of insurance as a source of recovery) and the law that a child born alive can bring a cause of action for injuries suffered in utero against a third party.[274] The court rejected the argument that either the unique relationship between a pregnant woman and her fetus or the potential deprivation of a woman's right to control her life during pregnancy should preclude liability.[275] The court held that a pregnant woman was "required to act with . . . *the same standard of care as that required of her once the child is born.*"[276]

---

rather than fetal, loss). Loss of consortium has been recognized as a cause of action at least since the time of Hammurabi, although his code explicitly calculated damages based on the social class of the pregnant woman. The Code of Hammurabi, §§ 209 and 213 (L.W.King trans.), available at http://www.leb.net/~farras/history/hammurabi.htm (declaring that a free born woman was entitled to receive ten shekels for her loss while a maid-servant was entitled to receive two shekels).

268. 301 N.W.2d 869 (Mich. Ct. App. 1981).

269. *Grodin*, 310 N.W.2d 869.

270. The court remanded the matter to the trial court to determine the "'reasonableness' of the [mother's] alleged negligent conduct." *Id.* at 871.

271. 616 A.2d 464 (N.H. 1992).

272. 807 So. 2d 86 (Fla. Dist. Ct. App. 2001).

273. *Bonte*, 616 A.2d at 464.

274. *Id.* at 465-66.

275. *Id.* at 466.

276. *Id.*

A similar result was reached by a Florida appellate court in *National Casualty Co.*, which ruled that a child could sue its mother for injuries allegedly caused by her negligent driving while pregnant, up to the amount of her automobile insurance coverage.[277] The court's brief opinion held that there was no reason to "den[y] . . . recovery merely because of the identity of the tortfeasor,"[278] rejecting concerns about the impact of its decision on the mother's privacy and personal health, and distinguishing *State v. Ashley*, in which the Florida Supreme Court held a pregnant woman who shot herself and caused the death of her fetus could not be charged criminally.[279]

In contrast, three courts have adamantly rebuffed suits brought by children against their mothers for injuries suffered in utero. In *Stallman v. Youngquist*,[280] the Illinois Supreme Court held that a child who suffered prenatal injuries in a car accident in which her mother was driving could not sue her mother for negligence.[281] The court first criticized the *Grodin* decision, suggesting that the Michigan court had confused the question of whether parental tort immunity should be abrogated with the different issue of whether a pregnant woman owed a tort duty to her fetus.[282] The *Stallman* court confronted the latter issue directly. It emphasized that the relationship between a pregnant woman and the fetus she was carrying was unique and "unlike the relationship between any other plaintiff and defendant,"[283] and thus could not be analogized to other negligence situations. The Illinois Supreme Court held that in view of the "fact of life" that a pregnant woman's "every waking and sleeping moment . . . shapes the prenatal environment which forms the world for the developing fetus,"[284] it was impermissible to impose a duty of care on a pregnant woman.[285]

The court asserted four grounds for its decision. First, it would be impossible to either limit or define the duty of a pregnant woman toward her fetus, since many actions taken in a woman's life, even prior to conception, could affect a fetus.[286] Second, it would be impossible to develop an objective standard applicable to women from diverse socioeconomic backgrounds, whose access to health care differed, and who might or might not know whether they were pregnant.[287] Third, the court recognized that creating a common law cause of action had the potential

277. *National Casualty*, 807 So. 2d 86.
278. *Id.* at 87.
279. *Id.* at 87-88 (citing State v. Ashley, 701 So. 2d 338 (Fla. 1997)).
280. 531 N.E.2d 355 (Ill. 1988).
281. *Stallman*, 531 N.E.2d at 361.
282. *Id.* at 358.
283. *Id.* at 360.
284. *Id.*
285. *Id.* at 359-61.
286. *Id.* at 360.
287. *Id.*

for "unprecedented intrusion into the privacy and autonomy of the [female] citizens of this State."[288] It held that if a duty was to be recognized, it must be by the legislature, "only after thorough investigation, study, and debate."[289] Finally, the court urged that "[t]he way to effectuate the birth of healthy babies is not . . . through after-the-fact civil liability in tort for individual mothers, but rather through before-the-fact education of all women and families about prenatal development."[290]

Chenault v. Huie[291] addressed the more difficult factual circumstances in which Huie, a pregnant woman (and her boyfriend) abused alcohol and other drugs while she was pregnant, and she gave birth to a child with developmental problems and cerebral palsy.[292] Huie's sister sued on behalf of the child, seeking compensatory and punitive damages for Huie's alleged negligence and gross negligence.[293] The Texas Court of Appeals declined to recognize a child's common law cause of action against its mother for prenatally-caused injuries. The court declared that while "the law wisely no longer treats a fetus as only a part of the mother, the law would ignore the equally important physical realities of pregnancy if it treated the fetus as an individual entirely separate from his mother."[294] The court pointed to the difficulty of establishing an objective, uniform standard of care for pregnant and potentially pregnant women, noting the inevitable subjectivity of after-the-fact jury decision-making, which would lead to inconsistent and unpredictable jury verdicts, as well as the invasion of women's autonomy and right to control their daily lives.[295] The court declared that recognizing a duty of women toward their fetuses was the province of the legislature, which alone could conduct the necessary "research and analysis of scientific and medical data . . . [and] evaluat[e] . . . broad matters of public policy."[296] Finally, the court expressed the concern that imposing civil liability might be counterproductive, because women who feared civil liability might not be candid with their physicians, and thus would receive less than adequate prenatal care.[297]

In 2004, in *Remy v. MacDonald*,[298] the Massachusetts Supreme Judicial Court declined to permit a child to sue its mother for prenatal harm.[299] In *Remy*, the plaintiff alleged that her mother drove negligently while pregnant, causing the plaintiff to be born prematurely with adverse

288. *Id.* at 361.
289. *Id.*
290. *Id.*
291. 989 S.W.2d 474 (Tex. Ct. App. 1999).
292. *Chenault*, 989 S.W.2d at 475.
293. *Id.*
294. *Id.* at 475-76.
295. *Id.* at 477-78.
296. *Id.* at 478.
297. *Id.*
298. 801 N.E.2d 260 (Mass. 2004).
299. *Remy*, 801 N.E.2d at 266.

health consequences. The court followed *Stallman* and *Chenault*, emphasizing the substantial disagreement about whether pregnant women should be held liable for causing fetal harm, observing that there were virtually unlimited circumstances in which a woman could be sued, and declaring that there was no principled way to limit the liability of pregnant women for causing fetal harm to the motor vehicle context.[300] The court explicitly rejected the reasoning of *Grodin*, *Bonte*, and *National Casualty Co.*, and found that courts should recognize "that there are inherent and important differences between a fetus, in utero, and a child already born, that permits [sic] a bright line to be drawn around the zone of potential tort liability of one who is still biologically joined to an injured plaintiff."[301]

The approach of *Remy*, *Stallman*, and *Chenault* is similar to that set forth by the Supreme Court of Canada, whose reasoning is instructive. The Court has held consistently over the last fifteen years that a pregnant woman and her fetus share a unique relationship, in which there is only one legal person, rather than two persons with potentially adverse positions.[302] In *Winnipeg Child & Family Services (Northwest Area) v. G. (D.F.)*, the Court noted, "[T]he law has always treated the mother and unborn child as one. To sue a pregnant woman on behalf of her unborn fetus therefore posits the anomaly of one part of a legal and physical entity suing itself."[303] In *Dobson v. Dobson*, the Court explained that there was no principled way to limit the circumstances under which the woman could be held liable, due to the extraordinarily close physical proximity between the woman and her fetus, and the enormous range of actions which the woman could take which could have a detrimental effect on fetal development.[304] "Everything the pregnant woman eats or drinks, and every physical action she takes, may affect the foetus."[305] The Court further identified two important public policy concerns "militat[ing] against the imposition of maternal tort liability for prenatal negligence[:] . . . the privacy and autonomy rights of women and . . . the difficulties inherent in articulating a judicial standard of conduct for pregnant women."[306] The Court emphasized that simply because a woman is pregnant, she does not lose "the right to make personal decisions, to control [her] . . . bodily integrity, and to refuse unwanted medi-

300. *Id.* at 264-66.

301. *Id.* at 267.

302. Winnipeg Child & Family Services (Northwest Area) v. G. (D.F.), [1997] 3 S.C.R. 925 ¶¶ 27-29 (holding that a pregnant woman addicted to solvents could not be civilly committed for treatment) and Dobson v. Dobson, [1999] 174 D.L.R. (4th) ¶ 1 (rejecting tort liability for a pregnant woman who allegedly drove negligently, causing injury to her fetus and declaring that, "[t]he relationship between a pregnant woman and her foetus is unique and innately recognized as one of great and special importance to society").

303. *Winnipeg Child & Family Services*, 3 S.C.R. at ¶ 27.

304. *See Dobson* 174 D.L.R.. at ¶ 20.

305. *Id.* at ¶ 27.

306. *Id.* at ¶ 21.

cal treatment."[307] The Court finally noted the difficulty in developing a workable judicial standard of conduct for pregnant women, stating it would be impossible to articulate an objective standard because the context of every pregnant woman's life is different, with women who are well-educated and ignorant, rich and poor, with and without access to good health care and good prenatal care.[308] The court observed that "the reasonable pregnant woman" standard would inevitably be interpreted in light of the trier of fact's prejudices about the proper conduct of pregnant women.[309]

2. Child Neglect and Proceedings to Terminate Parental Rights

In contrast to the split over whether pregnant women should be liable for prenatal torts, all states agree that a woman's use of alcohol or other drugs while pregnant is a proper trigger for taking custody of a child as "neglected," and may be the basis for terminating her parental rights.[310] The jurisprudence in this area can be succinctly summarized: "Courts rarely side with drug abusing parents where children are concerned."[311] Courts do differ, however, over the jurisdictional question of whether a fetus may be considered a child and the further question of whether a woman's substance abuse during pregnancy is in itself sufficient to justify the loss of parental rights.[312]

---

307. *Id.* at ¶ 32 (citing the ROYAL COMMISSION ON NEW REPRODUCTIVE TECHNOLOGIES, PROCEED WITH CARE: FINAL REPORT ON NEW REPRODUCTIVE TECHNOLOGIES 955-56 (1993)).

308. *Id.* at ¶ 54.

309. *Id.* at ¶ 53.

310. Some states statutes explicitly authorize courts to consider prenatal substance abuse. *See, e.g.*, COLO. REV. STAT. ANN §§ 19-3-102(1)(g) (West 2005) (declaring that a child is neglected or dependent if it is born with controlled substances in its system); OKLA. STAT. ANN. tit. 10, § 7001-1.3(14)(c) (West 2006) (declaring that a child born dependent on controlled substance is a "deprived child"). Other states have achieved the same result through judicial interpretation of more general child neglect criteria. *See, e.g., In re* Troy D., 263 Cal. Rptr. 869 (Cal. Ct. App. 1989) (applying CAL. WELF. & INST. CODE § 300 (a) (West 2006) to a child born to a mother who ingested drugs during pregnancy); *In re* Baby Boy Blackshear, 736 N.E.2d 462 (Ohio 2000) (holding that a newborn with a positive toxicology screen is *per se* an abused child under the Ohio civil child abuse statute); *see also In re* Stefanel Tyesha C., 556 N.Y.S.2d 280 (N.Y. App. Div. 1990) (quoting N.Y. FAM. CT. ACT § 1046 (a) (iii) (1981)) (holding that allegations that a mother admitted drug use while pregnant and that her infant had a positive toxicology test are sufficient to permit a child neglect proceeding to go forward).

311. Mary E. Taylor, Annotation, *Parent's Use of Drugs as Factor in Award of Custody of Children, Visitation Rights, or Termination of Parental Rights*, 20 A.L.R. 5th 534 , §2 [b] (2005).

312. In *In re Valerie D.*, the Supreme Court of Connecticut held that a woman's use of drugs while pregnant could not, standing alone, justify the termination of her parental rights, and also found that by taking the child away immediately after birth, the state made it impossible for the mother to establish that she had an appropriate and ongoing relationship with her child. 613 A.2d 748, 752-53 (Conn. 1992). *See also In re* Appeal in Pima County Juvenile Severance Action No. S-120171, 905 P.2d 555, 558 (Ariz. Ct. App. 1995) (holding that "while chronic use of drugs or alcohol by either parent during the mother's pregnancy may reflect a pattern of substance abuse and may be so telling of the kind of environment to which the child will be born as to justify the child's immediate removal from the parents at birth, chronic substance abuse during pregnancy in and of itself does not reflect an inability to parent that would justify severance of a parent's fundamental rights"); *see also supra* text accompanying notes 201-13 (regarding the jurisdiction of juvenile and family courts over fetuses, as opposed to children). Other courts have come to a different conclusion, *cf. In*

As always in child abuse and neglect cases, there are competing goals and strategies, which are magnified in the case of parents[313] who use alcohol and other drugs. The law's basic premise is that it is the child's, rather than the parent's, best interests, which are paramount. This leads to early intervention by child welfare authorities to protect a child at risk,[314] and means in practice that many courts have ruled that children born with positive toxicology screens for illegal drugs and/or other evidence of prenatal drug exposure may be temporarily removed from their parents' custody.[315] On the other hand, parental advocates may urge a watchful waiting period and argue that more family support services and drug treatment should be provided, in order to make it more likely that children may remain with their parents, whose interest in the enjoyment of a relationship with their children is constitutionally protected.[316] The Adoption and Safe Families Act of 1997 (AFSA)[317] gives states financial incentives to expeditiously terminate parental rights, with the laudable goal of not leaving children in foster care limbo while their parents struggle to get their lives in order.[318] Since the law translates "the best interests of the child" as "less time spent in foster care," the result has been faster termination of parental rights.[319] Many commentators have asserted that "the 12-month permanency clock for children ignores the clock of treatment for addiction, which is at best 24 months," and that the statute operates in a draconian and counter-productive manner in the

---

re Baby Blackshear, 736 N.E.2d 462, 464 (Ohio 2000), but in most cases there are factors in addition to the mother's prenatal drug use that are cited in support of the decision to terminate parental rights. *See, e.g., In re* W.A.B., 979 S.W.2d 804, 808 (Tex. Ct. App. 1999).

313. It is important to remember that there are often two parents involved, and frequently that the father is also a drug user or the one whose physical and emotional abuse exacerbates the mother's vulnerability to drug use. *See* Francisco G. v. Superior Court, 110 Cal. Rptr. 2d 679 (Cal. Ct. App. 2001) (upholding the termination of a father's parental rights where his alcohol abuse and domestic violence had not been addressed and he supported the mother's assertion that she did not have a drug problem when there was overwhelming evidence to the contrary).

314. *See, e.g., In re* Stefanel Tyesha C., 556 N.Y.S.2d at 284 ("[A] court cannot and should not 'await broken bone or shattered psyche before extending its protective cloak around [a] child pursuant to . . . article 10 of the Family Court Act . . . .'" (citations omitted)). Courts have justified their decision by citing the child's "right to begin life with a sound mind and body." *Stallman*, 531 N.E.2d at 358 (citations omitted).

315. *See, e.g., In re* Troy D., 263 Cal. Rptr. at 877 (holding that there was substantial evidence to support the juvenile court's finding that a newborn was "in need of proper and effective parental care or control" in light of the infant's positive toxicology results for opiates and amphetamines, post-birth behavior which suggested prenatal drug exposure, and the mother's continued drug use after losing custody of another child due to her drug use).

316. *See* Stanley v. Illinois, 405 U.S. 645 (1972); Santosky v. Kramer, 455 U.S. 745 (1982); *In re* Guardianship of K.H.O., 736 A.2d 1246, 1251 (N.J. 1999).

317. Pub. L. No. 105-89, 111 Stat. 2115 (1997) (codified as amended in scattered sections in 42 U.S.C.A.).

318. 42 U.S.C.A § 671(a)(15) (West 2006).

319. *See* Laureen D'Ambra, *Terminating the Parental Rights of the Mentally Disabled*, 49 R. I. BAR. J. 5, 7 (2001) (discussing the law's shorter time frames for children to remain in foster care, with the goal of promoting permanent placement for children); Dorothy Roberts, *The Challenge of Substance Abuse for Family Preservation Policy*, 3 J. HEALTH CARE L. & POL'Y 72, 73-76 (1999).

case of drug-abusing women. [320] The AFSA's push for a quick decision about termination of parental rights is also problematic in view of most states' inadequate drug treatment resources.[321] Many drug treatment programs are simply not tailored to mothers, who need both child care and a treatment philosophy different from the "confrontational" and individually focused style typical of drug treatment programs designed with the male drug addict in mind.[322]

B. The Failure to Protect Fetuses and Children from Work-Related Harms

In contrast to the emphasis of prosecutors, child abuse agencies, and some tort plaintiffs on the mother's body and behavior as the locus of both fetal harm and protection, government and private actors have largely been silent about systemic deficits in the American workplace that place parents and their offspring at risk. Two aspects of workplace life bear special scrutiny: the dangers to fetal and childhood development posed by male and female workers' exposure to toxic substances, and the lack of economically viable parental leave policies, which prevent many parents from adequately addressing the health and emotional needs of their newborns and young children.

1. Workplace Exposure to Hazardous Substances

Employers have responded to the risk that workplace exposure to toxins and other dangerous substances will injure future children by en-

---

320. *See* Roberts, *supra* note 319, at 77, 86 (quoting Cornelia Grunman, *Parents Give Advice on Reforming DCFS; Agency Criticized at Panel Hearings,* CHI. TRIB., Apr. 13, 1999, at Metro Chicago 3).

321. Federal law requires 5% of its Substance Abuse Prevention and Treatment (SAPT) Block Grant program to improve drug treatment access for pregnant women, but states may seek a waiver of this requirement if they can show that there are no access problems. *See* 45 C.F.R. § 96.124 (1993)(c)-(d). Only 14% of the drug treatment facilities in the United States have program specifically designed to treat pregnant and postpartum women. OFFICE of APPLIED STUDIES, SUBSTANCE ABUSE AND MENTAL HEALTH SERVICES ADMIN., DEP'T OF HEALTH AND HUMAN SERVICES, NATIONAL SURVEY OF SUBSTANCE ABUSE TREATMENT SERVICES (N-SSATS): 2003, DATA ON SUBSTANCE ABUSE TREATMENT FACILITIES 38 (Sept. 2004), http://wwwdasis.samhsa.gov/04nssats/nssats_rpt_04.pdf. Compounding the lack of adequate treatment facilities for pregnant women, only 35% of the facilities had programs for persons needing treatment for both substance abuse and mental illness. *Id.* at 37. Such individuals are especially likely to fall through the cracks of government sponsored program, as legislation is usually targeted at one, but not the other, of these illnesses. Georganne Chapin, *Sanctioning Substance-Abusing Home Relief Clients with the Loss of Medical Benefits—Legal and Policy Concerns,* 7 N.Y.S. BAR ASS'N HEALTH L. J. 35, 39 (Spring 2002).

322. *See* Roberts, *supra* note 319, at 78; Mary O'Flynn, *The Adoption and Safe Families Act of 1997: Changing Child Welfare Policy Without Addressing Parental Substance Abuse,* 16 J. CONTEMP. HEALTH L. & POL'Y 243, 262 (1999); Kathryn T. Jones, *Prenatal Substance Abuse: Oregon's Progressive Approach to Treatment and Child Protection Can Support Children, Women and Families,* 35 WILLAMETTE L. REV. 797 (1999); Holly A. Hills, Deborah Rugs & M. Scott Young, *Justice, Ethics, and Interdisciplinary Teaching and Practice: The Impact of Substance Abuse Disorders on Women Involved in Dependency Court,* 14 WASH. U. J.L. & POL'Y 359 (2004); Jane C. Murphy & Margaret J. Potthast, *Domestic Violence, Substance Abuse and Child Welfare: The Legal System's Response,* 3 J. HEALTH CARE L. & POL'Y 88 (1999).

acting "fetal protection" policies limiting the exposure of women of child-bearing age to such substances.[323] The Supreme Court ruled in 1991 in *International Union, UAW v. Johnson Controls*[324] that a fetal protection policy which excludes women from the workplace as a means of reducing the risk that a child will be harmed by prenatal toxic exposure constituted sex discrimination in violation of Title VII of the Civil Rights Act of 1964.[325] The employer's actions will be illegal unless the employer can demonstrate that its policy is a bona fide occupational qualification, that is, "reasonably necessary to the normal operation of the particular business."[326] However, many employers still implement "fetal protection" policies, admitting in effect that they would rather be sued for sex discrimination than for damages for causing prenatal injury.[327] These employers continue to focus exclusively on the risks of harm to future children posed via the female body (including harm to the woman's reproductive system and to the fetus) rather than acknowledge the harm that many substances pose to the male reproductive system. However, when studies show that a substance poses harm to the male reproductive system, industry and government frequently have acted to ban the substance entirely.[328]

2. Parenting Leaves

American law has also failed to mandate paid parenting leaves, which would permit parents to take care of newborns, as well as older children who become ill. Although the Family and Medical Leave Act (FMLA)[329] requires employers of more than fifty employees to grant

323. Although the Occupational Safety and Health Act, 29 U.S.C.A. §§ 651–678 (West 2006), and the Toxic Substances Control Act, 15 U.S.C.A. §§ 2601–2692 (West 2006), both mandate that the government set safe levels of workplace exposure to many dangerous substances, both acts have been weakened by court rulings about the level of scientific proof required for the government to demonstrate "significant risk." *See, e.g.*, Industrial Union Dep't., AFL-CIO v. American Petroleum Institute, 448 U.S. 607, 642 (1980) (plurality opinion) (holding that before promulgating any occupational safety and health standard, the Secretary of Labor must find that there are "significant risks" in a workplace which can be eliminated or decreased by a change in the standard).

324. 499 U.S. 187 (1991).

325. *Johnson Controls*, 499 U.S. at 207.

326. *Id.* at 200 (quoting 42 U.S.C.A. § 2000(e)(1)). In *Johnson Controls*, the defendant company manufactured batteries, in which lead was a major ingredient. Although the company initially permitted pregnant women to work in the manufacturing process, informing them of the dangers of lead exposure, after eight women tested with higher blood lead levels than recommended by OSHA, the company issued a "fetal protection" policy. This policy excluded "'all women except those whose inability to bear children is medically documented,'" but it made no provision for men to lower their lead exposure, which has also been shown to pose a risk of fetal harm. *Id.* at 190-92.

327. Draper, *supra* note 265, at 121.

328. The pesticide ethylene dibromide (EDB), for example, was cancelled because of its oncogenic and mutagenic risks, as well as reproductive risks to male workers. *See* EPA Limitation on Ethylene Dibromide, 48 Fed. Reg. 46,234 (Oct. 11, 1983); *see also* OSHA Limitation on Levels of Occupational Exposure to Ethylene Dibromide, 48 Fed. Reg. 45,956 (Department of Labor document regarding EDB's effects on male reproductive capacity); *cf. Johnson Controls*, 499 U.S. 187 (1991).

329. Pub. L. No. 103-3, 107 Stat. 6 (1993) (codified in scattered sections of 2 U.S.C.A, 5 U.S.C.A., 29 U.S.C.A.).

employees up to twelve weeks a year of unpaid leave for the birth or adoption of a child or for reasons related to illness,[330] many critics assert that this is insufficient to support vulnerable children in need of parental attention and that the unpaid character of such parental leave means that existing race and class hierarchies are not remedied.[331] The United States' policy stands in marked contrast to those of other developed nations, which give much more generous leaves to working parents.[332] Almost half of American workers are not covered by the FMLA,[333] and even among those who are, only a fraction avail themselves of its leave provisions, because they cannot afford not to work.[334] No federal law mandates paid parental leave for the period connected with pregnancy, childbirth, and the early stages of infancy,[335] and California and Ohio are the only two states to mandate any form of paid parental leave.[336] In contrast, other developed nations either mandate or offer paid parenting leave, at least for some portion of this critical stage of fetal and childhood health and development,[337] and many countries offer additional financial or child-care support to single parents, those who are most likely to need leave from work to care for a newborn or ill child and are simultaneously the least likely to be able to afford to do so.[338]

America's failure to provide paid leave for child-bearing and parenting is both physically and socially harmful to children, as well as economically short-sighted. Many studies indicate that breast-feeding provides important health benefits to newborns,[339] and it is certainly much

330. *See* 29 U.S.C.A. § 2611(4)(a)(i) (West 2006) (defining "Employer" as any person employing more than 50 people); 29 U.S.C.A. § 2612(a)(1)(B), (C) (entitling an employee to 12 weeks leave for the birth or adoption of a child).

331. Nancy E. Dowd, *The Family and Medical Leave Act of 1993: Ten Years of Experience: Race, Gender, and Work/Family Policy*, 15 WASH. U. J.L. & POL'Y 219, 222-31 (2004) [hereinafter *Ten Years of Experience*].

332. *Id.* at 231-36.

333. Erin Gielow, Note, *Equality in the Workplace: Why Family Leave Does Not Work*, 75 S. CAL. L. REV. 1529, 1539 (2002)

334. *Id.* at 1546.

335. In April 2006, Representative Caroline Maloney introduced HR 5148, the Federal Employees Paid Parental Leave Act. The Act would ensure paid leave for 6 of the 12 weeks that federal employees are authorized to take parental leave. H.R. 5148, 109th Cong. (2006) (referred to H. Comm. on Gov't Reform), *available at* http://thomas.loc.gov/cgi-bin/query/z?c109:H.R.5148:.

336. CAL. UNEMP. INS. CODE § 3300 (West 2006); OHIO REV. CODE ANN. § 124.136 (West 2006) (providing that permanent government employees shall receive 70% of their salary for four of the six weeks in which they are authorized to take parental leave).

337. Dowd, *supra* note 331 at 233-36 (summarizing European Union law, and comparing, *inter alia*, the approach of France, which mandates maternity leave and provides much more generous paid leaves to mothers than to fathers, and Sweden, which is gender-neutral in its paid parenting leave policies); *see also* Naomi S. Stern, *The Challenges of Parental Leave Reforms for French and American Women: A Call for a Revived Feminist-Social Theory*, 28 VT. L. REV. 321, 324-25 (2004) (describing the French statutory scheme).

338. Gielow, *supra* note 333, at 1547.

339. American Academy of Pediatrics, *Policy Statement: Breast Feeding and the Use of Human Milk*, 115 PEDIATRICS 496, 496-97 (2005); *see also* Shana M. Christrup, *Breastfeeding in the American Workplace*, 9 AM. U.J. GENDER SOC. POL'Y & L. 471, 474-76 (2001); Judy Heymann, Editorial, *We Can Afford to Give Parents a Break*, WASH. POST, May 14, 2006, at B07.

easier to breastfeed when not working full-time.[340] Both newborns and older children whose health needs are not met promptly are likely to be sicker for longer periods of time, adding to state and federal health care expenditures.[341] In addition, parents who are distracted by ill children left at home may be less productive workers.[342] Studies of parental leave practices in other countries show that parents who are given generous paid leaves to care for their children rarely abuse it,[343] and that children, parents, and employers benefit when the law provides a framework for parents to care for their children's health without jeopardizing the familial economy.[344]

C. Environmental Exposures

Children and fetuses are exposed to an astounding number of toxins, teratogens, mutagens, and carcinogens in the environment.[345] While the full effects of these exposures are not yet known,[346] and a comprehensive discussion of environmental hazards is beyond the scope of this paper, one recent example of government actions putting children at risk is instructive. In March 2005 the Environmental Protection Agency (EPA) promulgated the so-called Clean Air Mercury Rule,[347] which reversed a 2000 EPA rule and substantially expanded the ability of American power plants to continue to emit mercury and other toxic air pollutants. The EPA announced that coal- and oil-burning power plants were not subject to the requirements of §112 of the Clean Air Act, which would have required them to install new equipment to reduce mercury emissions. Instead, the new rule created a complicated system of "cap-and-trade" pol-

---

340. American Academy of Pediatrics, *supra* note 339 at 498; Christrup, *supra* note 339, at 480-81.

341. *Cf.* KAISER, IN A TIME OF GROWING NEED, *supra* note 160, at 12 (citing a study in Washington State showing that children who lack health care insurance were less likely than those who do to have had a clinic or physician visit within six months and are twice as likely to have received care at an emergency room during the same period).

342. *See* Gielow, *supra* note 333, at 1541-43.

343. DECKER, *supra* note 64, at 10 ("Contrary to the widely held belief that employees would abuse a liberal leave policy, the average usage rate of [Sweden's generous leave policy, which is 90 days per year to care for a sick child] . . . is seven days a year.").

344. *See* Gielow, *supra* note 333, at 1540-42.

345. In recognition of the risks posed by these exposures, in 1997 President Clinton formed a Task Force on Environmental Health and Safety Risks to Children, to improve inter-agency cooperation within the federal government to provide greater protections for children's health. Exec. Order No. 13,045, 62 Fed. Reg. 19,885 (April 23, 1997).

346. Robert L. Brent & Michael Weitzman, *The Current State of Knowledge About the Effects, Risks, and Science of Children's Environmental Exposures*, 113 PEDIATRICS 1158, 1159, 1164-65 (2004), *available at* http://pediatrics.aappublications.org/cgi/reprint/113/4/S1/1158 (noting the wide range of conclusions reached about the risk that pesticides and other environmental toxicants pose to children, urging that more, carefully controlled, studies be undertaken to evaluate issues such as the susceptibility of children compared with adults, the impact of exposure to multiple environmental hazards, and observing that where there is sufficient data to act, as in the case of lead, there is no reason to delay).

347. Mark D. Sullivan & Christine A. Fazio, *The EPA's New Clean Air Rules: Mixed Results for Air Quality*, N.Y. ST. B.J. 11, 15 (Jan. 2006); 70 Fed. Reg. 15,994 (Mar. 29, 2005) (to be codified at 40 C.F.R. pt. 63).

lution allowances which permitted the power plants to increase their mercury pollution by purchasing emissions allowances from other facilities.[348] Since half of all Americans live within thirty miles of a coal-burning power plant, children and their families living near a power plant were likely to be exposed to substantially more mercury emissions than they are now.[349] Eleven states filed suit to invalidate the new Rule.[350]

Congressional critics also sought to overturn the Rule, citing the significant risks that mercury poses to fetuses and children. They noted that mercury was a recognized neurotoxin, which causes devastating effects on fetuses and young children because it interferes with normal brain development. They further noted that 4.9 million women of childbearing age have elevated levels of mercury, 630,000 infants are born with elevated mercury levels, and 1,500 children are born each year with mental retardation due to in utero exposure to mercury.[351] Ultimately, this Congressional effort was defeated. Although EPA agreed to reconsider the rule,[352] in May 2006 the agency reaffirmed its original position.[353]

D. Lack of Health Care Access in the United States and Its Consequences

1. Uninsured Children and Adults

Americans continue to lack health insurance coverage in record numbers. In 2005, 51.4 million Americans (almost 18%) were uninsured for at least some part of the previous year, and 29.3 million (more than 10%) had been uninsured for more than a year.[354] Although children were more likely than adults to have health insurance, 9.2 million chil-

348. Sullivan & Fazio, *supra* note 347, at 15-16; Kim McGuire, *New Mercury Rules Decried, Environmental Coalition Plans Lawsuits over EPA Changes, Activists Say the New Limits Will Allow More Mercury Pollution from Coal-Burning Power Plants-Which They Say Violates the Clean Air Act*, DENVER POST, May 18, 2005, at B-02.

349. Mark Clayton, *In Bid to Cut Mercury, US Lets Other Toxins Through*, CHRISTIAN SCIENCE MONITOR, Mar. 31, 2005, at 13, *available at* http://www.csmonitor.com/2005/0331/p13s01-sten.html (noting also that coal-fired power plants also emit many other serious pollutants, including "vanadium, barium, zinc, nickel, hydrogen fluoride, hydrochloric acid, ammonia, and selenium").

350. Don Hopey, *11 States Sue EPA on Mercury Rules: Pennsylvania Joins suit Saying Emission Standards Inadequate*, PITTSBURGH POST-GAZETTE, May 19, 2005, *available at* http://www.post-gazette.com/pg/05139/507051.stm.

351. Senate Debate on Disapproval of EPA Rule Promulgation, 151 CONG. REC. S9912, S9913-14 (2005).

352. Sullivan & Fazio, *supra* note 347, at 16.

353. EPA, FINAL RULE RECONSIDERING TWO MERCURY ACTIONS: (1) RECONSIDERATION OF RULE REVISING EARLIER REGULATORY FINDING AND REMOVING CERTAIN ELECTRIC STEAM GENERATING UNITS FROM THE LIST OF SOURCE CATEGORIES; AND (2) RECONSIDERATION OF THE CLEAN AIR MERCURY RULE FACT SHEET 1, 2 (May 31, 2006), http://www.epa.gov/air/mercuryrule/pdfs/fs20060531.pdf.

354. ROBIN A. COHEN & MICHAEL E. MARTINEZ, HEALTH INSURANCE COVERAGE: ESTIMATES FROM THE NATIONAL HEALTH INTERVIEW SURVEY, JANUARY–SEPTEMBER 2005 1 (Mar. 2006), http://www.cdc.gov/nchs/data/nhis/earlyrelease/insur200603.pdf.

dren (more than 12%) lacked health insurance for some portion of the year and an additional 3.9 million (more than 5%) had been without health care coverage for more than a year.[355] There are also substantial racial and ethnic disparities in insurance status, with Hispanics suffering the greatest access problems, although African-Americans also lagged behind their white counterparts. One-third of all Hispanics were uninsured for at least part of the previous year and one-quarter lacked health care coverage for more than a year.[356] About 13% of pregnant women lack health insurance coverage,[357] despite efforts to expand Medicaid during the last two decades.[358] Medicaid does insure a greater proportion of pregnant than non-pregnant women,[359] and pays for a third of all American births.[360]

The consequences of lack of health insurance for adults and children are profound. While prenatal care is seen as an important factor in leading to good birth-outcomes,[361] many poor and low-income American women continue to lack prenatal care,[362] and those who receive some prenatal care often receive it late in pregnancy, when it is less effective.[363] Similarly, while pediatricians and health policymakers agree that well-child visits are essential to providing necessary screening and other preventative care,[364] half of uninsured children fail to have even one well-child visit a year.[365] Children with private health insurance coverage are much more likely to have received all necessary immunizations than those who are uninsured or have government health insurance.[366]

---

355. *Id.* at 8. This percentage that has not changed since 2002. *Id.* at 2-3.

356. *Id.* at 4. *See also* URBAN INSTITUTE, HEALTH INSURANCE TRENDS (2005), *available at* http://www.urban.org/toolkit/issues/healthinsurance.cfm?renderforprint=1.

357. KENNETH E. THORPE, JENNIFER FLOME & PETER JOSKI, THE DISTRIBUTION OF HEALTH INSURANCE COVERAGE AMONG PREGNANT WOMEN, 1999 5-6 (2001), *available at* http://www.marchofdimes.com/files/2001FinalThorpeReport.pdf. Forty-seven percent of poor pregnant women lacked Medicaid coverage in 1999. *Id.* at 10.

358. ROBERT WOOD JOHNSON FOUNDATION, *supra* note 32, at 5; U.S. GEN. ACCOUNTING OFFICE, GAO/PEMD-91-10: PRENATAL CARE: EARLY SUCCESS IN ENROLLING WOMEN MADE ELIGIBLE BY MEDICAID EXPANSIONS 7-8 (Feb. 1991), *available at* http://archive.gao.gov/d21t9/143346.pdf (describing legislative changes).

359. ROBERT WOOD JOHNSON FOUNDATION, *supra* note 32.

360. NATIONAL GOVERNORS ASSOCIATION CENTER FOR BEST PRACTICES, MATERNAL AND CHILD HEALTH UPDATE 2002: STATE HEALTH COVERAGE FOR LOW-INCOME PREGNANT WOMEN, CHILDREN, AND PARENTS 1 (June 10, 2003), http://www.nga.org/Files/pdf/MCHUPDATE02.pdf.

361. ROBERT WOOD JOHNSON FOUNDATION, *supra* note 32.

362. *Id.*

363. *Id.* at 8-9 (analyzing results of Medicaid expansions of the 1980's and early 1990's).

364. *See, e.g.*, MEDICAID FACTS, *supra* note 29, at 1. The American Academy of Pediatrics recommends that children receive 10-11 well-child visits before the age of three, in order for pediatricians to monitor child behavior and development and prevent illness and behavioral and educational problems in the future. Steven Blumberg, Neal Halfon, & Lynn M. Olson, National Survey of Early Childhood Health, *National Survey of Early Childhood Health: Parents' Views on Preventive Care for Infants and Toddlers*, 113 PEDIATRICS 1899 (2004), *available at* http://pediatrics.aappublications.org/cgi/reprint/113/6/S1/1899.

365. URBAN INSTITUTE, *supra* note 356.

366. National Survey of Early Childhood Health, National Survey of Early Childhood Health: *Parents' Views on Preventive Care for Infants and Toddlers*, *supra* note 364, at 6.

2. Medicaid and SCHIP Need Improvement

Medicaid is a partnership between state and federal governments, which provides health insurance to the very poorest of American children and adults, persons with disabilities, and elderly needing long-term care.[367] All states participating in Medicaid agree to provide the same set of federally mandated services for children,[368] which are known as Early and Periodic Screening, Diagnostic, and Treatment Services (EPSDT).[369] These provide essential care for children, particularly those with disabilities and other special needs.[370] However, many children have not received the mandated EPDST benefits, either because they cannot find a physician willing to accept the low Medicaid reimbursement rates,[371] or because some states have failed to adequately define the EPSDT services in their managed care contracts.[372] Several suits have been brought by groups of parents challenging the denial of benefits,[373] but it has been an uphill struggle to ensure that children enrolled in Medicaid receive all the

---

367. For a good overview of the Medicaid program, see U.S. GEN. ACCOUNTING OFFICE, GAO-01-749, MEDICAID: STRONGER EFFORTS NEEDED TO ENSURE CHILDREN'S ACCESS TO HEALTH SCREENING SERVICES 1-8 (July 2001) [hereinafter GAO, MEDICAID REPORT].

368. States differ in the extent to which they provide covered services to the near-poor, as well as the desperately poor. For example, all states provide coverage to parents who earn no more than the Federal Poverty Level, which is $16,600 for a family of three in the Lower Forty-Eight states. Only 14 states have raised their eligibility levels about the Federal Poverty Level. NATIONAL WOMEN'S LAW CENTER, POOR PARENTS ON MEDICAID TARGETED FOR CUTS 1-2 & 3 n.8 (Feb. 2006), *available at* http://www.nwlc.org/pdf/FSPoorParentsTargeted_06.06.pdf.

369. Some state variation is permitted in what services are covered under the rubric of "family planning," which includes birth control, treatment of sexually transmitted diseases, and sterilization, NATIONAL WOMEN'S LAW CENTER, MEDICAID, BIRTH CONTROL AND WOMEN'S *HEALTH* 1 & 2 n.2 (Mar. 10, 2006), *available at* http://www.nwlc.org/pdf/FSMedicaidBirthControlandWomens Health_05.31.06.pdf, thus risking reversal of recent gains in health outcomes for children and their parents, including the decrease in the rate of teenage pregnancy. ANNIE E. CASEY FOUNDATION, 2005 KIDS COUNT POCKET GUIDE 7 (2005), *available at* http://www.aecf.org/kidscount/sld/db05_pdfs/entire_db.pdf (noting that the teen birth rate has steadily declined since 1991, reaching its lowest level ever, 43 births per 1,000, in 2002). Medicaid provides two-thirds of all federal and state family planning funding. NATIONAL WOMEN'S LAW CENTER, MEDICAID, BIRTH CONTROL AND WOMEN'S HEALTH, *supra* at 1.

370. *See* MEDICAID FACTS, *supra* note 29, at 1. EPSDT services provide all necessary preventative, diagnostic, and screening care necessary to prevent and treat acute and chronic health conditions, including both physical and mental ailments. Medicaid's goal in insisting that they be provided is not only to ameliorate individual patient suffering, but also to prevent the development of more serious health problems, which are both expensive and debilitating. *Id.*

371. *Id.* at 2.

372. *Id.* Some contend that the increased enrollment of Medicaid beneficiaries into managed care is the source of this failure, *id.*, while others assert that managed care enhances access to health care services, because Medicaid beneficiaries now have a medical home, rather than being forced to hunt for a provider willing to accept the historically low Medicaid fees. *See* Vernon Smith & Linda Hamacher, *The "Good Olde Days" of Fee-for-Service Were Not So Good After All: Managed Care Has Made Things Better* 3, 6-8 (Ass'n of Health Ctr. Affiliated Health Plans, Working Paper, May 2003), *available at* http://www.ahcahp.org/publications/Working%20Papers/ffs%20is%20bad.pdf.

373. Suits have been brought in at least 28 states. *See* GAO, MEDICAID REPORT, *supra* note 367, at 9; *see, e.g.*, Westside Mothers v. Haveman, 289 F.3d 852, 863 (6th Cir. 2002) (holding that parents could seek injunctive relief against Michigan and managed care organizations with which Michigan had contracted to enforce the children's rights to receive EPSDT services); *see also* Frew v. Hawkins, 540 U.S. 431 (2004) (upholding a federal court consent decree in which Texas health officials agreed to provide EPSDT services to the more than one million child beneficiaries of the Texas Medicaid program, many of whom received services via managed care).

services to which they are entitled.[374] Other barriers to health care access under Medicaid are discussed below.

The State Child Health Insurance Program (SCHIP) was inaugurated in 1997, with the goal of giving states additional options in addressing the health care needs of children. In contrast to Medicaid, whose recipients are entitled to receive all mandated EPSDT services, SCHIP permits states to offer a less generous package of benefits, with the goal of reaching a larger group of enrollees.[375] SCHIP also authorizes states to provide services to the low-income parents of eligible children, and this opportunity for parental enrollment has increased the number of children who receive health care services.[376] However, children's health advocates, notably Sara Rosenbaum, have strongly criticized the SCHIP program for promising more than it actually delivers in terms of services to children.[377]

Unlike Medicaid,[378] the SCHIP program can be curtailed if states decide it is too generous or that they cannot afford it.[379] In the late 1990s, when the economy was strong, many states engaged in significant outreach activities, and Medicaid and SCHIP enrollments boomed, even though many eligible children and families were still not enrolled in the programs.[380] But as the economy faltered in the early twenty-first century, many states began to face large budget deficits. They scrambled to limit enrollment in SCHIP, either by freezing enrollment numbers, increasing procedural obstacles to enrollment, imposing cost-sharing measures, or limiting outreach activities.[381] As a result, SCHIP enrollment has fallen dramatically in many states.[382]

Further, as a result of the enactment of the Deficit Reduction Act of 2005,[383] federal and state spending on Medicaid will be significantly curtailed,[384] and some Medicaid enrollees will have to spend much more

374. GAO, MEDICAID REPORT, *supra* note 367, at 1-2, 7-8.

375. Rosenbaum, Markus & Sonosky, *supra* note 29, at 18-21.

376. KAISER, IN A TIME OF GROWING NEED, *supra* note 160, at 9-10.

377. Rosenbaum, Markus & Sonosky, *supra* note 29, at 18-21.

378. Medicaid provides states with some flexibility in provision of services through its waiver procedures, but historically it has been difficult for states to make major cuts in services. However, the Deficit Reduction Act of 2005, Pub. L. 109-171, has made significant changes in Medicaid. *See* discussion *infra* in text accompanying notes 384-388. KAISER COMMISSION ON MEDICAID AND THE UNINSURED, DEFICIT REDUCTION ACT OF 2005: IMPLICATIONS FOR MEDICAID 1-3 (Feb. 2006), [hereinafter KAISER, DEFICIT REDUCTION ACT], *available at* http://www.kff.org/medicaid/upload/7465.pdf.

379. KAISER, ENROLLING CHILDREN, *supra* note 57, at 2.

380. KAISER, IN A TIME OF GROWING NEED, *supra* note 160, at 2, 4

381. KAISER, ENROLLING CHILDREN, *supra* note 57, at 1-2.

382. *Id.*; *see also* KAISER, IN A TIME OF GROWING NEED, *supra* note 160, 1-4.

383. Pub. L. 109-171, 120 Stat. 4 (signed by President Bush Feb. 8, 2006).

384. The Congressional Budget Office estimates $4.8 billion in reductions over the period 2006 – 2010 and $26.1 billion over the next ten years. *See* KAISER, DEFICIT REDUCTION ACT, *supra* note 379, at 1.

out of pocket for their health care.[385] Congress made major changes in Medicaid, permitting states to charge families with incomes greater than 150% of the federal poverty level ($24,900 for a family of three in 2006) premiums and cost-sharing (co-pays, etc.) for health care services, although these cost-sharing requirements are not to be applied to pregnant women and certain eligible children.[386] In addition, the law makes it harder for certain groups of children to receive the preventative EPSDT services previously mandated.[387]

3. Insurance Alone is Not the Answer

Two decades of research on Medicaid and SCHIP have shown that merely making government insurance available is insufficient to ensure adequate care, for a number of reasons.[388] First, because Medicaid was originally conceived of as part of the welfare system, it lacks necessary political support,[389] and many health care professionals choose not to participate because of the very low reimbursement rates.[390] Medicaid recipients often feel stigmatized, and many eligible families are discouraged from enrolling.[391] As noted, in some states, the shift of poor children and families into Medicaid and SCHIP managed care programs has created access and service problems which parallel those of middle class families in managed care,[392] with children failing to receive preventative screenings or other mandated services.[393] In other states, however, Medicaid managed care delivers better health care services to its enrol-

385. Although it appears that the primary impact of the Deficit Reduction Act of 2005 (DRA) will be to curtail government spending on behalf of elderly and disabled adults, the DRA will also affect some children and their families. *Id.* at 1-6.

386. *Id.* at 1-3.

387. *Id.* at 3.

388. ROBERT WOOD JOHNSON FOUNDATION, *supra* note 32, at 5.

389. John K. Iglehart, *The American Health Care System: Medicaid*, 340 N. E. J. MED. 403, 407 (1999) (noting that "Medicaid underscores the ambivalence of a society that continually struggles with the question of which citizens deserve access to publicly financed medical care, and under what conditions").

390. GAO, MEDICAID REPORT, *supra* note 367, at 13-14.

391. Alexandra Marks, *Healthcare 'Crisis' Grows for Middle Class*, CHRISTIAN SCIENCE MONITOR, Apr. 3, 2002, at USA 3; Lauren Terrazzano, *More Kids Uninsured; L.I. Is One of the Nation's Richest Areas But Its Children Are Insurance-Poor in Greater Numbers*, NEWSDAY, Aug. 20, 2005, at A08.

392. Some access problems are inevitable in a system which provides disincentives to treat. *See* Pegram v. Herdrich, 530 U.S. 211 (2000) (describing the incentives inherent in both managed care and fee-for-service medicine). For a fuller discussion of the problems of managed care, see CLARK C. HAVIGHURST, JAMES F. BLUMSTEIN, & TROYEN A. BRENNAN, HEALTH CARE LAW AND POLICY 1180-1298 (2d ed. 1998). Of course, one should be cautious in bashing managed care, since the fee-for-service health care system also has undesirable incentives, particularly to over-treat, which can be equally bad for patient well-being. *Id.* at 160-83.

393. GAO, MEDICAID REPORT, *supra* note 367, at 9-10, 12-13. For example, only 19% of Medicaid enrolled children five and under had been screened for lead poisoning, even though this group of children is "almost five times more likely than others to have a harmful blood lead level." *Id.* at 12. Only a fifth of eligible children aged two to five had visited a dentist within the previous year. *Id.*

lees than the traditional fee-for-service model.[394] In any case, since Medicaid and SCHIP enrollees are frequently less well-educated, lack child care and convenient transportation, and are not native English speakers, it may be difficult for them to receive all of the care to which they are entitled.[395]

Further, either by design or inadvertence, Medicaid and SCHIP have substantial barriers to enrollment and utilization. These include complex eligibility rules (including in many states, denial of eligibility if the parents have even limited assets), cumbersome forms to fill out at inconvenient locations, and requirements of frequent reenrollment, as often as every six months.[396] SCHIP permits enrollees to be charged premiums or co-payments.[397] This can be a significant burden for low-income families enrolled in SCHIP.[398] Although it is necessary to ensure that enrollees meet the statutory means tests, and to acknowledge the possibility of "crowd-out" (the phenomenon by which consumers shift from privately funded health insurance to public programs),[399] if concerns about fraud or crowd-out become a major focus, many children will not have access to health care.[400]

4. Inadequate Substance Abuse Programs

The resources presently available to treat women who abuse alcohol and other drugs are woefully inadequate. There are three major problems with most substance abuse programs: 1) they fail to recognize the significant relationship between domestic violence and women's mental illness and substance abuse;[401] 2) they do not take into account the differing treatment needs of men and women;[402] and 3) they do not provide the complementary support necessary for pregnant women and mothers to succeed in beating their addiction.[403] Only 14% of the drug treatment

394. *See, e.g.*, Patrick J. Roohan et al., *Quality Measurement in Medicaid Managed Care and Fee-for-Service, the New York State Experience*, 21 AM. J. MED. QUALITY 185 (2006); Smith & Hamacher, *supra* note 372.

395. GAO, MEDICAID REPORT, *supra* note 367, at 14.

396. KAISER, IN A TIME OF GROWING NEED, *supra* note 160, at 4-10 (summarizing recent changes made by states in Medicaid and SCHIP and their impact on enrollment).

397. *Id.* at 7, 13.

398. *Id.* at 6-8.

399. *Id.* at 14-15. For a general discussion of the crowd-out phenomenon, see John V. Jacobi, *Government Reinsurance Programs and Consumer-Driven Care*, 53 BUFF. L. REV. 537, 571-73 (2005).

400. *See* URBAN INSTITUTE, *supra* note 356.

401. Only 35% of drug treatment facilities in the United States have programs for persons needing treatment for both substance abuse and mental illness. *See* SUBSTANCE ABUSE & MENTAL HEALTH SERVICES ADMIN. OFFICE OF APPLIED STUDIES, DEP'T OF HEALTH & HUMAN SERVICES, NATIONAL SURVEY OF SUBSTANCE ABUSE TREATMENT SERVICES (N-SSATS): 2003, DATA ON SUBSTANCE ABUSE TREATMENT FACILITIES 4 (Sept. 2004), *available at* http://wwwdasis.samhsa.gov/03nssats/nssats_rpt_03.pdf.

402. *See id.*

403. *See id.*

facilities in the United States have programs specifically designed to treat pregnant and postpartum women.[404]

Many women who abuse alcohol and other drugs were sexually abused or beaten as children and have significant mental health and self-esteem issues, which make it much more likely that they will misuse drugs.[405] Without acknowledgement of the causal connections between domestic violence, substance abuse, and mental illness and active intervention to prevent current domestic violence from continuing, women will not receive the support necessary to recover from addiction and mental illness.[406] Those who work in the government agencies that deal with domestic violence victims, including police, hospital staff, and social workers need more training in understanding the broad context of domestic violence, in order for their interventions to be appropriate and effective.[407]

Many drug treatment programs are not designed with the needs of women in mind. For example, traditional confrontational approaches, effective with male drug addicts, do not work well with women,[408] and women also have better treatment outcomes in programs that are for women only.[409] For women who are long-term abusers, residential programs are most effective,[410] but these programs must take into account the needs of women with children.[411] Child care, housing, health care, job training, and other vital supports are necessary if women are to stay "clean" and become self-sufficient.[412] Further, the Adoption and Safe Families Act should be amended to provide an exemption from its strict time limits, to acknowledge that addiction recovery does not fit neatly into the statutory timetable.[413] Finally, more programs must emphasize prevention, to treat addicted women before they become pregnant.[414]

5. The Consequences of America's Inadequate Health Care System

Research over the last several decades has made clear the consequences of inadequate health care for America's women and children. Both maternal and infant mortality are higher in the United States than in

---

404. *Id.*

405. Paltrow, *supra* note 12, at 477; WOMEN'S LAW PROJECT, *supra* note 12, at 1.

406. WOMEN'S LAW PROJECT, *supra* note 12, at 3.

407. *Id.*

408. SANDRA L. BLOOM, THE PVS DISASTER: POVERTY, VIOLENCE, AND SUBSTANCE ABUSE IN THE LIVES OF WOMEN AND CHILDREN 165 (2002).

409. *Id.* at 164; WOMEN'S LAW PROJECT, *supra* note 12, at 5, 23.

410. BLOOM, *supra* note 408, at 108 (citing M. Daley et al., *The Impact of Substance Abuse Treatment Modality on Birth Weight and Health Care Expenditures*, 33 J. PSYCHOACTIVE DRUGS 57-66 (2001)).

411. BLOOM, *supra* note 408, at 164.

412. *Id.* (citing NAT'L INST. OF DRUG ADDICTION, TREATMENT METHODS FOR WOMEN 13652 (1999)).

413. *See generally* Roberts, *Punishing Drug Addicts Who Have Babies*, *supra* note 9.

414. *See generally* WOMEN'S LAW PROJECT, *supra* note 12.

many other nations.[415] American children are more likely to be born preterm and at low or very low birthweight,[416] and are less likely to have preventative doctors' visits, obtain necessary immunizations, and access necessary reproductive and mental health care.[417] Study after study has shown that a focus on pregnant women as a vehicle for ensuring healthy children is too little, too late. Instead, comprehensive solutions, which address the systemic failure to take care of America's children, must be developed.

III. CONCLUSIONS AND RECOMMENDATIONS

If we are to truly assist children to become healthy adults who are able to embrace life's opportunities, we must explore new ways of thinking about the health of children and the women who bear and raise them. There are six areas where change is crucial: ending poverty, providing universal health care, expanding substance abuse prevention and treatment programs, enhancing environmental and workplace protections, instituting no-fault compensation for children who are harmed in utero, and ending criminal and civil actions against pregnant women who may be placing their fetuses at risk.

End Poverty

Even making this recommendation seems both naïve and incredibly ambitious; yet it is an inescapable fact that being poor has serious adverse consequences for children's health and development. Children living in poverty (who are also more likely to be malnourished and homeless) have more learning disabilities and mental retardation, lower IQs, and higher rates of mental illness, behavioral problems, and greater physical health problems, than middle-income children.[418] The effects of childhood poverty continue through adulthood, perpetuating the cyclical connection between inadequate parental income and childhood disease and dysfunction.[419] With twelve million American children living in families with incomes less than the federal poverty level (and five mil-

415. *See supra* text accompanying notes 52-59.

416. IOM Report, *supra* note 55.

417. KAISER, ENROLLING CHILDREN, *supra* note 57, at 1; GAO, MEDICAID REPORT, *supra* note 367, at 9, 12-13.

418. Charles Oberg, Maternal & Child Health Program, School of Public Health, University of Minnesota, *The Impact of Childhood Poverty on Health and Development*, HEALTHY GENERATIONS, May 2003, at 2 & 3 nn.7-10; *See also* Jane D. McLeod & Michael J. Shanahan, *Trajectories of Poverty and Children's Mental Health*, 37 J. HEALTH & SOCIAL BEHAVIOR 207, 207 (1996).

419. Anne Case, Darren Lubotsky & Christina Paxson, *Economic Status and Health in Childhood: The Origins of the Gradient*, 92 AM. ECON. REVIEW 1308, 1308-09 (2002). Studies indicate that malnutrition in utero has significant life-time consequences, which actually are more pronounced as people age. *See* Gina Kolata, *So Big and Healthy Nowadays, Grandpa Wouldn't Know You*, N.Y. TIMES, July 30, 2006, at A1.

lion of those children living on less than half that amount),[420] intervention is critical.

Provide Universal Health Care

Universal health care coverage is essential if we are to provide children with the health care services necessary for them to grow, learn, and develop into healthy and productive adults, who in turn will have healthy children of their own. While there are many historical and philosophical reasons for America's reliance on the market to provide health care for its citizens,[421] we can no longer afford to ignore the health care needs of the one-sixth of the population who lack health insurance of any kind.[422] Estimates of the cost of providing health care coverage for all Americans range from thirty-three to sixty-nine billion dollars annually,[423] potentially less than the amount the American government currently spends on the war in Iraq.[424] While several states have recently enacted laws expanding health care coverage[425] a comprehensive solution requires a federal effort.

At the very least, a uniform federal health care program for children with a comprehensive set of benefits and services should be established.[426] This would avoid the cyclical contractions and expansions of state Medicaid and SCHIP programs which presently accompany economic upswings and downturns and make it difficult for states to pay for

---

420. NATIONAL CENTER FOR CHILDHOOD POVERTY, MAILMAN SCHOOL OF PUBLIC HEALTH, COLUMBIA UNIVERSITY, WHO ARE AMERICA'S POOR CHILDREN 1, *available at* http://www.nccp.org/pub_cpt05b.html.

421. *See, e.g.*, Timothy S. Jost, *Why Can't We Do What They Do? National Health Reform Abroad*, 32 J. L. MED. & ETHICS 433 (2004).

422. URBAN INSTITUTE, *supra* note 356 (finding that 46.6 million Americans did not have insurance in 2005). As a practical matter, the uninsured do receive some health care through emergency room visits. Such care is expensive and often time-consuming. It is estimated that one-third of the care provided at hospital emergency departments is inappropriate. Ceci Connolly, *Some Finding No Room at the ER; Screening Out Non-Urgent Cases Stirs Controversy*, WASH. POST, Apr. 26, 2004, at A01. The high costs of providing emergency room care required under EMTALA, the Emergency Medical Treatment and Active Labor Act, and state anti-dumping laws are borne by hospitals and ultimately, the tax-payer. *See* Aliessa v. Novello, 754 N.E.2d 1085, 1093 (N.Y 2001) (discussing the problem of immigrants who are denied care until their medical situation becomes an emergency).

423. URBAN INSTITUTE, *supra* note 356. These estimates appear to be in 2004 dollars. If it is not possible to establish health care for the entire population immediately, then, at a minimum, full coverage for children's health care should be established now, before the mass of Baby Boomers age into retirement, Medicare, and the need for long-term care.

424. In February, 2006, the Department of Defense stated that it was spending $4.5 billion a month (or $54 billion a year) on the Iraq war. Mark Mazetti & Joel Haveman, *Iraq War is Costing $100,000 per Minute*, SEATTLE TIMES, Feb. 3, 2006; Mark Silva, *$70 Billion Sought for War Costs; White House Says Another $50 Billion Needed for 2007*, CHI. TRIB., Feb. 3, 2006, at C1.

425. The most notable are the Massachusetts Health Care Access and Affordability Act, ch.. 58 (2006), *available at* http://www.mass.gov/legis/laws/seslaw06/sl060058.htm, and the Maryland Fair Share Health Care Fund Act, MD. CODE ANN., LAB. & EMPL. § 8.5-101 *et. seq, partially invalidated by* Retail Indus. Leaders Ass'n v. Fielder, 435 F. Supp. 2d 481 (D. Md. 2006) (finding the act preempted by the Employment Retirement Income Security Act of 1974).

426. The Medicaid EPSDT program should be seen as a floor, not a ceiling. *See* Rosenbaum, Markus & Sonosky, *supra* note 29, at 43; *see also supra* text accompanying notes 370-401.

adequate health care in times of fiscal exigency.[427] With national universal coverage, children will not lose access to vital health services because their parents move or change jobs, fail to fill out cumbersome paperwork, earn slightly more or less income, or are unable to afford premiums and co-payments.[428] Further, boys and girls who receive good health care will become adults who are more likely to bear healthy children.

Ultimately, however, if we want to meet the goal of having more American children in good health, it will be necessary to provide health care for all adults as well. Issues of nutrition, infertility, sterility, sexually transmitted diseases, and reproductive problems must be addressed in the adult population if we are to achieve better birth outcomes.

Provide Expanded and Targeted Substance Abuse Programs for Pregnant Women and Addicts Who Are Likely to Become Pregnant

As noted above, substance abuse education and treatment programs must be expanded and improved in order to serve both addicted women and the children they bear. Treatment which takes into account the special needs of women with drug and alcohol problems has been shown to be effective and to save money over the long run.[429]

Improve Environmental, Workplace, and Public Health Protections

At the same time, medical treatment alone is insufficient to ensure children's health. Environmental, workplace, and other public health laws must be strengthened to protect children from exposure to toxic substances, whether exposure is in utero or after birth. In addition, the government should mandate paid parental leave so that parents will be able to care for their children when they are infants or ill.

Establish a No-Fault Program to Compensate Children Who Suffer Prenatal Harm

One way that the government can respect the autonomy of pregnant women, compensate children who are harmed due to prenatal injury or exposure to toxic substances, and respond to the fears of employers and others about tort liability is to establish a national prenatal injury compensation program. Such a program could be funded by modest contributions by employers and manufacturers of toxic substances, including alcohol. Such a program could be modeled on the National Childhood Vaccine Injury (NCVI) program,[430] enacted in 1986 to encourage vac-

427. Smith & Hamacher, *supra* note 372, at 12-13.

428. *See generally* KAISER, IN A TIME OF GROWING NEED, *supra* note 160.

429. BLOOM, *supra* note 408, at 163-75.

430. The National Childhood Vaccine Injury Act of 1986, 42 U.S.C. §§ 300aa-1-33, created a program to provide compensation children injured by state-mandated vaccines, as a means of ensuring that children who suffer injury from vaccination will be compensated, that parents will be en-

cine manufacturers to continue producing vaccines for childhood diseases while simultaneously compensating the small number of children who were injured as a result of vaccination, and thus reassure parents who might otherwise decide not to vaccinate their children. The NCVI has proved extraordinarily successful in meeting all three of its goals, and has been touted as a model for other tort-based consumer protection problems.[431]

The program I propose acknowledges that women, men, and children face risks from the food and drink they consume, the environment, and the workplace, and that manufacturers of dangerous substances should be held responsible for the harm caused by in utero exposure, even when they try to minimize those risks. Children who suffer harm from an otherwise socially desired or valued product must be compensated, just as they are under the National Childhood Vaccine Injury Program. The proposed program would involve a compromise, limiting liability of manufacturers and employers in exchange for guaranteeing compensation to prenatally injured children. These trade-offs are superior to the current approach of excluding women from the workplace or otherwise penalizing women for their conduct during pregnancy. Manufacturers would have an incentive to minimize the exposure to the toxic substances they use and produce, whether it is lead used to make batteries, mercury and other environmental contaminants, nicotine and tar in cigarettes, or alcohol in wine, beer, and liquor.[432] Even sellers and distributors of illegal drugs could be made to contribute to the PIC fund, by requiring monetary victim restitution as part of their criminal sentences.[433]

---

courage to vaccinate their children, and that manufacturers will not be discouraged from entering and participating in the vaccine market because of fears of liability for products that are "unavoidably unsafe." Theodore H. Davis, Jr. & Catherine B. Bowman, *No-Fault Compensation for Unavoidable Injuries: Evaluating the National Childhood Vaccine Injury Compensation Program*, 16 U. DAYTON L. REV. 277, 279 (1991).

431. Derry Ridgway, *No-Fault Vaccine Insurance, Lessons from the National Vaccine Injury Compensation Program*, 24 J. HEALTH POL. POL'Y & LAW 59, 76-88 (1999); Geoffrey Evans, *Update on Vaccine Liability in the United States: Presentation at the National Vaccine Program Office Workshop on Strengthening the Supply of Routinely Recommended Vaccines in the United States, 12 February 2002*, 42 CLINICAL INFECTIOUS DISEASES S130-S137 (Mar. 1, 2006), abstract *available at* http://www.journals.uchicago.edu/cgi-bin/resolve?id=doi:10.1086/499592&erFrom=7871817745696282041Guest. Under the National Vaccine Injury Compensation Program, manufacturers pay a tax of $0.75 per dose of vaccine administered. Ridgway, *supra* at 62, http://content.nejm.org/cgi/reprint/340/5/403.pdf.

432. Thus, for example, alcohol manufacturers and distributors would have an economic incentive to make warning labels about the effects of alcohol during pregnancy clearer, more conspicuous, and more explicit. Press Release, Center for Science in the Public Interest, Alcohol Warning Labels Go Unnoticed, Poll Finds (Aug. 20, 200), *available at* http://www.cspinet.org/booze/batf_labels2001_press.htm.

433. *Cf.* U.S. SENTENCING GUIDELINES MANUAL §§ 5E1.1, 8B1.1 (2005), *available at* http://www.ussc.gov/2005guid/5e1_1.htm and http://www.ussc.gov/2005guid/8b1_1.htm (requiring individual and organizational defendants to pay restitution to their victims).

End Civil Commitment and Involuntary Medical Treatment of Pregnant Women

The involuntary restraint and compulsory medical treatment of pregnant women is counterproductive, deterring women from seeking medical and psychological help. Physicians and hospitals are fallible. They do a profound disservice to those whom they wish to help when they rely on court orders rather than trying to advise and persuade pregnant patients about what is in their (and their fetuses') best interest. There is substantial evidence that health care providers are relying on racial and class stereotypes when they decide when to seek judicial intervention.[434] Similarly, there is no reason for legislatures to disable women from exercising the right to self-determination when pregnant, by rendering their advance medical directives invalid.

End Civil and Criminal Liability of Pregnant Women for Causing Prenatal Harm

The thesis of this article is that holding women civilly or criminally liable for their actions while pregnant is bad public policy. Imposing criminal or civil liability deters women from seeking medical care, including treatment for drug and alcohol addiction, leading to worse, rather than better, birth outcomes, and raises significant normative questions about who is the reasonable pregnant woman.[435] Some American courts,[436] as well as those in other countries,[437] have acknowledged that there is no way to prescribe the standard of appropriate behavior while pregnant with any certainty, that making a judgment about recklessness or negligence is inevitably subjective, and thus is freighted with the possibility of prejudice and bias. Further, most efforts at criminal prosecution or civil commitment have focused on poor women and women of color,[438] despite evidence that drug usage during pregnancy is equivalent across racial and economic lines, with the only difference being that white and middle-class women tend to use alcohol, a legal drug, rather than cocaine.[439]

Attacking pregnant women provides a simplistic solution to a complex problem. Courts and legislatures should avoid this meretricious solution, recognizing the unique relationship between pregnant woman

434. *See generally* Chasnoff, *supra* note 10; Roberts, *Unshackling Black Motherhood*, *supra* note 9.

435. Chenault v. Huie, 989 S.W.2d 474, 477-78 (Tex. App. 1999).

436. *Chenault*, 989 S.W.2d at 477-78; Stallman v. Youngquist, 531 N.E.2d 355, 360 (Ill. 1988); Kilmon v. State, 905 A.2d 306, 310-15 (Md. 2006).

437. Dobson v. Dobson, 2 S.C.R. 753 (Can. 1999); Winnipeg Child & Family Services (Northwest Area) v. G. (D.F.), 3 S.C.R. 925 (Can. 1997); Paton v. United Kingdom, App. No. 8416/78, 3 Eur. H.R. Rep. 408, 415 (1980).

438. BLOOM, *supra* note 408, at 109; Chasnoff, *supra* note 10; Roberts, *Unshackling Black Motherhood*, *supra* note 9.

439. BLOOM, *supra* note 408, at 109-10.

and the fetus.[440] Government policymakers should acknowledge that the vast majority of pregnant women want only the best for the fetus whom they are nourishing, and that in almost all cases women who are not acting in the best interests of their fetus are facing heavy burdens of poverty, addiction, lack of access to quality health care, and domestic abuse. The way to help such women, and the children they will bear, is to change the system in which they are now struggling, not to make pregnant women the scapegoat for that system's failures.

440. *Stallman*, 531 N.E.2d at 360.

THE CONSTITUTIONAL RIGHT TO PRIVACY

LEE GOLDMAN†

INTRODUCTION

The Due Process Clause of the Fourteenth Amendment provides in part that no State shall "deprive any person of life, liberty, or property, without due process of law . . . ."[1] This Clause "guarantees more than fair process"—it imposes substantive restraints on government power.[2] Although the Court's substantive due process doctrine often has been criticized,[3] it is now well established[4] and provides protection for so-called fundamental rights.[5]

According to traditional doctrine, if government action substantially interferes with a fundamental right, the state must demonstrate that the

---

† Professor, University of Detroit Mercy School of Law, J.D. 1979.

1. U.S. CONST. amend. XIV, § 1. The counterpart Due Process Clause of the Fifth Amendment imposes an identical restraint on the federal government. U.S. CONST. amend. V.

2. Troxel v. Granville, 530 U.S. 57, 65 (2000) (quoting Washington v. Glucksberg, 521 U.S. 702, 719 (1997)).

3. *See, e.g.*, CHARLES L. BLACK, JR., A NEW BIRTH OF FREEDOM: HUMAN RIGHTS, NAMED & UNNAMED 3 (1997) ("This paradoxical, even oxymoronic phrase—'*substantive* due *process*'—has been inflated into a patched and leaky tire on which precariously rides the load of some substantive human rights not named in the Constitution."); ROBERT H. BORK, THE TEMPTING OF AMERICA: THE POLITICAL SEDUCTION OF THE LAW 31 (1990) (explaining substantive due process is a "momentous sham"); JOHN HART ELY, DEMOCRACY AND DISTRUST: A THEORY OF JUDICIAL REVIEW 18 (1980) ("'[S]ubstantive due process' is a contradiction in terms—sort of like 'green pastel redness.'").

4. *See Glucksberg*, 521 U.S. at 719-20; Zablocki v. Redhail, 434 U.S. 374, 384 (1978); *see also* David Crump, *How Do Courts Really Discover Unenumerated Fundamental Rights? Cataloguing the Methods of Judicial Alchemy*, 19 HARV. J. L. & PUB. POL'Y 795, 838 (1996) ("The Supreme Court consistently has . . . recognize[d] unenumerated fundamental rights . . . ."); James E. Fleming, *Securing Deliberate Autonomy*, 48 STAN. L. REV. 1, 13 (1995) ("Griswold today is a case that any nominee, to stand a chance of being confirmed, has to say is rightly decided.").

5. Although substantive due process and fundamental rights doctrine sometimes are used interchangeably, they are not equivalents. Substantive due process, in addition to securing certain fundamental rights, *see infra* note 13, protects against arbitrary government action, *see Glucksberg*, 521 U.S. at 766 (Souter, J., concurring); Daniels v. Williams, 474 U.S. 327, 331 (1986); Poe v. Ullman, 367 U.S. 497, 543-44 (Harlan, J., dissenting), safeguards individuals from conduct by government officers that "shocks the conscience," County of Sacramento v. Lewis, 523 U.S. 833, 846 (1998); United States v. Salerno, 481 U.S. 739, 746 (1987) (quoting Rochin v. California, 342 U.S. 165, 172 (1952)), and limits the size of civil punitive damages, State Farm Auto Ins. Co. v. Campbell, 538 U.S. 408, 416 (2003). Nevertheless, this article focuses on the fundamental rights branch of substantive due process generally, and the right to privacy, specifically.

action is narrowly tailored to further a compelling government interest.[6] If no fundamental right is involved, the government need only establish a rational basis for the challenged action.[7] Thus, determining whether there is a fundamental right involved becomes critical. Unfortunately, given the political differences of the Justices and the lack of any clear conceptualization in this area, determining whether a fundamental right exists has proven to be a Herculean task.

The Supreme Court Justices have adopted two, often conflicting, approaches to determine whether a case involves a fundamental right. The more liberal Justices, seeking to protect minority interests, ask whether a right is central to personal dignity and autonomy or is at the heart of liberty.[8] The more conservative Justices, fearing judicial activism at the expense of democratic preferences, insist that a right is not fundamental unless it is "deeply rooted in this Nation's history and tradition" and "implicit in the concept of ordered liberty."[9] The former test is easily criticized as too indeterminate,[10] the latter as protecting only those rights that don't need protection.[11] The difficulty in determining whether a fundamental right exists is compounded by disingenuous application of the Court's compelling government interest and rational basis review.[12] It is not surprising then that there is little clarity on questions ranging from the constitutionality of bans on same-sex marriages or the sale of sex toys to criminalization of adultery, incest, or the use of marijuana for medical purposes.

This article proposes a conceptualization of a central branch of the fundamental rights doctrine—the constitutional right to privacy,[13] which

6. *See* Lawrence v. Texas, 539 U.S. 558, 593 (2003) (Scalia, J., dissenting); *Glucksberg*, 521 U.S. at 721; Reno v. Flores, 507 U.S. 292, 302 (1993); Carey v. Population Serv. Int'l, 431 U.S. 678, 686 (1977) (citing Roe v. Wade, 410 U.S. 113, 155-56 (1973)). Despite what traditional doctrine provides, this article argues that a sliding scale approach to fundamental rights issues best balances competing government and individual interests and is consistent with actual Supreme Court practice. *See infra* notes 177-234 and accompanying text.

7. *See Glucksberg*, 521 U.S. at 766-67 & n.9 (Souter, J., concurring); Lyng v. Castillo, 477 U.S. 635, 638-39 (1986).

8. *See* Planned Parenthood of Se. Pa. v. Casey, 505 U.S. 833, 851 (1992) (plurality opinion); *see also Lawrence*, 539 U.S. at 562 ("Liberty presumes an autonomy of self that includes freedom of thought, belief, expression, and certain intimate conduct.").

9. *Lawrence*, 539 U.S. at 593 (Scalia, J., dissenting) (quoting *Glucksberg*, 521 U.S. at 721).

10. *See, e.g.*, Crump, *supra* note 4, at 854-56.

11. *See* Michael H. v. Gerald D., 491 U.S. 110, 140-41 (1989) (Brennan, J., dissenting); *see also infra* notes 96-97 and accompanying text; Adam B. Wolf, *Fundamentally Flawed: Tradition and Fundamental Rights*, 57 U. MIAMI L. REV. 101, 115 (2002).

12. *See infra* notes 106-26 and accompanying text.

13. In addition to providing protection for privacy interests, the fundamental rights branch of substantive due process, *see supra* note 5, incorporates key provisions of the Bill of Rights against the states, *see* Duncan v. Louisiana, 391 U.S. 145, 150 n.14 (1968); Peter J. Rubin, *Square Pegs and Round Holes: Substantive Due Process, Procedural Due Process, and the Bill of Rights*, 103 COLUM. L. REV. 833, 842 (2003), and includes protection for the rights to vote, *see* Harper v. Virginia State Bd. of Elections, 383 U.S. 663, 670 (1966); to travel interstate, *see* United States v. Guest, 383 U.S. 745, 757 (1966); to access the courts, M.L.B. v. S.L.J., 519 U.S. 102, 107 (1996); Griffin v. Illinois, 351 U.S. 12, 17 (1956), and to be free of totalitarian legislation, *see Poe*, 367 U.S. at 521-22 (1961) (Douglas, J., dissenting) (quoting Robert L. Calhoun, *Democracy and Natural Law*,

remedies some of the deficiencies in the Court's jurisprudence. Specifically, this article argues for a Lockean view[14] of the Constitution as a pact between individuals and the government to forego certain rights that are necessary to further society's interests, but with a reservation of rights in certain private areas where the government does not belong. When the government regulates in an area where it does not belong, presumptively it needs the regulation to be narrowly tailored to achieve a compelling government interest. If regulation is in an area where the government does belong and the regulation does not significantly affect private interests, presumptively the regulation is valid as long as there is a rational basis for the regulation. However, if a regulation in an area where the government belongs significantly affects private interests, a balancing test should be applied, giving deference to the legislature's initial determination of the appropriate balance. By specifically indicating the areas where the government does and does not belong and identifying the most important variables in the balance when balancing is appropriate, this article hopes to bring a degree of clarity, or at least honesty and consistency, to an area in which it too long has been lacking. The proposed conceptualization appears to be consistent with the views of the Framers and early political philosophers,[15] as well as most of the Court's case law.[16] By providing a conceptualization and admitting to balancing in some cases, the recommended approach provides more honest analysis, better guidance to lower courts, and desired flexibility in evaluating regulations impacting important individual interests.

Part I of this article provides a brief review of existing fundamental rights law and the problems associated with both defining fundamental rights and applying the Court's standard of review. Part II defines the proposed right of privacy, specifying and justifying the areas where the government does and does not belong. It explains the factors to be considered in the proposed balancing or sliding scale test and responds to anticipated criticisms of the balancing approach. Finally, Part III ad-

---

5 NAT. L. F. 31, 36 (1960)); Pierce v. Soc'y of the Sisters of the Holy Names of Jesus and Mary, 268 U.S. 510, 535 (1925); Jeb Rubenfeld, *The Right of Privacy*, 102 HARV. L. REV. 737, 806 (1989), rights that are deemed fundamental to the structure of our governmental system. The constitutional right to privacy, as used in this article, refers to the unenumerated right to privacy protected by substantive due process. The article does not address privacy interests protected by specific provisions of the Constitution, for example, the Fourth Amendment's right to be free of unreasonable search and seizures, *see* U.S. CONST. amend. IV; Katz v. United States, 389 U.S. 347, 361-62 (1967) (Harlan, J., concurring), or tort concepts of privacy. *See generally* Daniel J. Solove, *Conceptualizing Privacy*, 90 CAL. L. REV. 1087, 1100 (2002); Samuel D. Warren, & Louis D. Brandeis, *The Right to Privacy*, 4 HARV. L. REV. 193, 195 (1890).

14. *See* JOHN LOCKE, TWO TREATISES OF GOVERNMENT 353 (Peter Laslett ed., Cambridge Univ. Press 1988) (1690); *see also* RANDY E. BARNETT, RESTORING THE LOST CONSTITUTION: THE PRESUMPTION OF LIBERTY 323-29 (2004); A. JOHN SIMMONS, THE LOCKEAN THEORY OF RIGHTS 218-20 (1992); Mark C. Niles, *Ninth Amendment Adjudication: An Alternative to Substantive Due Process Analysis of Personal Autonomy Rights*, 48 U.C.L.A. L. REV. 85, 110-12 (2000).

15. *See* BARNETT, *supra* note 14, at 68-76; Niles, *supra* note 14, at 108.

16. *See infra* Part I.

dresses many of the "hot" substantive due process questions, such as the validity of bans on gay marriage and the use of medical marijuana, to illustrate application of the recommended approach.

I. EXISTING FUNDAMENTAL RIGHTS/PRIVACY LAW

A. Determining Whether a Fundamental Right Exists

Ironically, much of the modern fundamental rights/privacy doctrine derives from dissents in a case dismissed for lack of justiciability. In *Poe v. Ullman*,[17] plaintiffs challenged a Connecticut statute forbidding the giving of contraceptive advice and the use of contraceptives.[18] The Court held that there was no justiciable controversy based on its finding that Connecticut had chosen not to enforce the statute.[19] Both Justice Douglas and Justice Harlan dissented.[20] Foreshadowing the Court's decision in *Griswold v. Connecticut*,[21] the Justices found the statute unconstitutional as an invasion of privacy, a liberty interest protected by the Fourteenth Amendment.[22] Their approaches, however, presaged what would become a continuing controversy for the Court. Justice Douglas found that privacy is a right "implicit in a free society,"[23] finding support for the right in both the "totality of the Constitutional scheme" and the common law right "to be let alone."[24] He specifically rejected the notion that tradition was a suitable basis for defining protection under the Fourteenth Amendment, stating

> The due process clause . . . guarantees basic rights, not because they have become petrified as of any one time, but because due process follows the advancing standards of a free society as to what is deemed reasonable and right. It is to be applied, according to this view, to facts and circumstances as they arise, the cases falling on one side of the line or the other as a majority of the nine justices appraise conduct as either implicit in the concept of ordered liberty or as lying without the confines of that vague concept.[25]

Justice Harlan, finding the contraceptives ban an "invasion of privacy in the conduct of the most intimate concerns of an individual's personal life,"[26] agreed that the Connecticut statute violated the fundamental rights belonging "to the citizens of all free governments."[27] Justice

17. 367 U.S. 497 (1961).
18. *Poe*, 367 U.S. at 498.
19. *Id.* at 508.
20. *Id.* at 509, 522.
21. 381 U.S. 479 (1965).
22. *Poe*, 367 U.S. at 517 (Douglas, J., dissenting); *id.* at 539 (Harlan, J., dissenting).
23. *Id.* at 521 (Douglas, J., dissenting).
24. *Id.* & n.12.
25. *Id.* at 518 n.9 (quoting OWEN J. ROBERTS, THE COURT AND THE CONSTITUTION 80 (1951)).
26. *Id.* at 539 (Harlan, J., dissenting).
27. *Id.* at 541 (Harlan, J., dissenting).

Harlan, citing Justice Brandeis' dissent in *Olmstead v. United States*,[28] also found that the liberty interest of the Fourteenth Amendment "conferred, as against the government, the right to be let alone—the most comprehensive of rights and the right most valued by civilized men."[29] However, Justice Harlan, unlike Justice Douglas, did not feel comfortable allowing judges to roam at large.[30] Rather, he thought the balance between the liberty of the individual and the demands of an organized society should have "regard to what history teaches are the traditions from which it developed as well as the traditions from which it broke."[31] This disagreement concerning the proper role of tradition and the perspective from which it should be defined has been a continuing controversy for the Court.[32]

The Court as a whole explicitly recognized a right to privacy and invalidated Connecticut's contraceptives ban in *Griswold*.[33] Two years later in *Loving v. Virginia*,[34] the Court held that a ban on interracial marriages violated the Due Process Clause of the Fourteenth Amendment, finding the freedom to marry a "vital personal right[] essential to the orderly pursuit of happiness by free men."[35] In *Eisenstadt v. Baird*,[36] the right to use contraceptives recognized in *Griswold* was extended to unmarried couples.[37] The Court stated: "If the right of privacy means anything, it is the right of the *individual*, married or single, to be free from unwarranted governmental intrusion into matters so fundamentally affecting a person as the decision whether to bear or beget a child."[38] The Court next held a ban on abortions unconstitutional in *Roe v. Wade*,[39] stating that the "right of privacy . . . founded in the Fourteenth Amendment's concept of personal liberty . . . is broad enough to encompass a woman's decision whether or not to terminate her pregnancy."[40] The right of privacy was further expanded in *Moore v. City of East Cleveland*.[41] In that case, the Court ruled unconstitutional a zoning ordinance that limited occupancy in dwelling units to families narrowly defined to

28. 277 U.S. 438, 478 (1928) (Brandeis, J., dissenting).

29. *Poe*, 367 U.S. at 550 (Harlan, J., dissenting) (quoting *Olmstead*, 277 U.S. at 478 (Brandeis, J., dissenting)).

30. *Id.* at 544 (quoting Rochin v. California, 342 U.S. 165, 170 (1952).

31. *Id.* at 542. Justice Harlan went on to recognize that tradition "is a living thing," and defined the relevant tradition broadly as privacy in the individual's marital relations. *Id.* at 539, 542, 552.

32. *See infra* notes 62-84 and accompanying text; *see also* Michael H. v. Gerald D. 491 U.S. 110, 123 (1989).

33. 381 U.S. 479, 485-86 (1965).

34. 388 U.S. 1 (1967).

35. *Id.* at 12. The Court first held that the Virginia statute violated the Equal Protection Clause of the Fourteenth Amendment. *Id.* at 11-12.

36. 405 U.S. 438 (1972).

37. *Id.* at 454-55.

38. *Id.* at 453 (emphasis added).

39. 410 U.S. 113 (1973).

40. *Id.* at 153.

41. 431 U.S. 494 (1977) (plurality opinion).

exclude the plaintiff and her two grandsons.[42] The plurality opinion affirmed that the Court had "long recognized that freedom of personal choice in matters of marriage and family life is one of the liberties protected by the Due Process Clause of the Fourteenth Amendment."[43] In *Cruzan v. Director, Missouri Department of Health*,[44] the Court, although affirming Missouri's right to require clear and convincing evidence of an incompetent's wishes to withdraw life-sustaining medical treatment, assumed and strongly suggested that the Due Process Clause protected the right to refuse unwanted lifesaving medical treatment.[45] Most recently, in *Lawrence v. Texas*,[46] the Court, overruling *Bowers v. Hardwick*,[47] held unconstitutional Texas' statute making homosexual sodomy illegal.[48] The Court indicated that "Liberty protects the person from unwarranted government intrusions into a dwelling or other private places . . . . [It] presumes an autonomy of self that includes freedom of thought, belief, expression, and certain intimate conduct."[49]

Even as the right of privacy was expanding, several Justices, fearing a return to the *Lochner* era, expressed concern about the potentially unlimited reach of the Court's expansive language and ad hoc identification of fundamental rights.[50] The Court's retrenchment began with *Bow-*

---

42. *Id.* at 506.
43. *Id.* at 499 (quoting Cleveland Bd. of Educ. v. LaFleur, 414 U.S. 632, 639-40 (1974)).
44. 497 U.S. 261 (1990).
45. *Id.* at 281.
46. 539 U.S. 558 (2003).
47. 478 U.S. 186 (1986).
48. *Lawrence*, 539 U.S. at 578.
49. *Id.* at 562.
50. *See, e.g.*, Moore v. City of Cleveland, 431 U.S. 494, 502 (1977) (plurality opinion). In *Moore*, the plurality stated:

> Substantive due process has at times been a treacherous field for this Court. There are risks when the judicial branch gives enhanced protection to certain substantive liberties without the guidance of the more specific provisions of the Bill of Rights. As the history of the *Lochner* era demonstrates, there is reason for concern lest the only limits to such judicial intervention become the predilections of those who happen at the time to be Members of this Court. That history counsels caution and restraint.

Moore, 431 U.S. at 502 (footnote omitted). In addition, Justice White in his dissent stated:

> That the Court has ample precedent for the creation of new constitutional rights should not lead it to repeat the process at will. The Judiciary, including this Court, is the most vulnerable and comes nearest to illegitimacy when it deals with judge-made constitutional law having little or no cognizable roots in the language or even the design of the Constitution . . . . [Given] that much of the underpinning for the broad, substantive application of the Clause disappeared in the conflict between the Executive and the Judiciary in the 1930's and 1940's, the Court should be extremely reluctant to breathe still further substantive content into the Due Process Clause . . . .

Id. at 544 (White, J. dissenting). In *Griswold*, Justice Goldberg explained:

> In determining which rights are fundamental, judges are not left at large to decide cases in light of their personal and private notions. Rather, they must look to the 'traditions and [collective] conscience of our people' to determine whether a principle is 'so rooted [there] . . . as to be ranked as fundamental.

Griswold v. Connecticut, 381 U.S. 479, 493 (1965) (Goldberg, J., concurring) (quoting Snyder v. Massachusetts, 291 U.S. 97, 105 (1934)). Similarly, Justice Black, dissenting in *Griswold*, noted:

> The Due Process Clause . . . was liberally used by this Court to strike down economic legislation in the early decades of this century, threatening, many people thought, the

ers v. Hardwick,[51] a challenge to Georgia's sodomy statute. Justice White, writing for the Court, warned:

> The Court is most vulnerable and comes nearest to illegitimacy when it deals with judge-made constitutional law having little or no cognizable roots in the language or design of the Constitution. That this is so was painfully demonstrated by the face-off between the Executive and the Court in the 1930's, which resulted in the repudiation of much of the substantive gloss that the Court had placed on the Due Process Clauses of the Fifth and Fourteenth Amendments. There should be, therefore, great resistance to expand the substantive reach of those Clauses, particularly if it requires redefining the category of rights deemed to be fundamental.[52]

The Court, reversing the Court of Appeals, found the sodomy statute constitutional and rejected the claim that the Fourteenth Amendment provided protection of private sexual conduct between consenting adults.[53] Justice White limited the reach of the Due Process Clause by defining the right to be protected narrowly, asking whether there is a fundamental right for homosexuals to engage in acts of consensual sodomy.[54] By focusing on the specific conduct, rather than aspirational goals,[55] Justice White was easily able to conclude that such a right was not "'implicit in the concept of ordered liberty' such that 'neither liberty nor justice would exist if [they] were sacrificed'" [56] or "deeply rooted in this Nation's history and tradition,"[57] the two alternative tests he identified for determining fundamental rights.[58]

Justice Scalia, in *Michael H. v. Gerald D.*,[59] sought to further restrict the Court's fundamental rights jurisprudence. In *Michael H.*, a putative natural father whose blood tests indicated a 98.07% probability of paternity challenged a California statute creating a presumption that

---

> tranquility and stability of the Nation. That formula, based on subjective considerations of 'natural justice,' is no less dangerous when used to enforce this Court's views about personal rights than those about economic rights.

Griswold, 381 U.S. at 522 (Black, J., dissenting) (citing Lochner v. New York, 198 U.S. 45, 64 (1905)).

51. 478 U.S. 186, *overruled by Lawrence*, 539 U.S. at 558.

52. *Bowers*, 478 U.S. at 194-95.

53. *Id.* at 196.

54. *Id.*

55. The four dissenting justices challenged the majority's definition of the right involved stating, "This case is no more about 'a fundamental right to engage in sodomy' . . . than *Stanley v. Georgia* . . . was about a fundamental right to watch obscene movies . . . ." *Bowers*, 478 U.S. at 199 (Blackmun, J., dissenting) (quoting *id.* at 191 (majority opinion)). "Rather, this case is about 'the most comprehensive of rights and the right most valued by civilized men,' namely, 'the right to be let alone.'" *Id.* (Blackmun, J., dissenting) (quoting Olmstead v. United States, 277 U.S. 438, 478 (1928)).

56. *Id.* at 191-92 (majority opinion) (quoting Palko v. Connecticut, 302 U.S. 319, 325-26 (1937)).

57. *Id.* at 192 (quoting Moore v. City of E. Cleveland, 431 U.S. 494, 503 (1977)).

58. *Id.* at 191-92.

59. 491 U.S. 110 (1989) (plurality opinion).

the father of a child born to a married woman was the woman's husband.[60] Justice Scalia's plurality opinion began by quoting Justice White's reasons for being "extremely reluctant to breathe . . . further substantive content into the Due Process Clause."[61] To "limit and guide interpretation of the Clause," Justice Scalia insisted that a liberty interest be "rooted in history and tradition."[62] No alternative test was offered, and unlike in *Poe*, the focus of the Court's review of tradition was historical.[63] Moreover, writing for himself and the Chief Justice, Justice Scalia explicitly adopted Justice White's strategy of defining the relevant tradition narrowly.[64] Justice Scalia opined that the appropriate inquiry is

> [T]o the most specific level at which a relevant tradition protecting, or denying protection to, the asserted right can be identified. If, for example, there were no societal tradition, either way, regarding the rights of the natural father of a child adulterously conceived, we would have to consult and (if possible) reason from, the traditions regarding natural fathers in general.[65]

In *Washington v. Glucksberg*,[66] Chief Justice Rehnquist, writing for a majority, upheld the state of Washington's ban on assisted suicide and placed his own limiting gloss on the test for fundamental rights.[67] He first reiterated the Court's reluctance to "expand the concept of substantive due process . . . ."[68] Although acknowledging that many of the rights and liberties previously recognized by the Court sounded in personal autonomy, [69] the Chief Justice refused to recognize any right to make all important, intimate, and personal decisions.[70] To determine if a fundamental right existed, the Chief Justice first required a "'careful description' of the asserted fundamental liberty interest."[71] By finding the "careful description" of the fundamental right asserted by reference to

---

60. *Id.* at 114-15.

61. *Id.* at 122 (quoting *Moore*, 431 U.S. at 544 (White, J., dissenting)).

62. *Id.* at 122-23.

63. *Id.* at 124-25.

64. *Id.* at 127 n.6.

65. *Id.* Justice Brennan, joined by Justices Marshall and Blackmun, vigorously dissented. He challenged Justice Scalia's reliance on tradition, his strictly historical perspective, and his definition of the relevant right at the most specific level of generalization. *Id.* at 136 (Brennan, J. dissenting). Justice White, joined by Justice Brennan, filed a separate dissent. *Id.* at 157 (White, J., dissenting).

66. 521 U.S. 702 (1997).

67. *Glucksberg*, 521 U.S. at 720.

68. *Id.* (quoting Collins v. City of Harker Heights, 503 U.S. 115, 125 (1992)). To justify this reluctance, the Chief Justice observed: "'[G]uideposts for responsible decisionmaking in this uncharted area are scarce and open-ended.' By extending constitutional protection to an asserted right or liberty interest, we, to a great extent, place the matter outside the arena of public debate and legislative action." *Id.* (quoting *Collins*, 503 U.S. at 125).

69. *Id.* ("In a long line of cases, we have held that, in addition to the specific freedoms protected by the Bill of Rights, the 'liberty' specially protected by the Due Process Clause includes the rights to marry, to have children, to direct the education and upbringing of one's children, to marital privacy, to use contraception, to bodily integrity, and to abortion." (citations omitted)).

70. *Id.* at 727-28.

71. *Id.* at 721 (quoting Reno v. Flores, 507 U.S. 292, 302 (1993)).

the statute being challenged,[72] the Chief Justice effectively garnered a majority for Scalia's previously unadopted "most specific level" of generalization rule, or something very close to it.[73] The Chief Justice further limited expansion of substantive due process rights by requiring the asserted right to be both "'deeply rooted in this Nation's history and tradition' *and* 'implicit in the concept of ordered liberty' such that 'neither liberty nor justice would exist if they were sacrificed . . . .'"[74] *Glucksberg*'s conjunctive test necessarily is more restrictive than *Bowers*' disjunctive test or *Michael H.*'s focus solely on tradition. *Glucksberg*'s conjunctive test necessarily is more restrictive than *Bowers*' disjunctive test or *Michael H.*'s focus solely on tradition. As in *Michael H.*, *Glucksberg*'s inquiry into relevant traditions was historical.[75]

Although not overruling prior cases establishing fundamental rights, the conservative majority, through *Bowers*, *Michael H.*, and *Glucksberg*, appeared to completely transmogrify fundamental rights/privacy doctrine. In effect, the Court was saying, "this much but not more." It was against this background that the Court decided *Lawrence v. Texas*.[76]

Lawrence not only overruled the Court's earlier decision in *Bowers*, but contained broad open-ended language reminiscent of the Court's earlier fundamental rights/privacy case law. Justice Kennedy, writing for the majority, began his opinion by stating, "Liberty protects the person from unwarranted government intrusions into a dwelling or other private places . . . . [It] presumes an autonomy of self that includes freedom of thought, belief, expression and certain intimate conduct."[77] He then specifically rejected Justice White's narrow framing of the relevant issue in *Bowers*.[78] According to Justice Kennedy, "[t]o say that the issue in *Bowers* was simply the right to engage in certain sexual conduct demeans the claim the individual put forward, just as it would demean a married couple were it to be said marriage is simply about the right to have sexual

---

72. *See id.* at 723.

73. *See supra* note 65 and accompanying text. Statutes prohibit distribution of contraceptives, abortion or interracial marriage; they do not make the decision whether to beget a child or to marry illegal or ban privacy or personal autonomy.

74. *Glucksburg*, 521 U.S. at 721 (quoting *Moore*, 431 U.S. at 503 (plurality opinion); *Palko*, 302 U.S. at 326 (emphasis added)).

75. *Id.* at 721.

76. 539 U.S. 558 (2003).

77. *Lawrence*, 539 U.S. at 562. Later in the opinion, Justice Kennedy quoted at length the broad description of liberty contained in the Court's opinion in *Planned Parenthood of Se. Pa. v. Casey*, 505 U.S. 833, 851 (1992):

> These matters [decisions relating to marriage, procreation, contraception, family relationships, child rearing and education], involving the most intimate and personal choices a person may make in a lifetime, choices central to personal dignity and autonomy, are central to the liberty protected by the Fourteenth Amendment. At the heart of liberty is the right to define one's own concept of existence, of meaning, of the universe, and of the mystery of human life. Beliefs about these matters could not define the attributes of personhood were they formed under compulsion of the State.

Lawrence, 539 U.S. at 571.

78. *Id.* at 566-67.

intercourse."[79] Justice Kennedy instead focused on whether the government could interfere with personal relationships between consenting adults.[80] In answering that question, Justice Kennedy derided the *Bowers* Court's exclusive reliance on history.[81] He opined, "[h]istory and tradition are the starting point but not in all cases the ending point of the substantive due process inquiry,"[82] and found the laws and practices of the past half-century of the most relevance.[83] Moreover, he defined recent history broadly as showing "an emerging awareness that liberty gives substantial protection to adult persons in deciding how to conduct their private lives in matters pertaining to sex."[84]

Although *Lawrence*'s broad definitions of the liberty interests involved and its limitations on the use of tradition seemingly are a resounding rejection of the conservative trilogy of *Bowers, Michael H.,* and *Glucksberg*, the case has not been so read by lower courts.[85] Many lower courts[86] refuse to view *Lawrence* as a fundamental rights case at

79. *Id.* at 567.

80. *Id.* at 567, 571.

81. *Id.* at 571-72.

82. *Id.* at 572 (quoting County of Sacramento v. Lewis, 523 U.S. 833, 857 (1998) (Kennedy, J., concurring)).

83. *Id.* at 571-72.

84. *Id.* at 572 (citing *Lewis*, 523 U.S. at 857 (Kennedy, J., concurring)).

85. *See* Abigail Alliance For Better Access to Developmental Drugs v. von Eschenbach, 445 F.3d 470, 477 n.8 (D.C. Cir. 2006) (citing cases); *see also infra* note 86.

86. *See, e.g.*, Muth v. Frank, 412 F.3d 808, 817-18 (7th Cir.), *cert. denied* 126 S.Ct. 575 (2005); Williams v. Att'y Gen. of Ala., 378 F.3d 1232, 1236 (11th Cir. 2004); Lofton v. Sec'y of the Dep't of Children & Family Servs., 358 F.3d 804, 815-16 (11th Cir 2004), *cert. denied* 543 U.S. 1081(2005); Loomis v. United States, 68 Fed. Cl. 503, 517-19 (2005); Hernandez v. Robles, 2006 WL 1835429 (N.Y. 2006) (Graffeo, J., concurring); Martin v. Ziherl, 607 S.E.2d 367, 370 (Va. 2005); State v. Clinkenbeard, 123 P.3d 872, 878 (Wash. Ct. App. 2005); *see also* Lawrence v. Texas, 539 U.S. 558, 586 (2003) (Scalia, J., dissenting). These courts appear to give a very cramped interpretation of *Lawrence*. In addition to the broad language quoted in the text, the opinion states the issue as "whether the petitioners were free as adults to engage in the private conduct in the exercise of their liberty under the Due Process Clause . . . ." *Lawrence*, 539 U.S. at 564. To answer that question, the Court reviews many of its earlier fundamental rights cases, *id.* at 564-66 (discussing *Griswold, Eisenstadt* and *Roe*), not cases decided under a rational basis standard. Additionally, the Court quotes the *Bowers* Court's statement of the issue as "whether the Federal Constitution confers a *fundamental right* upon homosexuals to engage in sodomy . . . . " *Id.* at 566 (emphasis added). The *Lawrence* Court found that statement failed "to appreciate the extent of the liberty at stake" and overruled the Bowers decision. *Id.* at 567. In overruling *Bowers,* the Court focused on the *Bowers*' historical review, *id.* at 567-73, and referred to the subsequent broad language of the Court in *Casey* as casting doubt on *Bowers*. *Id.* at 573-74. That analysis implies that the *Bowers* Court erred by failing to find a right entitled to heightened scrutiny, rather than by overvaluing the justification offered by the State. That implication is reinforced by the Court's endorsement of Justice Stevens' dissent in *Bowers*. *Id.* at 577-78. In Stevens' view, the Court's fundamental rights case law precluded criminalization of sodomy as to all citizens. *Bowers*, 478 U.S. at 216-18 (Stevens, J., dissenting). The *Lawrence* Court's conclusion that there was "no legitimate state interest which [could] justify its intrusion into the personal and private life of the individual," *Lawrence*, 539 U.S. at 578, is not necessarily inconsistent with heightened review. First, the Court might have meant that because there was no legitimate justification, there was no need to inquire if the State's justification was compelling or outweighed the appellants' liberty interest. Second, the Court might have meant that although society's interest in morality is legitimate, there is no sufficient interest to justify intrusion into the personal and private life of the individual, an interest subject to heightened protection. *See* Dale Carpenter, *Is Lawrence Libertarian?*, 88 MINN. L. REV. 1140, 1157 (2004). Finally, the phrase could have been used in the same sense in which Justice Stevens used it in his *Bowers* dissent, an

all because the Court does not speak of creating a fundamental right and is viewed as applying a rational basis test based upon its conclusion that, "[t]he Texas statute furthers *no legitimate state interest* which can justify its intrusion into the personal and private life of the individual."[87] Because the conservative approach appears to remain dominant, the next section will highlight its shortcomings.

B. Problems with "History and Tradition" as the Basis for Defining Fundamental Rights

The primary advantage of the "history and tradition" test is its greater objectivity, binding judicial discretion so that the courts do not interfere with the democratic process.[88] Actual application of the test, however, has proven that the greater objectivity is more theoretical than real. First, the definition of the right being asserted will often determine the outcome. For example, is there a tradition of government non-interference with private intimate relations between consenting adults or a tradition supporting sodomy?[89] Second, even where there is agreement concerning the right involved, historical research often will be disputed.[90] There can be disagreement about the relevant time as well as the relevant sources.[91] Lastly, traditions are often conflicting. For example, there is a tradition outlawing adultery or sodomy, but there is also

---

analysis the Court specifically found should be controlling. *Lawrence*, 539 U.S. at 578. In *Bowers*, after finding that prior cases precluded application of the sodomy statute to the public generally, Justice Stevens analyzed whether the State could justify selective enforcement of the law. *Bowers*, 478 U.S. at 218-20 (Stevens, J., dissenting). Justice Steven found no "legitimate interest" for doing so. *Id.*

This author speculates that the Court intended heightened review, but was afraid to say so explicitly. The Supreme Court has stated that classifications that burden fundamental rights are subject to heightened review. *See, e.g.*, Romer v. Evans, 517 U.S. 620, 631 (1996); Zablocki v. Redhail, 434 U.S. 374, 388 (1978). By protecting homosexual conduct, the Court may have feared they effectively would have made homosexuals a protected class. The Court clearly was not prepared to address the consequences of such a holding, going out of its way to clarify that its decision did not address the constitutionality of state laws prohibiting gay marriages. *See, e.g.*, *Lawrence*, 539 U.S. at 567, 578; *id.* at 585 (O'Connor, J., concurring).

87. *Lawrence*, 539 U.S. at 578 (emphasis added).

88. *See* Cass R. Sunstein, *The Right to Marry*, 26 CARDOZO L. REV. 2081, 2105-06 (2005); Crump, *supra* note 4, at 863.

89. *Compare Lawrence*, 539 U.S. 558 (2003) *with Bowers*, 478 U.S. 186, *rehearing denied* 478 U.S. 1039 (1986), *overruled in* Lawrence v. Texas, 539 U.S. 558 (2003). Justice Scalia's lowest level of specificity test is designed to overcome this shortcoming. However, Justice Scalia's test is itself malleable. Results vary depending on how factually detailed one makes the statement of the issue and which facts are excluded when moving to the next level of abstraction. *See* Laurence H. Tribe & Michael C. Dorf, *Levels of Generality in the Definition of Rights*, 57 U. CHI. L. REV. 1057, 1092-93 (1990). More fundamentally, the lowest level of specificity test is inconsistent with Supreme Court precedent. *See, e.g.*, Loving v. Virginia, 388 U.S. 1 (1967) (interracial marriage); Griswold v. Connecticut, 381 U.S. 479 (1965) (contraceptives); Roe v. Wade, 410 U.S. 113 (1973) (abortion).

90. *See* Rebecca L. Brown, *Tradition and Insight*, 103 YALE L. J. 177, 202-03 (1993); Tribe & Dorf, *supra* note 89, at 1087-89; ELY, *supra* note 3, at 60, 103.

91. *See* Marybeth Herald, *A Bedroom of One's Own: Morality and Sexual Privacy After* Lawrence v. Texas, 16 YALE J. L. & FEMINISM 1, 11 (2004). This is particularly true when issues involve medical and technological advances unanticipated by earlier generations. *Id.* at 12; *see also* Crump, *supra* note 4, at 862-63.

a tradition of non-enforcement of such laws; there is a tradition respecting equality, but also a tradition of subjecting various groups to a variety of forms of ostracism or prejudice.[92] Not surprisingly, Justices often will resolve these conflicts based on what best furthers their own predilections.[93]

The fundamental problem with the "history and tradition" test, however, is not its failure to achieve increased objectivity, but its inconsistency with the structure of the Constitution. As Professor Ely observed, tradition's "overtly backward-looking character highlights its undemocratic nature: it is hard to square with the theory of our government the proposition that yesterday's majority . . . should control today's."[94] This is especially true when yesterday's majority was composed primarily, if not solely, of white, wealthy, straight men.[95] Moreover, if the only rights receiving protection were those historically and currently valued by society, there would be no need for the fundamental rights doctrine,[96] at least other than to provide protection from the maverick state. This would ignore the Court's Constitutional role as a protector of minority interests, and is inconsistent with the Fourteenth Amendment's anti-majoritarian purposes.[97]

This article does not suggest that tradition, if one can be agreed upon, is irrelevant. Tradition, representing the combined wisdom of generations, often will have much to recommend it.[98] Moreover, if the Court breaks too radically from ongoing traditions, it risks institutional credibility.[99] What is objectionable is the blind adherence to tradition. Although some traditions are worthy, others reflect ignorance, prejudice, or inequalities in power.[100] One should learn from history, not mechanically follow it. It is for this reason that this article recommends adoption of a right to privacy defined more specifically and considers tradition only as part of its sliding scale review, and then only if the circumstances upon which the tradition was based have not changed.[101]

---

92. *See* Brown, *supra* note 90 at 203; ELY, *supra* note 3, at 61.

93. *See* Wolf, *supra* note 11, at 126-128; Brown, *supra* note 90, at 210-11.

94. ELY, *supra* note 3, at 62. Professor Ely further argues that if the Framers wanted to freeze tradition, they would have wrote out the tradition rather than seek to protect it through open-ended language. *Id.*

95. Wolf, *supra* note 93, at 126-27.

96. *See Michael H.*, 491 U.S. at 140-41 (Brennan, J., dissenting); Wolf, *supra* note 11, at 115.

97. *See* THE FEDERALIST NO. 10 (James Madison); THE FEDERALIST NO. 78 (Alexander Hamilton); ELY, *supra* note 3, at 62; Niles, *supra* note 14, at 118; Crump, *supra* note 4, at 861.

98. *See* Sunstein, *supra* note 88, at 2106.

99. *See* Poe v. Ullman, 367 U.S. 497, 542 (1961) (Harlan, J., dissenting).

100. *See* Sunstein, *supra* note 88, at 2106; Niles, *supra* note 14, at 141. The fact that the government has a longstanding tradition of violating individual rights does not make it legitimate. *See id.*, Tribe & Dorf, *supra* note 89, at 1088.

101. *See infra* notes 207-12 and accompanying text.

C. The Appropriate Standard of Review

Once a fundamental right is found, the Court repeatedly has stated that the government cannot infringe upon that right unless the infringement is narrowly tailored to serve a compelling state interest.[102] In the absence of a fundamental right, the government can justify its action by demonstrating a mere rational basis for its conduct.[103] Because "the review standard for ordinary liberties is so deferential, and the standard for preferred liberties so rigid,"[104] outcomes often are ordained by the designation of rights as fundamental or not.[105]

Despite the clarity of the Court's doctrine, there is much in the Court's language and practice that suggests balancing of interests is appropriate. Indeed, Justice Harlan's influential opinion in *Poe*[106] seemingly required a balancing of interests. He opined, "'liberty' is not a series of isolated points," but a "rational continuum."[107] It includes not only freedom from arbitrary restraints, but also "recognizes, what a reasonable and sensitive judgment must, that certain interests require particularly careful scrutiny of the state needs asserted to justify their abridgment."[108] Justice Harlan further argued that due process, through the course of the Court's decisions, "represent[s] the balance which our Nation, built upon postulates of respect for the liberty of the individual, has struck between that liberty and the demands of an organized society."[109]

One technique the Court has employed to balance interests is to impose a "substantial" or "undue" burden threshold for determining whether a fundamental right has been infringed. The paradigm example is Justice O'Connor's plurality opinion in *Planned Parenthood of Southeastern Pennsylvania v. Casey*.[110] In *Casey*, Justice O'Connor reaffirmed the fundamental right to choose an abortion.[111] However, she opined that,

---

102. *See supra* note 6.

103. *See supra* note 7.

104. Ira Lupu, *Untangling the Strands of the Fourteenth Amendment*, 77 MICH. L. REV. 981, 1030 (1979).

105. *Id.* at 1029-30; *see also* ERWIN CHEMERINSKY, CONSTITUTIONAL LAW: PRINCIPLES AND POLICIES 417 (1997). The two-tiered approach appears to be particularly popular among the conservative justices. That approach avoids the always-feared subjective decision-making required by a balancing of interests. It also discourages a court, realizing that the compelling interest test makes most government regulation improper, from finding a right fundamental in the first instance.

106. 367 U.S. 497 (1961).

107. *Id.* at 543 (Harlan, J., dissenting).

108. *Id.*; *see also* Planned Parenthood of Se. Pa. v. Casey, 505 U.S. 833, 848 (1992) (plurality opinion); Moore v. City of E. Cleveland, 431 U.S. 494, 502 (1977) (plurality opinion); Roe v. Wade, 410 U.S. 113, 169 (1973).

109. *Poe*, 367 U.S. at 542; *see also Glucksberg*, 521 U.S. at 765; *Casey*, 505 U.S. at 850; *Moore*, 431 U.S. at 501 (plurality opinion).

110. 505 U.S. 833 (1992); *see also Glucksberg*, 521 U.S. at 767 n.8 (Souter, J., concurring); *Youngberg*, 457 U.S. at 320; Zablocki v. Redhail, 434 U.S. 374, 386-87 (1978).

111. *Casey*, 505 U.S. at 846.

> [t]he fact that a law which serves a valid purpose . . . has the incidental effect of making it more difficult or more expensive to procure an abortion cannot be enough to invalidate it. Only where the state regulation imposes an undue burden on a woman's ability to make this decision does the power of the State reach into the heart of the liberty protected by the Due Process Clause.[112]

The Court necessarily also balances interests when applying an intermediate standard of review. In *Carey v. Population Services International*,[113] although invalidating a restriction on the distribution of contraceptives to persons under sixteen, the Court explicitly indicated that a lower level of scrutiny was appropriate when minors claimed an infringement of their right to privacy.[114] The Court also has found intermediate review appropriate in cases deciding when the government may involuntarily administer anti-psychotic drugs to a mentally ill patient.[115] In such cases, the court must find (1) the government interest *important*; (2) "involuntary medication will *significantly further* those concomitant state interests"; (3) "involuntary medication is *necessary* to further those interests"; and (4) "administration of the drugs is *medically appropriate*."[116] Similarly, in *Moore v. City of East Cleveland*,[117] the Court, invalidating a local regulation limiting who could live together, stated, "[w]hen the government intrudes on choices concerning family living arrangements, this Court must examine carefully the importance of the governmental interests advanced and the extent to which they are served by the challenged regulation."[118]

The Court even has explicitly balanced interests. In *Youngberg v. Romeo*,[119] the Court found that an involuntarily committed mental patient had a liberty interest in minimally adequate training.[120] The Court defined adequacy "as that training which is reasonable in light of identifiable liberty interests and the circumstances of the case."[121] *Youngberg* was cited in *Cruzan*.[122] The Court in that case stated, "determining that a person has a 'liberty interest' under the Due Process Clause does not end the inquiry, 'whether respondent's constitutional rights have been vio-

112. *Id.* at 874 (citations omitted). Applying this standard, Justice O'Connor found that although the spousal notification provision was invalid, the 24-hour waiting period, informed consent and reporting and record-keeping requirements were not. *Id.* at 881-91.

113. 431 U.S. 678 (1977).

114. *Id.* at 693 n.15 (The Court reasoned that lesser scrutiny was appropriate because the right of privacy implicated "'the interest in independence in making certain kinds of decisions', and the law has generally regarded minors as having a lesser capability of making important decisions." (quoting Whalen v. Roe, 429 U.S. 589, 599-600 (1977))).

115. *See* Sell v. United States, 539 U.S. 166, 180-81 (2003).

116. *Id.*

117. 431 U.S. 494 (1977) (plurality opinion).

118. *Id.* at 499 (citing Poe v. Ullman, 367 U.S. 497, 554 (1961) (Harlan, J., dissenting).

119. 457 U.S. 307 (1982).

120. *Youngberg*, 457 U.S. at 318-19.

121. *Id.* at 319 n.25.

122. 497 U.S. 261, 279 (1990).

lated must be determined by balancing [the] liberty interests against the relevant state interests."[123]

Perhaps it was not coincidence that in the two most recent Supreme Court substantive due process decisions, *Lawrence*[124] and *Troxel v. Granville*,[125] the Court didn't even state what standard of review it was applying.[126] The Court finally may have begun to recognize the limits of tiered analysis and acknowledge its frequent practice of balancing interests. Nonetheless, lower courts have consistently cited to traditional doctrine.[127] They require the government to demonstrate that an infringement of a fundamental right is narrowly tailored to achieve a compelling government interest.[128] It is for this reason that tiered analysis is critiqued below.

D. Problems with the Tiered Review

The primary problem with tiered review is its inflexibility.[129] As Justice Harlan recognized, liberty is a "rational continuum."[130] It makes little sense to assume that unless a regulation must be narrowly tailored to a compelling government interest it is valid except if irrational, no matter how overbroad or how much it infringes an individual's liberty. The problem is particularly acute when tiered review is combined with the Court's narrow definition of fundamental rights.[131] For example, consider a law that makes it illegal for overweight people to eat pies, cake, ice cream, bread, potatoes, or pasta. The Court would have a difficult time identifying a right to be overweight or to eat those foods as fundamental under its current jurisprudence and the government has a rational basis for the regulation—to reduce the health risks attendant to

---

123. *Cruzan*, 497 U.S. at 279.

124. 539 U.S. 558 (2003).

125. 530 U.S. 57 (2000).

126. *See* Note, *Assessing the Viability of a Substantive Due Process Right to In Vitro Fertilization*, 118 HARV. L. REV. 2792, 2807-08 (2005).

127. *See, e.g.*, Abigail Alliance for Better Access to Developmental Drugs v. von Eschenbach, 445 F.3d 470 (D.C. Cir. 2006); Williams v. Att'y Gen. of Ala., 378 F.3d 1232, 1240 (11th Cir. 2004); Littlefield v. Forney Indep. Sch. Dist., 268 F.3d 275, 288 (5th Cir. 2001); Hodgkins v. Peterson, No. 1:04-CV-569-JDT-TAB, 2004 WL 1854194, at *7-*8 (S.D. Ind. Jul. 23, 2004); Loomis v. United States, 68 Fed. Cl. 503, 517 (2005); State v. Clinkenbeard, 123 P.3d 872, 878-79 (Wash. Ct. App. 2005); State v. J.P., 907 So.2d 1101, 1109-1110 (Fla. 2004); State v. Saunders, 381 A.2d 333, 341 (N.J. 1977). *But cf.* Doe v. Heck, 327 F.3d 492, 519 (7th Cir. 2003) (citation omitted) (recognizing that "it is well established that when a fundamental constitutional right is at stake, courts are to employ the exacting strict scrutiny test," but questioning if *Troxel v. Granville* means courts are to apply some other standard of heightened scrutiny to claims alleging violation of the fundamental right to familial relations).

128. *See supra* note 127.

129. *Cf.* Dennis v. United States, 341 U.S. 494, 524 (1951) (Justice Frankfurter stated that it is better to decide by "candid and informed weighing of the competing interests . . . than by announcing dogmas too inflexible for the non-Euclidian problems to be solved").

130. *See* Poe v. Ullman, 367 U.S. 497, 543 (1961) (Harlan, J., dissenting).

131. Of course, the Court has tried to narrowly define rights, in part, because under the two-tiered approach, government regulation has little chance of surviving when the Court finds a right to be fundamental. *See supra* note 102-04 and accompanying text.

excess weight.[132] Yet, the statute seems to be drastically overbroad and infringe significant liberty interests.[133] Perhaps, recognition of the difficulties created by tiered review's inflexibility explains the Supreme Court's sometime disingenuous application of its enunciated standard.[134] Nonetheless, it is better to candidly acknowledge the weighing of competing interests. Only then can we "actually increase the possibility of accountability and ultimately hope to reduce the power of idiosyncratic decisionmaking."[135]

The fact that the Court's application of tiered review often is disingenuous is reason enough to reject it. However, the Court's repeated refusal to acknowledge the realities of balancing and overrule tiered review also has created practical problems in the lower courts. First, fear that a valid government interest might not survive the tightest fit of strict scrutiny makes courts disinclined to find a liberty interest in the first instance.[136] Indeed, this fear, combined with Supreme Court warnings about its reluctance to create new fundamental rights,[137] virtually paralyzes courts from recognizing rights by analogy.[138] Consequently, an individual effectively can challenge a regulation infringing a liberty interest not specifically recognized by the Court only if she is willing to

---

132. It might be argued that paternalistic concerns should not be considered legitimate even under rational basis review. However, the government could still justify the regulation as rational by claiming that the increased risk of the overweight person's suffering sudden heart failure endangers other drivers.

133. A possible response to this hypothetical is that it is just that—a hypothetical; that we can trust legislators not to enact such a silly law. One reply is that the framers of the Constitution established a tripartite system of government precisely because they didn't trust legislators always to act wisely. *See* THE FEDERALIST NO. 78, at 476-77 (Alexander Hamilton) (Bantam Ed. 2003); *Williams*, 378 F.3d at 1240 n.11. However, the problems suggested by the hypothetical are not limited to silly laws. Consider a law banning the use of drugs. If the liberty interest at stake is defined by reference to the statute, as suggested by Chief Justice Rehnquist in *Glucksberg*, no fundamental right is involved. 521 U.S. 702, 720 (1997). The government has an obvious legitimate basis for the regulation, for example to reduce the incidence of driving while impaired. However, the statute's application to a bed-ridden terminal cancer patient for whom the drug is the best or only form of relief from excruciating pain seems to infringe significant liberty interests. *Cf.* Gonzales v. Raich, 545 U.S. 1 (2005) (holding that application of Controlled Substance Act to users of marijuana for medical purposes, despite state law allowing such use, did not violate Commerce Clause); *see also* Note, *Last Resorts and Fundamental Rights: The Substantive Due Process Implications of Prohibitions on Medical Marijuana*, 118 HARV. L. REV. 1985, 1985-87 (2005); *infra* notes 263-70 and accompanying text.

134. *See supra* notes 106-126 and accompanying text.

135. Brown, *supra* note 90, at 215.

136. *See, e.g.*, *Williams*, 378 F.3d at 1240 (declining to find a right to sexual privacy by consenting adults because such a right would subject activities such as incest, prostitution, and obscenity to strict scrutiny, something the court was not prepared to do). A possible problem created by this reluctance to subject a regulation to strict scrutiny is what might be called a tyranny of labels. A court fearing strict scrutiny in one context might reject an asserted fundamental right that it might wish to recognize in another. For example, in *Kelley v. Johnson*, 425 U.S. 238 (1976), the Court, upholding a police regulation on personal grooming, rejected the asserted liberty interest in choice of one's hair length. One might question whether a law requiring all citizens to shave their heads, which might be justified as reducing the incidence of head lice (even in the absence of an epidemic, or even an increase in frequency), should be upheld merely because it is rational.

137. *See supra* note 50.

138. *See, e.g.*, *von Eschenbach*, 445 F.3d at 487 (Griffith, J., dissenting).

incur the costs of litigation through an appeal to the Supreme Court. Additionally, lower courts have repeatedly refused to recognize a right to private consensual sexual relations among adults, a liberty interest seemingly found by the Court in *Lawrence*,[139] because the Court did not specify it was applying strict scrutiny.[140] If the Court continues its trend of not identifying the standard of review it is applying, the law in the lower courts will only be further distorted.

It is time that Supreme Court policy and lower court practice be harmonized as to both the identification of fundamental rights and the standard of review to apply to such rights. To do this, there needs to be a better conceptualization of the area and a more honest statement of the standard of review actually applied. The proposals in the following section seek to do just that.

II. A PROPOSED APPROACH TO PRIVACY RIGHTS UNDER SUBSTANTIVE DUE PROCESS

A. The Conceptualization of Privacy Rights

The conceptualization proposed by this article is heavily influenced by the writings of John Locke[141] and John Stuart Mill.[142] As suggested earlier, the beginning premise is that the Constitution should be viewed as a pact between individuals and the government to forgo certain rights that are necessary to further society's interests, but with a reservation of rights in certain private areas where the government does not belong.[143] Broadly speaking,[144] the government, to further society's interests, has a right to regulate the individual's interaction with the larger world. It should be presumed, however, that the government does not properly control the world of the self, defined as one's thoughts, feelings, bodily integrity, and private intimate relationships[145] with consenting adults.[146]

---

139. *See supra* note 86; *see also* Carpenter, *supra* note 86, at 1155; Paul M. Secunda, *Lawrence's Quintessential Millian Moment and Its Impact on the Doctrine of Unconstitutional Conditions*, 50 VILL L. REV. 117 (2005); Herald, *supra* note 91, at 38 (all finding such a right).

140. *See, e.g.*, Muth v. Frank, 412 F.3d 808, 817-18 (7th Cir.) *cert. denied* 126 S.Ct. 575 (2005); Lofton v. Sec'y of the Dep't of Children & Family Servs., 358 F.3d 804, 815-16 (11th Cir. 2004) *cert. denied* 543 U.S. 1081 (2005); Loomis v. United States, 68 Fed. Cl. 503, 517-19 (2005); Stanhardt v. Arizona, 77 P.3d 451, 457 (Ariz. Ct. App. 2003); State v. Clinkenbeard, 123 P.3d 872, 878 (Wash. Ct. App. 2005).

141. *See* LOCKE, *supra* note 14.

142. *See* JOHN STUART MILL, ON LIBERTY (Legal Classics Library Ed. 1992) (1859).

143. *See supra* note 14 and accompanying text.

144. Precise definition of the areas the government does not belong is not possible without some recourse to intuition. *See* Tom Gerety, *Redefining Privacy*, 12 HARV. CIV. RTS.-CIV. LIBERTIES L. REV. 233, 236-42 (1977). Yet, as Justice O'Connor observed in *Casey*, "[l]iberty must not be extinguished for want of a line that is clear." Planned Parenthood of Se. Pa. v. Casey, 505 U.S. 833, 869 (1992) (plurality opinion).

145. What constitutes an "intimate" relationship cannot be defined precisely. To decide whether a group is sufficiently personal to warrant protection under the right to intimate association, the Court considers "'factors such as size, purpose, selectivity, and whether others are excluded from critical aspects of the relationship.'" Boy Scouts of Am. v. Dale, 530 U.S. 640, 698 n.26 (2000) (quoting Bd. of Dirs. of Rotary Int'l v. Rotary Club of Duarte, 481 U.S. 537, 546 (1987)).

In effect, those areas are viewed as controlled by private governments, that of the individual and the consenting adults.[147]

The Supreme Court on many occasions has endorsed this "right to be let alone" or "area the government does not belong."[148] As early as 1928, Justice Brandeis stated:

> The makers of our Constitution . . . recognized the significance of man's spiritual nature, of his feelings and of his intellect. They knew that only a part of the pain, pleasure and satisfactions of life are to be found in material things. They sought to protect Americans in their beliefs, their thoughts, their emotions and their sensations. They conferred, as against the Government, the right to be let alone—the most comprehensive of rights and the right most valued by civilized men.[149]

---

146. It may seem anomalous to include intimate relations, which are necessarily dependant on another, as part of the world of the self. However, intimacy is necessary for the full development of the self. Intimacy requires the ability to care and be cared for and "has a great deal to do with the formation and shaping of an individual's sense of his own identity." Kenneth L. Karst, *The Freedom of Intimate Association*, 89 YALE L.J. 624, 633-36 (1980). The Supreme Court itself has observed that Constitutional protection for such relationships is warranted and "reflects the realization that individuals draw much of their emotional enrichment from close ties with others." Roberts v. Jaycees, 468 U.S. 609, 619 (1984); *see also* Lawrence v. Texas, 539 U.S. 558, 578 (2003) *Casey*, 505 U.S. at 851. Almost by definition, private intimate relations exclude the outside world and hence should exclude the government.

Professor Ely has questioned why food, housing or jobs are not considered fundamental rights, suggesting that the Court only favors "upper middle class" rights. *See* ELY, *supra* note 3, at 59. Certainly, it might be argued that jobs, housing, or food, in some sense, are necessary to the development of the self. They are at least fundamental to a person's existence. While these interests are important, the right to privacy should not cover them. The purpose of the Bill of Rights, including the Due Process Clause, was to protect the individual from government excesses, *see, e.g.*, Michael C. Dorf, *Incidental Burdens on Fundamental Rights*, 109 HARV. L. REV. 1175, 1187-88 (1996), not to define a minimal level of subsistence to be enforced by the judiciary. Moreover, privacy is concerned with preventing the government from infringing on the individual's prerogative. The proper decision-maker concerning government benefits, obviously, must be the government. *See infra* notes 146-148 and accompanying text. By contrast, if government regulations allocated jobs universally or defined what foods must be eaten, the government would need to justify this infringement of individual prerogative. Of course, the government can choose to, and to some extent does, provide for minimum subsistence. However, that is more properly decided by consensus through the legislature, the branch of government that controls the purse, rather than by judicial fiat. *See* Dorf, *supra,* at 1235; Crump, *supra* note 4, at 903.

147. *Cf.* MILL, *supra* note 142, at 22 ("Over himself, over his body and mind, the individual is sovereign.").

148. *See, e.g.*, *Lawrence*, 539 U.S. at 578; *Casey*, 505 U.S. at 847; Carey v. Population Servs. Int'l., 431 U.S. 678, 685 (1977); Eisenstadt v. Baird, 405 U.S. 438, 453 n.10 (1972); Stanley v. Georgia, 394 U.S. 557, 564 (1969); Griswold v. Connecticut, 381 U.S. 479, 494 (1965).

149. Olmstead v. United States, 277 U.S. 438, 478 (1928) (Brandeis J., dissenting), *overruled by* Katz v. United States, 389 U.S. 347 (1967) and Berger v. New York, 388 U.S. 41 (1967). It is logical to ask if there is textual support for the asserted right to be let alone. The most logical source is the Ninth Amendment. U.S. CONST. amend. IX ("The enumeration in the Constitution of certain rights shall not be construed to deny or disparage others retained by the people."). However, the Supreme Court has shown little inclination to use that Amendment and the weight of legislative history suggests the Amendment was designed to protect state, not individual, rights. *See* Kurt T. Lash, *The Lost Original Meaning of the Ninth Amendment*, 83 TEX L. REV. 331 (2004); Kurt T. Lash, *The Lost Jurisprudence of the Ninth Amendment*, 83 TEX. L. REV. 597, 598-600 (2005); Russell L. Caplan, *The History and Meaning of the Ninth Amendment*, 69 VA. L. REV. 223, 227-28 (1983). The right to be let alone also can be supported by the overall structure of the Constitution

There are many reasons to protect this realm of private decision-making. First, the individual has the most information concerning their personal preferences. Second, self-definition is an end in itself and is necessary for complete development and happiness.[150] Third, the freedom to develop oneself can lead to genius that benefits society.[151] Conversely, the inability to control one's life can lead to unrest and divisiveness that undermines societal stability. Fourth, government control of private decisions risks the tyranny of the majority feared by Madison and the other Framers.[152] For that matter, the whole concept of deliberative democracy is meaningless if there is no individuality because the self is formed by the state.[153]

Although there is no bright line between the "public" and "private" worlds, analogy may be made to the concept of boundaries in psychology. Just as one knows that a mother should not order food for an adult child in a restaurant, the government should not decide in which private sexual acts consenting adults can engage. The goal in both situations is to allow the individual to become fully actualized. Although neither the mother nor the government can or are under an obligation to make the individual happy, they should allow the necessary condition for happiness, proper ego boundaries,[154] to develop.

---

and its concern for individual rights. Although the "right to be let alone" may not be specifically enumerated, the Constitution is a short document that can't be expected to have specified every right protected. *See* Poe v. Ullman, 367 U.S. 497, 540 (1961) (Harlan, J., dissenting); BARNETT, *supra* note 14, at 259; J. Braxton Craven, Jr., *Personhood: The Right to be Let Alone*, 1976 DUKE L.J. 699, 704 n.35. Nor does the enumeration of some rights necessarily preclude protection for others. For example, if someone you know to be a "clean freak" lends you their car and tells you not to have any food or drink in the car, not to write with pencil, pen or marker while in the car, and not to take any non-toilet trained babies in the car, you should know that taking your dog in the car and leaving its poop on the passenger seat is beyond the bounds of your authorization, despite its lack of specification. A full discussion of the textual and historical support for fundamental rights is beyond the scope of this article. It is enough to observe that the Supreme Court repeatedly has recognized that the Constitution does protect certain unenumerated fundamental rights. *See supra* note 4.

150. *See, e.g.*, MILL, *supra* note 142, at 102, 106; RUEVEN BAR-LEVAV, THINKING IN THE SHADOW OF FEELINGS 193, 330 (1988).

151. *See* MILL, *supra* note 142, at 117-18.

152. *See supra* note 140.

153. *See* James E. Fleming, *Securing Deliberate Autonomy*, 48 STAN. L. REV. 1, 23 (1995); Note, *Last Resorts and Fundamental Rights: The Substantive Due Process Implications of Prohibitions on Medical Marijuana*, 118 HARV. L. REV. 1985, 1987 (2005).

154. Ego boundaries, as psychoanalytic thinkers define them, are the boundaries conceived to exist between the self and the outside world. *See* Matthew Maibaum, *A Lewinian Taxonomy of Psychiatric Disorders*, THE INT'L SOC'Y FOR GESTALT THEORY & ITS APPLICATIONS, 2001, http://gestalttheory.net/archive/maibaum.html; BAR-LEVAV, *supra* note 150, at 330; CHARLES BRENNER, AN ELEMENTARY TEXTBOOK OF PSYCHOANALYSIS 59 (rev. ed. Anchor Books 1974) (1955). It would probably be more accurate to speak of self-boundaries than of ego boundaries, but the latter phrase is the psychoanalytic term or art. *See* Maibaum, *supra*. Strong boundaries are a prerequisite for a fully developed self and allow the individual to achieve true intimacy and happiness. *See* BAR-LEVAV, *supra* note 150, at 150, 158-59, 193, 330-31; JOHN BRADSHAW, BRADSHAW ON: THE FAMILY 43, 47, 55 (1988).

B. Non-Private Areas

Admittedly, the phrase "area where the government doesn't belong" is amorphous and specifying the area as including one's thoughts, feelings, bodily integrity and one's intimate relationships doesn't fully remedy this problem. Perhaps what best clarifies what should be considered private is a description of what is not private. For the reasons explained below, the government should be presumed to act where it belongs[155] when it provides government benefits, regulates commercial activity and activity in public areas,[156] and seeks to prevent harm to others.[157]

1. Government Benefits

There can be no privacy right to government benefits, under this article's conceptualization because rights dependent on the government, by definition, can't be an area the government does not belong. A denial of benefits also does not interfere with fundamental rights in the same way as regulation of such rights—individuals can do as they wish and they are in no worse a position than if the government did not exist. For similar reasons, there is no obligation on behalf of the government to publicly fund fundamental rights.[158] Although the government does not have the obligation to support private choices, it should not be able to deny benefits to which an individual would otherwise be entitled, absent a rational relation to the purposes of the benefit.[159] For example, although the government does not have an obligation to fund abortions, it would be improper for it to deny food stamps to persons who have had an abortion.

---

155. The presumption may be rebutted when the purpose of the government action is to infringe individual rights. Improper purpose should be presumed if the government lacks a rational basis for its actions. *See* State v. Clinkenbeard, 123 P.3d 872, 878 (Wash. Ct. App. 2005) (discussing rational basis review for a challenged statute).

156. The government also "belongs" in the area of foreign affairs. This is specifically provided for in the Constitution and necessarily involves the government. U.S. CONST. art. 1, § 8, cls. 3, 10, 11; U.S. CONST. art. 2, § 2, cl. 2. Apparently, this is so obvious that the Court has never had to address a fundamental rights challenge in this area. A proper role of the government also includes resolution of competing fundamental rights. *See* JOHN LOCKE, TWO TREATISES OF GOVERNMENT 324 (Peter Laslett ed., Cambridge University Press 1988) (1689).

157. Government action in an area it belongs often will affect rights in an area it does not belong. For example, a law that limits the amount that can be charged for an abortion, a commercial regulation, can impact the women's right to choose whether to have a child and infringe her right to bodily integrity. In such cases, this article recommends evaluation of the government action under a sliding scale analysis. *See infra* notes 237-54 and accompanying text.

158. *See, e.g.*, Rust v. Sullivan, 500 U.S. 173, 196-200 (1991); Youngberg v. Romeo, 457 U.S. 307, 317 (1982); Harris v. McRae, 448 U.S. 297, 314-17 (1980); Maher v. Roe, 432 U.S. 464, 469 (1977).

159. To this extent, substantive due process interacts with equal protection. *See* Zablocki v. Redhail, 434 U.S. 374, 395 (1978) (Stewart, J., concurring); Loving v. Virginia, 388 U.S. 1, 12 (1967); *cf.* Schware v. Bd. of Bar Examiners, 353 U.S. 232, 238-39 (1957) (finding that denial of admission to bar must have rational connection to applicant's fitness or qualifications to practice law).

A denial of benefits in such a case is tantamount to a fine and must be treated as a regulation of the fundamental right.[160]

2. Commercial Activity

There is no privacy right to commercial activity under this article's proposal. It is difficult to characterize commercial activity as private.[161] It is engaged in with others and does not by itself involve any sense of intimacy. More importantly, one of the primary reasons for the "more perfect union" was the need for the government to be able to regulate commercial transactions.[162] Periods of laissez-faire economics proved economic regulation necessary to prevent harm to others. Thus, regulation of commerce, specifically provided for in the Constitution,[163] must be treated as an area in which the government belongs.[164] A return to Lochnerism is not a risk of this article's proposal.[165]

3. Activity in Public Areas

Tautologically, activity in public areas is not private. As part of the Lockean pact, the government has the right to regulate the individual's interaction with the larger world.[166] Public areas represent the commons and demand regulation by the people rather than the individual. This would explain why public nudity can be prohibited and why environmental regulations are legitimate. The paradigm non-public area, of course, is the home.[167] The government's ownership interest also allows it to make rules on government owned property. For example, the government is not under any obligation to permit abortions to be conducted in public hospitals.[168]

---

160. See *infra* notes 177-207 and accompanying text for treatment of government regulations of fundamental rights.

161. *See* Paris Adult Theatre v. Slaton, 413 U.S. 49, 65 (1973); Mark C. Niles, *Ninth Amendment Adjudication: An Alternative to Substantive Due Process Analysis of Personal Autonomy Rights*, 48 UCLA L. REV. 85, 149 (2000); MILL, *supra* note 142, at 170.

162. *See* THE FEDERALIST NOS. 11-13 (Alexander Hamilton); Gonzales v. Raich, 545 U.S. 1, 16 (2005); Niles, *supra* note 161, at n.120.

163. U.S. CONST. art. I, § 8, cl. 3.

164. To offset an infringement of privacy rights, however, the government must have a rational basis to believe the marketplace needs regulation. *See Clinkenbeard*, 123 P.3d at 878. This article also would require the government's justification to be the actual, rather than a hypothesized, reason for the regulation. LAURENCE H. TRIBE, AMERICAN CONSTITUTIONAL LAW 1307 n.13 (2nd ed. 1988). Where the activity the government seeks to regulate has been ongoing, a court should require documentation for the need for the regulation. *See* Hodgkins v. Peterson, No. 1:04-CV-569-JDT-TAB, 2004 WL 1854194, at *8 (S.D. Ind. July 23, 2004).

165. Lochnerism, named after *Lochner v. New York*, 198 U.S. 45, 64 (1905), here refers to the Court's use of substantive due process to replace a state's reasonableness assessment with its own on matters of economic policy.

166. *See* Niles, *supra* note 161, at 111 & n.14.

167. *See* Lawrence v. Texas, 539 U.S. 558, 562 (2003).

168. *See* Webster v. Reprod. Health Servs., 492 U.S. 490, 507 (1989).

4. Harm to Others

The basic role of the government in the Lockean system is to prevent harm to others, both physical and economic.[169] Without the impartial magistrate that is the government, power rather than justice would determine rights.[170] Mill also recognized that the individual's natural rights ended where they caused harm to others.[171] This principle is hardly controversial and the Supreme Court has recognized that it is a government function to protect non-consenting parties from harm.[172] This is just part of the government's police powers.

However, the concept of harm to others needs to be refined, lest this exclusion eliminates all fundamental rights. Unless one is a hermit living in the woods, all actions can cause some harm to others. A person's choice to discontinue life-support affects all who know him or her. Knowledge that an individual engages in sodomy may offend or cause psychological harm to those with conservative sexual preferences. An important limiting principle was suggested by Hobbes—that one's rights as against another should be limited by what he would allow other men against himself.[173] Thus, finding a person's appearance or habits offensive to contemplate should not be considered harm to others because others would not want their appearance or habits subject to approval of the individual.[174] Another limiting principle, suggested by Mill, is that harm to others cannot be solely derivative of the harm to the individual.[175] This limitation is necessary because any decision one makes can harm oneself and therefore harm someone who cares about you. A third limitation is that harm to others generally must result from action, not omission.[176] This follows from common law principles and protects against laws such as one requiring donation of body parts to another.

---

169. *See* LOCKE, *supra* note 156, at 276 ("*Civil Government* is the proper Remedy for the inconveniences of the State of Nature, which must certainly be Great, where Men may be Judges in their own Case, since 'tis easily to be imagined, that he who was so unjust as to do his Brother an Injury, will scarce be so just as to condemn himself for it."); BARNETT, *supra* note 14, at 70-71.

170. *See* LOCKE, *supra* note 156, at 271-72.

171. *See* MILL, *supra* note 142, at 21-22, 140.

172. The government also has an interest in preventing harm to minors or incompetent persons. Such persons can be deemed incapable of a valid consent. *See Paris Adult Theatre*, 413 U.S. at 57.

173. *See* THOMAS HOBBES, LEVIATHAN 188, 190 (C.B. McPherson ed., Penguin Books 1968) (1651).

174. *See* TRIBE, *supra* note 164, at 1409. Professor Tribe finds this result completely analogous to recognized First Amendment principles. *Id.* "The expression of ideas or emotions cannot be shut off to protect unwilling viewers or hearers without 'a showing that substantial privacy interests are being invaded in an essentially intolerable manner,' since any 'broader view . . . would effectively empower a majority to silence dissidents simply as a matter of personal predilections.'" *Id.* (citations omitted).

175. *See* MILL, *supra* note 142, at 26.

176. *See id.* at 24-25.

C. Standard of Review

This article recommends a multi-factored sliding scale standard of review in which the balance is presumptively predetermined in two categories of cases. Specifically, when the government regulates in an area that it belongs, and there is no significant effect on privacy rights, the government regulation is presumed valid unless irrational. By contrast, when the government directly regulates in an area it does not belong, the regulation is presumed invalid unless it is narrowly tailored to achieve a compelling justification. This mirrors current rational basis and compelling interest review. More commonly, however, the government will regulate in an area in which it belongs, but the regulation will have a substantial effect on privacy rights. In that case, a court should directly balance interests using factors described below, giving deference to the legislative determination.[177] This approach resembles that of the Court under the First Amendment,[178] the constitutional provision most closely analogous to privacy rights.[179] The section below describes the sliding scale, its relevant factors and presumptive categories, and responds to criticism of balancing jurisprudence.

1. The Sliding Scale

Common sense dictates that the more central the right and the greater the infringement, the more compelling should be the government justification and the tighter the fit should be between the regulation and the justification.[180] A weak justification or poor fit relative to the privacy infringement creates an inference that the regulation was improperly motivated or irrational.[181] That alone should be sufficient to invalidate it.[182] To determine the suggested balance and ferret out improper

---

177. Deference is appropriate to the legislative determination because it represents the results of the democratic process. Although deference is a vague term, the Court has had much experience giving it meaning when reviewing lower court decisions. *See* Anderson v. City of Bessemer City, 470 U.S. 564, 574 (1985).

178. In the First Amendment context, incidental restrictions on speech that do not have a significant effect on free expression are upheld unless irrational. *See* Arcara v. Cloud Books, Inc., 478 U.S. 697, 702-04 (1986); Citizen Publ'g Co. v. United States, 394 U.S. 131, 140 (1969); Dorf, *supra* note 146, at 1201. Laws that discriminate on the basis of content or viewpoint are subject to strict scrutiny. *See* Boos v. Barry 485 U.S. 312, 321-22 (1988); Alan Brownstein, *How Rights Are Infringed: The Role of Undue Burden Analysis in Constitutional Doctrine*, 45 HASTINGS L.J. 867, 920 (1994). An incidental restriction on speech that targets or disproportionately effects expressive activity is subject to balancing. *See Arcara*, 478 U.S. at 710; United States v. O'Brien, 391 U.S. 367, 377 (1968); Dorf, *supra* note 146, at 1202.

179. The First Amendment protects free expression, in part, because it allows for self-actualization, much like privacy rights. *See* C. Edwin Baker, *The Scope of First Amendment Freedom of Speech*, 25 UCLA L. REV. 964, 995 (1978); THOMAS I. EMERSON, TOWARD A GENERAL THEORY OF THE FIRST AMENDMENT 5 (1966).

180. *See* Washington v. Glucksberg, 521 U.S. 702, 772 n.12 (1997) (Souter, J., concurring); Poe v. Ullman, 367 U.S. 497, 543 (Harlan, J., dissenting); Niles, *supra* note 161, at 132-33.

181. *See* ELY, *supra* note 3, at 145-47; Niles, *supra* note 161, at 133 & n.155; Turner v. Safley, 482 U.S. 78, 89-90 (1987).

182. *See Casey*, 505 U.S. at 877 (plurality opinion); Dorf, *supra* note 146, at 1182-83; TRIBE, *supra* note 164, at 1312.

motivation, courts should consider the importance of the right infringed and the extent of the infringement, the alternatives available to the individual, the directness of the infringement, as well as the justification and fit. Because privacy protection is, in part, designed to protect against the tyranny of the majority,[183] the court also should consider whether those affected are underrepresented or if there are other reasons to question the validity of the democratic process. Finally, to ensure institutional stability and avoid the tyranny of the minority, tradition may be considered. Although the Court does not acknowledge applying a balancing test, these factors have support in the case law.[184] The application of these factors will be illustrated in Part III.

a. Importance of the Right Infringed and the Extent of the Infringement

This article does not suggest that fine gradations between rights are possible. However, few would argue that an invasion of bodily integrity in the form of a restriction on hair length is indistinguishable from a law requiring all individuals to donate a kidney. The primary determinant for measuring the importance of the right infringed should be the consequence the regulation has on the life of the individual.[185] This is one reason the right to choose whether to have a child is so valued. Having a child alters your way of life dramatically. A child requires years of financial and emotional support and reduces the parent's freedom of movement and activity.[186] Although for most, the rewards of parenthood far outweigh the burdens, for those where that is not true, parenthood can negatively reshape their lives.[187] In any event, pregnancy also has extreme consequences for the woman's bodily integrity. Women experience weight gain, distortion of their bodies, emotional changes, fatigue, and often sickness.

No matter how important is the right infringed, if the extent of infringement is small, the consequences to the individual are diminished. Thus, a ban on abortions must be treated much more seriously than a regulation requiring a 24-hour waiting period before an abortion is performed. The Court explicitly recognized this factor in *Casey*.[188]

183. *See Casey*, 505 U.S. at 847 (citing *Poe*, 367 U.S. at 541 (Harlan, J., dissenting)).

184. *See infra* notes 185-217 and accompanying text.

185. *See, e.g., Eisenstadt*, 405 U.S. at 453; Jed Rubenfeld, *The Right of Privacy*, 102 HARV. L. REV. 737, 783-84 (1989).

186. *See* Karst, *supra* note 146, at 641 n.90.

187. For some, placing the child up for adoption may reduce the burden of having a child. However, for many, psychologically or morally, that is not a valid option. *Id.*

188. *Casey*, 505 U.S. at 886-87; *see also Zablocki*, 434 U.S. at 386; *id.* at 396 (Stewart, J., concurring).

b. Alternatives Available to the Individual

Closely related to the effect on the life of the individual are the alternatives available to the individual despite the government regulation.[189] For this reason a ban on the sale of condoms should not be treated the same as a ban on the sale of all contraceptives. The latter makes sexual intercourse a much greater risk and has much greater potential to significantly alter the life of the individual. For similar reasons, a federal regulation might be treated slightly more harshly than an identical state regulation. Although the option to move to another state is not available for many, the option to move to a different country, for most, is even more theoretical than real.[190] Also relevant to the availability of alternatives is the penalty imposed.[191] If the consequence of violating the law is only a small fine, the option of a knowing violation remains an option. A penalty requiring, or even allowing for, years of imprisonment effectively removes that option.

c. The Directness of the Infringement

To the individual whose rights are infringed, it may not matter if the infringement is direct or incidental. Nonetheless, this factor is relevant to determine whether the government regulation should be treated as the cause of the infringement as well as whether the government had an improper purpose.[192] If the infringement is too indirect, it may not be considered the proximate cause of the individual's injury. For example, environmental regulations may force a business to close, which in turn may deprive one of the laid-off workers from having the money to fund an abortion. This should not give rise to a claim for violation of the worker's privacy rights. It is equally clear that the environmental regulation, with such an indirect affect on privacy rights, could not have been motivated by a purpose to infringe those rights.

d. Justification and Fit

When considering the government's justification, the Court should only consider the legislature's actual purpose, not merely some hypothesized purpose. As explained by Professor Tribe:

---

189. *Cf. Turner*, 482 U.S. at 90 (explaining that a relevant factor in determining the reasonableness of prison restrictions infringing a constitutional right is "whether there are alternative means of exercising the right that remain open to prison inmates").

190. A local regulation leaves even more geographic options available to the individual than a federal or state regulation. *See* Moore v. City of E. Cleveland, 431 U.S. 494, 550 (1977) (White, J., dissenting). Nonetheless, once one goes below the federal level, a competing consideration offsets, to some extent, the availability of other geographic options. Specifically, the greater the number of locations that do not have the challenged restriction, the weaker is the local government's argument that the regulation is needed. *See Poe*, 367 U.S. at 526 (Harlan, J., dissenting).

191. *See, e.g.*, Loving v. Virginia, 388 U.S. 1, 10 (1967); Lofton v. Sec'y of the Dep't of Children & Family Servs., 358 F.3d 804, 817 (11th Cir. 2004); State v. J.P., 907 So. 2d 1101, 1110 (Fla. 2004).

192. *See Zablocki*, 434 U.S. at 391 (Burger, J., concurring); Dorf, *supra* note 146, at 1182-83.

> Knowing why government chose to enact a particular requirement, and why it is being enforced on a given occasion, bears on the way in which the requirement is likely to be perceived and hence the degree of affront it is likely to carry; it bears on the extent to which government's action is likely to chill protected choices in adjacent areas by persons who will inevitably understand not only what government has demanded but the principle on which government appears to have acted; it illuminates the degree to which invalidation of the requirement would serve to educate government itself with respect to the sorts of designs those in power should resist; and it assists courts in the inevitably difficult task of deciding how much weight to give to an alleged concern, recognizing that the history of how an argument found its way into a case—whether by hindsight or more genuinely—sheds at least some light on how the doubts regarding the argument's validity ought to be resolved.[193]

While a legislature can always reenact the invalidated regulation, providing a record that the purpose was as the government attorney hypothesized, many legislators would not be willing to engage in such a charade. In any event, the reason for deference to the legislature is that it represents the democratic process.[194] If the hypothesized purpose was not what the legislature intended, there should be no deference.

Given the importance of the rights at stake, the Court should require the government to provide support for the need for its regulation unless judicial notice would be proper.[195] The greater the infringement, the weaker the justification, and the poorer the fit, the greater should be the government's burden for evidentiary support. A record would be instrumental in determining the government's actual purpose and help evaluate the weight of the government's interest. For example, if the government justifies a regulation prohibiting midwives from providing abortion services based on safety concerns, it would be significant if midwife abortions led to complications in fifty percent or .0001 percent of abortions.

Assuming the government has a valid justification, a court must still consider the fit between the regulation and the government's asserted purpose.[196] Of course, perfect fit is rarely possible. The degree to which the fit is imperfect, however, should be used to diminish the weight of the government's justification. If the regulation is seriously underinclusive, it suggests that the asserted purpose either was not the true purpose or is not especially weighty. If a regulation is significantly over-

193. TRIBE, *supra* note 164, at 1307 n.13.
194. *See* Atkins v. Virginia, 536 U.S. 304, 323 (2002).
195. *See Turner*, 482 U.S. at 97; Stanley v. Georgia, 394 U.S. 557, 566-67 (1969); *Hodgkins*, 2004 WL 1854194, at *8.
196. *See Zablocki*, 434 U.S. at 396 (Stewart, J., concurring).

inclusive,[197] there are individuals whose rights are infringed for no apparent reason. In that case, the government should be required to pass a more tailored regulation.

The most common and significant justification that the government asserts is to prevent harm to minors[198] and non-consenting adults.[199] In analyzing this justification, a court should consider the severity of the injury, how speculative it is, and how directly it affects others. In judging the severity of the injury, physical injury presumptively should be considered more significant than economic or psychological injury. Whatever the severity of the injury, it should be downgraded by the likelihood that it will not occur.[200] When the severity of the injury, downgraded for its speculativeness, is small, or the injury is too indirect, an inference is raised that the injury was not the true purpose of the government regulation or, even if the true purpose, is insufficient to outweigh a significant infringement of privacy rights. This would be particularly true to the extent the fit was poor.

e. Likelihood of Defects in the Democratic Process

The principal argument against balancing and judicial scrutiny of legislative judgments is that it interferes with the democratic process.[201] Accordingly, when there are potential defects in the democratic process, greater judicial scrutiny is justified.[202] The primary factor to consider here, derived from *Caroline Products*' famous footnote four,[203] is whether the individuals whose rights are infringed were underrepresented in the legislative process.[204] It is likely that legislators tend to undervalue the interests of minority groups to which they do not be-

197. A regulation should not be considered over-inclusive if there is no reasonably satisfactory alternative to eliminate the over-inclusiveness that still achieves the government purpose. For example, consider a drug regulation that prohibits marijuana distribution. Even if the government is not concerned with distribution to someone who uses marijuana for medical purposes, the regulation is not necessarily over-inclusive to the extent it covers a distributor to such a person if the exclusion of such a person would defeat enforcement of the statute. Such might be true if uncertainty as to a distributor's customers prevented arrest for possession of large amounts of marijuana. Admittedly, determining the necessity of what might otherwise be considered over-inclusive often will be a difficult fact question.

198. Minors are presumed to be incapable of informed consent. The government also has an interest in ensuring that adult consent is informed and real. *See Paris Adult Theatre*, 413 U.S. at 57.

199. For Mill, prevention of harm to others is the only justification for infringing an individual's liberty interests. *See* MILL, *supra* note 142, at 24.

200. *Cf. Zablocki*, 434 U.S. at 395 (Stewart, J., concurring) ("The invasion of constitutionally protected liberty and the chance of erroneous prediction are simply too great.").

201. Iddo Porat, *The Dual Model of Balancing: A Model for the Proper Scope of Balancing in Constitutional Law*, 27 CARDOZO L. REV. 1393, 1429 (2006).

202. *See* United States v. Carolene Prods. Co., 304 U.S. 144, 152 n.4 (1938).

203. *Carolene Prods.*, 304 U.S. at 152 n.4 ("[P]rejudice against discrete and insular minorities may be a special condition, which tends seriously to curtail the operation of those political processes ordinarily to be relied upon to protect minorities, and which may call for a correspondingly more searching judicial inquiry.").

204. *See* Alexander Aleinikoff, *Constitutional Law in the Age of Balancing*, 96 YALE L.J. 943, 984 (1987).

long.[205] Greater scrutiny also is justified when the interests infringed may not have been fully considered given the hurly-burly of politics, or belong to unpopular groups.[206]

f. Tradition

As suggested earlier, tradition is not a valid test for defining fundamental rights.[207] Nonetheless, it may be useful to consider tradition, to the extent one can be agreed upon,[208] when balancing individual rights and government interests. Traditions are likely to be wise as they represent the considered judgment of many people over an extended period.[209] Knowledge of historical traditions also can inform judgments about legislative intent.[210] Finally, institutionally, the Court would lose respect if its judgments regularly were in opposition to public sentiment.[211] Tradition loses its persuasive force, however, to the extent the assumptions underlying the tradition have changed. For example, if the basis for the tradition against acknowledging gay marriages is the inability of gay couples to have children, innovations in artificial reproductive methods diminish the weight to be given to that tradition.[212]

2. Presumptive Categories

Presumptive categories are not truly exceptions to the sliding scale approach. Rather, they are situations where a rebuttable presumption exists as to the result of the balancing process. Again, the two categories where rebuttable presumptions exist are where government action in an area it belongs does not have a significant effect on individual rights, and where the government directly infringes privacy rights.

a. Government Activity Where It Belongs that Does Not Have a Significant Effect on Individual Rights

The threshold requirement of a significant effect to trigger heightened or full sliding scale review when the government operates in an area it belongs recognizes that almost any government regulation can indi-

205. *Id.* at 984 n.243 (explaining that the Court plays a representation-reinforcing role) (citing ELY, *supra* note 3, at 88-104).

206. Aleinikoff, *supra* note 204, at 984.

207. *See supra* notes 88-97 and accompanying text.

208. *See* Sunstein, *supra* note 88, at 2106; *supra* notes 88-93 and accompanying text.

209. *See id.*

210. *See* Brown, *supra* note 90, at 189.

211. *See Poe*, 367 U.S. at 542 (Harlan, J., dissenting).

212. Often, there may be disagreement about whether the premises behind a tradition have changed. For example, is there an increased recognition that gay couples may have children from prior heterosexual marriages or a greater frequency of adoption by gay couples that would undermine the assumption that gay marriages do not involve children? The burden of establishing that postulates upon which a tradition is based are no longer valid should be upon the person seeking to diminish the significance of the tradition. Conversely, the person relying on tradition should have the burden of establishing the tradition. In any event, tradition is not intended to be a binding, or even a critical factor, but merely a variable worth considering.

rectly or incidentally affect privacy interests. For example, laws against theft can prevent a poor thief from obtaining the funds necessary for an abortion. A husband who murders his wife might assert that the homicide laws interfere with his right to choose with whom to have an intimate relation. It would not be efficient to require extended review of such due process challenges, and in the absence of a significant effect, the constitutional right to privacy is, at most, only marginally implicated.[213]

The question remains, however, what is a significant effect? No precise definition is possible and once again, a certain measure of intuition is required. Nonetheless, some general principles suggest themselves. The more remote the effect, the fewer the number of people affected, the greater the alternatives available to the individual, the less likely a regulation should be considered significant.[214] The Court already imposes similar threshold requirements in the privacy and First Amendment areas.[215] Application of the significant effect requirement should prove no more elusive.

b. Direct Regulation of Privacy Interests

When the government directly regulates in an area it does not belong, the regulation should be presumed invalid unless the government can satisfy the traditional strict scrutiny test. When a law singles out protected conduct and is not narrowly tailored to serve a compelling government interest, "it is a fair inference that the law's principal purpose is the illegitimate one of frustrating the exercise of a right."[216] Again, that alone should be sufficient to invalidate the law.[217]

3. Criticisms of the Sliding Scale Approach Do Not Recommend Against Its Adoption

A sliding scale or balancing approach is not without its critics.[218] They argue that such an approach is too unpredictable, is just an excuse for justices to legislate their personal preferences, usurps the role of the legislature, and actually waters down protection of important rights.[219]

---

213. Of course, the government action still must be rational. *See* Washington v. Glucksberg, 521 U.S. 702, 766 (1997) (Souter, J., concurring); Daniels v. Williams, 474 U.S. 327, 331 (1986); *Poe*, 367 U.S. at 543 (Harlan, J., dissenting).

214. To this extent, this category involves balancing. However, the threshold is designed to allow dismissal of the extreme cases, such as the hypothetical situations posited in the text, without any significant litigation. In cases of doubt, the assumption should be that there is a significant effect so that the matter can be decided after a more detailed sliding scale review.

215. *See supra* notes 110-12, 163.

216. *See* Dorf, *supra* note 146, at 1235.

217. *See supra* note 182.

218. *See, e.g.*, Aleinikoff, *supra* note 204, at 943; Crump, *supra* note 4, at 906; Porat, *supra* note 201, at 1395.

219. *See* Aleinikoff, *supra* note 204, at 984-94; Crump, *supra* note 4, at 906, 910; Michael A. Scaperlanda, *Illusions of Liberty and Equality: An "Alien's" View of Tiered Scrutiny, Ad Hoc Balancing, Governmental Power, and Judicial Imperialism*, 55 CATH. U. L. REV. 5, 9, 31-33 (2005).

These arguments cannot justify courts' continued adherence to tiered review. A balancing test can be unpredictable and an excuse for justices to legislate their personal preferences. However, virtually all constitutional interpretation is value-laden.[220] Indeed, the Framers probably recognized as much, which likely is one reason why the Framers subjected judges to two layers of approval.[221]

Certainly, existing law is not better. One commentator has gone so far as to describe the Court's current approach as "like the methods of modern alchemy, conjuring up mystical formulae to conceal the sleight of hand by which a judge transforms the base metal of personal inclinations into the gold of fundamental rights."[222] The recommended approach, by providing transparency, should reduce unpredictability as well as judicial discretion. "By candidly acknowledging and celebrating the exercise of judgment, we can actually increase the possibility of accountability and ultimately hope to reduce the power of idiosyncratic decisionmaking." [223]

The more serious criticism of balancing is that it is undemocratic—by substituting the judgment of nine justices for the representatives of the people, balancing usurps the role of the legislature.[224] First, this article's proposal does not call for the Court to substitute its judgment for that of the legislature. The Court should give deference to legislative findings. Nor can the Court affirmatively legislate. It can only protect rights from government infringement. The need for the Court to fulfill its role as the primary protector of individual rights is particularly acute given the reality that Congress and state legislatures often don't consider the Constitutionality of the laws they pass.[225] However, it is not necessary to view the judiciary as better equipped to protect individual rights than the legislature.[226] It is obvious that the protection of important rights by two branches of the government is better than protection by just one. A multi-branch veto to legislation viewed as infringing rights may stymie majority preferences and to that extent seem undemocratic. However, the Framers were concerned with abuses of power and consequently created a government with a system of checks and balances.[227] This arti-

220. *See* ELY, *supra* note 3, at 67; Tribe & Dorf, *supra* note 89, at 1060, 1062.

221. *See* Brown, *supra* note 90, at 220 n.230.

222. Crump, *supra* note 4, at 805.

223. Brown, *supra* note 90, at 215; *see also* OLIVER WENDELL HOLMES, *The Path of the Law*, *in* COLLECTED LEGAL PAPERS 184 (1920) ("[Judges] themselves have failed adequately to recognize their duty of weighing considerations of social advantage. The duty is inevitable, and the result of the often proclaimed judicial aversion to deal with such considerations is simply to leave the very ground and foundation of judgments inarticulate, and often unconscious . . . .").

224. *See* Aleinikoff, *supra* note 204, at 984.

225. *See* Brown, *supra* note 90, at 186 (citing James B. Thayer, *The Origin and Scope of the American Doctrine of Constitutional Law*, 7 HARV. L. REV. 129, 155-56 (1893)).

226. *Cf.* ELY, *supra* note 3, at 67 (although recognizing undemocratic potential of legislature, questioning whether there is any reason to believe that the judiciary is any better).

227. *See* Tribe & Dorf, *supra* note 89, at 1064.

cle's proposal merely asks the judiciary to fulfill its role in that system. Moreover, given the vagueness of the language of the Fourteenth Amendment, reliance on precedent may be the best way to achieve predictability and consistency.[228] There is no one better to interpret precedent than the Supreme Court.[229]

The argument that balancing waters down the protection of the most important rights is somewhat disingenuous as those least interested in protecting individual rights are the ones who typically present it.[230] In any event, important rights will be protected if in fact it is agreed that they are truly important. Indeed, balancing may result in greater protection of important rights because courts, fearing the restrictiveness of strict scrutiny, will no longer need to avoid calling important rights fundamental.

The fact is that balancing is the way most people resolve issues. They consider the pros and cons of a course of action and try to decide which outweighs the other. Balancing "is the mark of a reasonable, rational, subtle mind."[231] The Supreme Court balances in numerous areas,[232] including in the area of fundamental rights.[233] This article only asks that the balancing process involving privacy rights be done explicitly and clearly.

III. APPLICATIONS OF THE PROPOSED STANDARD

This section applies the suggested proposal to several of the actual substantive due process issues confronting the courts today.[234] Although I would like to pretend that the proposal provides complete predictability and is totally value-neutral, that would be nonsense. Judgment is required for all Constitutional questions and such judgment necessarily is affected by the decision-makers' values and experiences. What recommends the suggested approach is its transparency. Competing interests are identified and reasons for their weighting are specified. Litigators will know how to build an appropriate factual record. Admittedly, unresolved questions abound. However, the more the Court resolves these

228. *Id.*
229. *Id.* at 1065.
230. *See* Aleinikoff, *supra* note 204, at 1004.
231. *Id.* at 962.
232. *See id.* at 963-73.
233. *See supra* notes 106-26 and accompanying text.
234. Without knowing the specific regulation enacted, the purpose of the enactment, or the factual record presented, it is not possible to resolve all issues that might arise. Alternative bases for constitutional challenge, such as the Equal Protection Clause, also are beyond the scope of this paper.

issues with clarity, the easier it will be to answer the remaining unresolved questions.[235]

A. Abortion

There is little question that regulation of abortion impacts important fundamental rights—the rights to bodily integrity and to choose one's intimate relations.[236] It is equally clear that the government is acting where it belongs when it regulates abortions. Bans on abortions, in at least some contexts, are commercial regulations, designed to prevent harm to others, and control activity in public places. Accordingly, a balancing of interests is required under this article's proposal. This section concludes that a ban on all abortions is unconstitutional. Although analysis of every type of lesser regulation is beyond the scope of this article, treatment of such regulations should mirror existing law. After all, the Court's undue burden standard is itself a balancing test.[237]

1. Extent of Infringement

A complete ban on abortions, as suggested earlier,[238] can have a huge effect on the life of the individual. Pregnancy plays havoc with a woman's body and parenthood involves a lifetime of emotional and financial commitments. There are few, if any, adequate alternatives for someone wanting an abortion. Giving birth and placing the child up for adoption may limit the lifetime effects on the mother, but does nothing to eradicate the infringement on the woman's bodily integrity. It also carries significant emotional consequences for the mother as well as the child. Abstinence is obviously unsatisfactory and birth control is not foolproof or available to some. The penalty for illegal abortions has traditionally been penal[239], making civil disobedience an unsatisfactory option. A complete ban is a direct regulation of the rights of women, a group that remains under-represented in the legislative branch. Although historically abortions may have been prohibited, recent tradition is to the contrary.[240]

---

235. For example, if the Court barred adultery prosecutions in the case of spousal consent, *see infra* notes 274-76 and accompanying text, it likely would preclude an action for statutory rape when there was parental consent, *see infra* note 280 and accompanying text.

236. Rubenfeld, *supra* note 13, at 739-40.

237. Although the analysis would be similar, as with all balancing tests, one can disagree about the conclusions. *See id.* at 878 ("Even when jurists reason from shared premises, some disagreement is inevitable . . . . That is to be expected in the application of any legal standard which must accommodate life's complexity.").

238. *See supra* notes 186-88 and accompanying text.

239. *See* Reva Siegel, *Reasoning from the Body: A Historical Perspective on Abortion Regulation and Questions of Equal Protection*, 44 STAN. L. REV. 261, 314 (1992).

240. *See Planned Parenthood*, 505 U.S. at 846.

2. Government Justifications

The government might seek to justify a ban on abortions as necessary to prevent harm to others, whether the fetus, the husband or the mother, to preserve the sanctity of human life, to prevent overreaching by unscrupulous practitioners of desperate or emotionally distraught women, or to regulate public morality. None of these justifications for a ban on abortions should be considered sufficient to offset a ban's significant infringement of individual rights.

a. Harm to Others

Harm to the fetus as a justification for a ban on abortions falters because a fetus is not a person and does not have rights under the Constitution.[241] If it were otherwise, states not only could forbid abortions, but would be compelled to, at least unless the health of the mother was implicated.[242] Similarly, a pregnant woman who consumed alcohol or caffeine might be arrested for child abuse. Nor can the government imbue the fetus with rights at the expense of an existing person's constitutional rights. As Professor Dworkin explains:

> The suggestion that states are free to declare a fetus a person . . . assumes that a state can curtail some persons' constitutional rights by adding new persons to the constitutional population. The constitutional rights of one citizen are of course very much affected by who or what else also has constitutional rights, because the rights of others may compete or conflict with his. . . .
>
> . . . If a state could declare trees to be persons with a constitutional right to life, it could prohibit publishing newspapers or books in spite of the First Amendment's guarantee of free speech, which could not be understood as a license to kill.[243]

Although a fetus cannot be imbued with rights, it might be argued that "killing" a fetus harms society by diminishing the value of life. This argument, however, has more in common with political sound bites than legal reasoning.[244] The "value of life" is employed selectively. For example, those who oppose abortion often support the death penalty.[245] The Court itself did not let the "sanctity of human life" stop it from im-

241. *Id.* at 860.

242. Ronald Dworkin, *Unenumerated Rights: Whether and How* Roe *Should Be Overruled,* 59 U. CHI. L. REV. 381, 398-99 (1992).

243. *Id.* at 400-01.

244. When vetoing expanded federal support for embryonic stem cell research, President Bush, a strong proponent of the death penalty and opponent of abortion, said the bill violated his principles on the sanctity of human life. *See* Sheryl Gay Stolberg, *First Bush Veto Maintains Limits on Stem Cell Research,* N.Y. TIMES, July 20, 2006, at A1.

245. *See* Arthur L. Rizer III, *Does True Conservatism Equal Anti-Death Penalty?,* 6 HOWARD SCROLL SOC. JUST. L. REV. 88, 115 (2004).

plicitly recognizing a right to die.[246] Moreover, abortion no more diminishes the value of life than prohibitions on abortion reduce the value of having children raised by people who want them. In any event, such abstract harm cannot outweigh the very real, direct, and substantial infringement of women's privacy rights.[247]

Abortion can harm the interests of a potential father who desperately wants to become a parent. The problem, however, is that when a mother and father disagree about whether to have a child, the view of only one can prevail. As between the two, the balance weighs in favor of the woman who physically bears the child and who is most affected by the pregnancy.[248]

Finally, abortions can harm the women who choose to have them.[249] However, if a woman's privacy interests mean anything, it must mean that a woman has the right to decide whether the potential harms to her from an abortion are outweighed by her desire to discontinue the pregnancy.[250]

b. Commercial Justifications

The government has an interest in regulating the commercial arena.[251] If the legislature found that abortion providers were taking advantage of women seeking abortions, it would be justified to impose reasonable regulations in the form of licensing, safety requirements, or even price regulation. If evidence showed that providers, interested in a quick profit, pressured women to have abortions that they later regretted,

---

246. *See* Cruzan v. Dir. Mo. Dep't of Health, 497 U.S. 261, 280 (1990).

247. *See Casey*, 505 U.S. at 914-16 (Stevens concurring in part, dissenting in part).

248. *See id.* at 896 (plurality opinion); Radhika Rao, *Reconceiving Privacy: Relationships and Reproductive Technology*, 45 UCLA L. REV. 1077, 1109 (1998). This is not to suggest that the government doesn't have any interest in protecting the rights of the father. A regulation that encourages a woman to consult the father that does not threaten the abuses identified by Justice O'Connor in *Casey*, 505 U.S. at 888-93, should be permissible. For example, a law might require the abortion provider, at the time an appointment is made, to suggest consultation with the father or discuss the impact the abortion decision has on others. Such a law would not significantly infringe the woman's privacy interests, yet would encourage communication with and the involvement of the father, which if nothing else, should lead to more stable relationships.

249. *See* Andrew A. Adams, *Aborting* Roe*: Jane Roe Questions the Viability of* Roe v. Wade, 9 TEX. REV. L. & POL. 325, 331-32 (2005).

250. Again, while the government interest in protecting the mother isn't sufficient to ban an abortion, it could justify reasonable regulations to ensure that the woman's consent is informed and considered. Although this article disagrees with *Casey*'s conclusion that a 24-hour waiting period does not unduly burden women's right to an abortion (particularly given the district court's findings of fact, *see Casey*, 505 U.S. at 885-86 (plurality opinion)), it would view a one-or two-hour waiting period, with an exception for medical emergencies, as permissible. For many women, the 24-hour wait would significantly increase the difficulty in obtaining an abortion and, absent more than anecdotal evidence, this article would assume that the government's interest could be served almost as well by the lesser wait. If the government was able to develop a record that showed that a significant percentage of women regretted having an abortion and would have changed their mind if required to wait 24 hours, the balance could come out differently. Requiring abortion clinics to develop follow-up records for this purpose should be permissible.

251. *But see supra* note 164.

a ban on commercial abortions might be justified. There is no reason to believe this currently to be the case. The government certainly has not developed evidence establishing such a problem. Finally, even if such evidence were developed, and given the extent of the infringement it would have to be a strong record, the government could not justify extending the ban on abortions on that basis to provision of abortion by free clinics.

c. Justifications Based Upon Public Activity or Effects

A government ban on abortions also might be based upon its police power to regulate public places and public morality. The obvious limitation of this justification is its inapplicability to abortions performed in private. As suggested earlier, the offense created by the mere knowledge that abortions are being conducted cannot support a ban on abortions.[252] The government should be able to ban abortions in government owned buildings, public places, or even places open to public view.[253] In such places, the privacy interests of the individual are diminished and the infringement generally will not be significant given the alternatives available, i.e., private abortions. Making abortions unavailable in public hospitals may deprive low-income women, particularly in rural areas, of the abortion option. Nonetheless, the government has the right not to support abortions under this article's proposal.[254]

In short, although numerous regulations affecting the abortion decision might be possible, a complete ban of abortion would be invalid under this article's proposal.

B. Gay Marriage

The initial question when analyzing a ban on gay marriages is whether there is an infringement of any protected right. Although under traditional doctrine, marriage is a fundamental right,[255] this article's pro-

252. *See supra* note 174 and accompanying text.

253. For example, the government should be able to require abortion providers to have blinds on their windows.

254. *See supra* note 158.

255. *See* Turner v. Safley, 482 U.S. 78, 95 (1987); Zablocki v. Redhail, 434 U.S. 374, 384 (1978); Loving v. Virginia, 388 U.S. 1, 12 (1967). Even under traditional doctrine, there is the question whether the fundamental right to marriage is limited to relationships between members of the opposite sex. A number of courts have said that it is. *See, e.g.*, Hernandez v. Robles, No. 05239, 2006 WL 1835429 (N.Y. July 6, 2006); Lewis v. Harris, 908 A.2d 196 (N.J. 2006). If so, a ban on gay marriage need only have a rational basis. This interpretation of the fundamental right to marriage seems to be a strained and result-oriented reading of Court precedent. In no case, recognizing a fundamental right to marriage, has the court suggested such a limitation. Indeed, the Court has said, "the right to marry is of fundamental importance *for all individuals*." *Zablocki*, 434 U.S. at 384 (emphasis added). The Court has recognized that marriages are "expressions of emotional support and public commitment" that for many have "spiritual significance, *Turner*, 482 U.S. at 95-96, and are "one of the vital personal rights essential to the orderly pursuit of happiness by free men." *Zablocki*, 434 U.S. at 383; *Loving*, 388 U.S. at 12. These reasons to treat marriage as fundamental are equally applicable to gay marriages. Furthermore, if marriage was limited to relationships of the

posal would not consider marriage a protected privacy right. Rather, it is just a benefit provided by the government. Accordingly, a government can choose not to support gay marriages, refusing to give it official recognition.[256] What the government cannot do is deprive gays of benefits because they are not married unless it has a rational basis unrelated to infringing gays' privacy rights.[257] Given the literally hundreds of rights that follow from marriage, the government would be hard-pressed to justify each and every one.[258] For example, there doesn't seem a rational basis for denying gay couples' access to their "significant other" in public hospitals that limit visitation to relatives. Similarly, it is hard to justify depriving gay couples the right of inheritance without a will. The Court in *Hernandez v. Robles*,[259] attempted to justify these types of discrimination as rationally related to inducement of marriage and its attendant benefits to children. Such a justification is bootstrapping. It is equivalent to saying that only persons who have not had an abortion are "bubbas." The benefit of government recognition as a bubba can be rationally justified as supporting childbirth which is necessary for the continuation of our society. The government can't then say only bubbas can receive food stamps and justify it as inducing bubbadom and its attendant benefits to society.

---

opposite sex based on the traditional definition of marriage, it could just as well be limited to relationships of the same race. That position was rejected in *Loving*.

Of course, under traditional doctrine if there were a fundamental right for gays to marry, a government ban on such marriages would be subject to strict scrutiny. *See supra* note 6. Given the number of childless heterosexual marriages and homosexual marriages with children, the typical child-based justifications for treating gay marriages differently could not satisfy the narrow tailoring required by that test. *See* Lawrence v. Texas, 539 U.S. 558, 605 (2003) (Scalia, J., dissenting); *Hernandez*, 2006 WL 1835429 (Kaye, J., dissenting); Goodridge v, Dep't of Pub. Health, 798 N.E.2d 941, 961-64 (Mass. 2003) (finding procreation or child-rearing justifications unable to meet rational basis scrutiny).

Contrary to the suggestions of some, see Elizabeth F. Emens, *Monogamy's Law: Compulsory Monogamy and Polyamorous Existence*, 29 N.Y.U. REV. L. & SOC. CHANGE 277, 279-80 nn.6, 8, 13-15, recognition of gay marriages would not necessitate acceptance of polygamous marriages. In the latter case, consent is not easily refused. *See Lawrence,* 539 U.S. at 569, 578. Without a valid consent, the government can deny recognition of polygamous marriages because they run the risk of psychological harm to existing spouses. Lack of consent would also distinguish denial of marriage to one's pet or favorite flower. *See* Sunstein, *supra* note 88, at 2083.

256. *See supra* note 158 and accompanying text. In any event, the government might justify its support of only heterosexual marriage as rationally related to its interest in supporting stability where accidental children are possible. *See, e.g.*, Morrison v. Sadler, 821 N.E.2d 15, 24 (Ind. 2005).

257. *See supra* note 159-160 and accompanying text. There is a question whether a ban on gays' right to marry does infringe their privacy rights. This article believes that denying gays the right to marry infringes the individual's choice in intimate sexual relationships, a right recognized in *Lawrence*. *See supra* note 86. If one believes *Lawrence* did not recognize such a right, or that this right is not significantly infringed, the ban on gay marriage would fall under my first presumptive category, *see supra* notes 155-57 and accompanying text, and be valid. One taking that position might suggest that whether or not the state recognizes gay marriage, the individual is free to have sexual relationships with whom they wish. This article would reject that argument as an exercise in semantics. The reality is that depriving gays of the benefits of marriage is punishing them for their sexual preferences. *See* Alaska Civil Liberties Union v. State, 122 P.3d 781, 788 (Alaska 2005).

258. *See, e.g.*, Hernandez v. Robles, No. 05239, 2006 WL 1835429, at *2 (N.Y. July 6, 2006) (counsel identified 316 benefits based upon marriage); *Goodridge*, 798 N.E.2d at 955-57 (Mass. 2003); *Lewis*, 908 A.2d at 196.

259. No. 05239, 2006 WL 1835429, at *21 (N.Y. July 6, 2006).

Even if one agrees that the unwillingness to allow gays to marry deprives them of many benefits without a rational basis unrelated to infringing gays' privacy rights, there remains the question of remedy. It may be that the proper remedy simply is to invalidate each of the benefits dependent on marriage that discriminate against gays. However, if the purpose of such benefits is to encourage stable, committed relationships, the proper relief may be to invalidate bans on gay marriage,[260] or require the enactment of a civil union provision providing gay couples with identical privileges as married heterosexuals.[261]

C. Medical Marijuana

The government's general regulation of recreational drugs is not jeopardized by this article's proposal. Such regulations would be valid under my first presumptive category.[262] The government's interest in preventing harm to others clearly outweighs the indirect and limited infringement on privacy rights and precluding drug enforcement would be a clear break from tradition.[263] The difficult issue is whether the ban should apply to users of recreational drugs such as marijuana for medical purposes. Of course, if there were other reasonable alternatives to marijuana to deal with the medical concerns, the ban should apply to such users. The denial of one drug when equivalent ones are available does not affect the life of the individual one iota more than the ban on marijuana use for recreational users. However, what if the only satisfactory relief from pain is provided by marijuana? This article would put the burden of establishing the unique medical benefits of marijuana on the user. Given the findings of the attorney general that marijuana does not have such unique uses, the burden would be heavy.[264] If the individual user were able to sustain that burden, this article likely would preclude the government from prosecuting such an individual for possession or use of small amounts of the drug.

---

260. The conclusion suggested in text is not dependent upon finding gays a protected class. Rather, it follows from a lack of a rational basis for depriving gays of many of the benefits from marriage other than to punish them for their choice of committed, intimate relationship.

261. *See* Baker v. State, 744 A.2d 864, 886 (Vt. 1999); *see also* VT. STAT. ANN. tit. 15, ch. 23; *Lewis*, 908 A.2d at 196.

262. *See supra* notes 155-57 and accompanying text.

263. The life of the recreational user will not be dramatically altered by a ban on recreational drugs, particularly given the availability of prescription drugs and alcohol. On the other hand, the government has an interest in preventing the secondary effects from marijuana use such as theft, assault, driving under the influence and accidental injuries. These effects are concerns even if the individual is using drugs in the privacy of her own home.

264. This article does not address the appropriate forum for an individual user's challenge. It may be that the individual should first challenge the attorney general's findings in an administrative forum. *See* Gonzales v. Raich, 545 U.S. 1, 33 (2005); County of Santa Cruz v. Ashcroft, 279 F. Supp. 2d 1192, 1203 (N.D. Cal. 2003).

Assuming marijuana uniquely relieved pain, a ban on its use would significantly affect the individual's interest in bodily integrity.[265] Given that the government is acting in an area it belongs (commercial regulation, preventing harm to others), a sliding scale analysis is appropriate. By assumption, the extent of the infringement would be great and there would be no satisfactory alternatives available to the individual. Although the infringement would be indirect,[266] the government's interest would not be especially strong. Many of the alleged harms from marijuana use are indirect, speculative, or unlikely to occur. For example, it is not clear what effect marijuana use has on crime and most marijuana users do not get into car accidents.[267] One also might question how strong the government's interest is when it does not regulate alcohol, a substance with very similar effects to marijuana. This is not to suggest that the government has no interest in marijuana enforcement.[268] Rather, that it is not sufficiently strong to outweigh the medical marijuana users' significant privacy interest in controlling their pain.[269] The government might argue that an exemption for medical marijuana would undermine the government's general drug enforcement efforts. Nonetheless, even if this interest was sufficient to outweigh the individual's privacy interest, given the experience with medical marijuana exemptions in other coun-

265. Although the Court has never held that relief from pain implicated an individual's privacy rights, five justices in *Glucksberg*, suggested there was a fundamental right to use physician-recommended medication to alleviate pain and suffering to control the circumstances of their death. *See* Washington v. Glucksberg, 521 U.S. 702, 737 (1997) (O'Connor, J., concurring, joined by Ginsberg, J.); *id.* at 743, 745 (Stevens, J., concurring); *id.* at 777 (Souter, J., concurring); *id.* at 790 (Breyer, J., concurring). In *Regina v. Parker*, 49 O.R.3d 481 (Ont. C.A. 2000), the Canadian court explicitly found that a ban on the use of medical marijuana infringed the individual's right to bodily integrity.

266. Normally, the indirectness of the infringement would suggest that the government had a proper purpose. This is certainly true with respect to the Narcotics Act generally. One wonders, however, whether the inclusion of marijuana as a schedule one drug (a drug subject to abuse, devoid of legitimate medical uses, and lacking accepted safety under medical supervision 21 U.S.C. 812) in 1970 wasn't, at least in part, motivated by animus toward the "hippie" generation.

267. *See* Marcia Tiersky, *Medical Marijuana: Putting the Power Where it Belongs*, 93 NW. U.L. REV. 547, 574 (1999).

268. Even the Canadian Justice who found the use of medical marijuana protected under the Canadian Constitution found that the government did have a valid interest in marijuana enforcement generally. *See* Regina v. Clay, 49 O.R.3d 577, 592 (C.A. Ont. 2000). Obviously, the more dangerous the drug, the stronger the government's interest. Thus, an exemption for marijuana use does not mandate an exemption for cocaine or heroine use.

269. The remaining sliding scale factors are not especially helpful. The users of medical marijuana may be under-represented in the legislature, but legislators likely would view their interests sympathetically. Although there is at least a recent tradition of regulating marijuana, that tradition would have little weight given that the premise on which the tradition was based, the lack of a unique medical use, has changed by assumption. Moreover, an increasing number of states have recognized the medical benefits of marijuana, *see* Note, *Last Resorts and Fundamental Rights: The Substantive Due Process Implications of Prohibitions on Medical Marijuana,* 118 HARV. L. REV. 1985, 2005 n.107 (2005) and the U.N. Single Convention on Narcotic Drugs (1961) which requires marijuana to be listed as a controlled substance, exempts use for medical purposes. *See Clay*, 49 O.R.3d 577, 590 (C.A. Ont. 2000). If anything, the experience in other countries raises questions about the government's need for regulation.

tries,[270] the burden should be on the government to demonstrate the validity of this concern.

D. Criminal Sexual Conduct

This article's proposal will not have a significant practical effect on prosecution of criminal sexual conduct cases. The two primary types of statutes that might be invalidated, fornication and adultery statutes, are currently rarely enforced.[271] Cases involving rape, prostitution, and incest would proceed largely unchanged.

1. Fornication

A fornication statute could not properly be enforced under this article's proposal as to consenting adults having sex in private.[272] Such a statute directly infringes on the individual's privacy rights. Because the government would be acting in an area it did not belong, it would need to demonstrate a justification that was narrowly tailored to achieve a compelling government interest. The government could not credibly claim any such compelling interest given such statutes' history of non-enforcement.[273]

The government could regulate public exhibition of sexual acts. That is an area the government belongs and does not significantly infringe any privacy rights. Such regulation would be presumptively valid. The government also could deny certain benefits to persons having private sexual relations with a consenting adult. For example, a public university could preclude professors from fornicating with students. There are rational reasons unrelated to infringing the professors' privacy rights for such a denial, including fear of sexual harassment suits or perceived

270. *See supra* note 269. For a description of the Canadian system for exemption of medical marijuana, see Marijuana Medical Access Regulations, SOR/2001-227, discussed in R. Wood, [2006] N.B.J. No. 254, 1, 7-8 (N.B.C. June 20, 2006).

271. *See* Karst, *supra* note 146, at 670, 674. The non-enforcement of criminal statutes does not moot the question of its constitutionality. The constitutionality of criminal statutes also can affect civil liability. For example, in *Martin v. Ziherl*, 269 Va. 35 (Va. 2005), the plaintiff alleged that the defendant engaged in sexual relations with her knowing that he was infected with the herpes virus and that he was contagious, and failed to inform her of those facts. If the fornication statute was constitutional, the plaintiff would have been guilty of criminal conduct and, under a Virginia law analogous to the "unclean hands" doctrine, her tort action would have been barred. *Martin*, 269 Va. at 38. The Court found the statute unconstitutional, reversed the lower court decision, and allowed the plaintiff's tort case to proceed. *Id.* at 38, 42-43.

272. If it is true that *Lawrence* recognized a right to private consensual sex between consenting adults, as this article has argued, *see supra* note 86, fornication statutes would be invalid under existing law also. Even if *Lawrence* didn't recognize such a right, it could be argued that the right to beget a child, a right the Court has recognized, *see, e.g.*, Lawrence v. Texas, 539 U.S. 558, 565 (2003); Eisenstadt v. Baird, 405 U.S. 438, 453 & n.10 (1972), necessarily includes the right to do what is a prerequisite for that right. *See* State v. Saunders, 381 A.2d 333, 339-40 (N.J. 1977).

273. *See* Poe v. Ullman, 367 U.S. 497, 554 (Harlan, J., dissenting). The government has a strong interest in ensuring that the sexual relations are truly consensual. However, the rape statutes fully protect that interest.

bias by other students.[274] The school also could want an announced policy so that students wouldn't erroneously confuse faculty concern for sexual advances.

2. Adultery

Adultery statutes directly interfere with private sexual relations between consenting adults. The government, however, could assert that it is seeking to prevent harm to the spouse and children of the adulterer. A sliding scale analysis would then be appropriate. To the extent the adulterer is deprived of sexual relations with someone she cares deeply for, it is a significant infringement. The infringement is direct. The harm asserted by the government is speculative, indirect, and does not involve physical injury. There also would be an issue with fit. If the adulterer has no children and the spouse has not complained or consented, there does not appear to be reason for the statute. On the other hand, divorce is an alternative that would avoid application of the statute, there is a tradition against adultery, and, unfortunately, adulterers probably are not under-represented in the legislature. What might be the deciding factor is the history of non-enforcement of adultery statutes.[275] As with fornication statutes, that seriously undermines the government's assertion of need.[276]

3. Rape

Under this article's proposal a basic rape statute would be valid. A rape statute does not interfere with any privacy right[277] and the government has a strong interest in preventing harm to the non-consenting party. Less clear is the proposal's effect on statutory rape provisions. By definition, statutory rape does not involve consenting adults, and the government has an interest in protecting minors from harm. Nonetheless, the Court has recognized that minors also have constitutional rights, including the right to privacy.[278] A statutory rape statute does burden the minor's right to intimate relations. Although, the minor may be deemed incapable of a valid consent, by analogy to the abortion cases, [279] a statutory rape statute, to be constitutional, might need to have an exemption

274. In the case of a teacher-student relationship, there also would be a question if consent could be easily refused. *See Lawrence*, 539 U.S. at 578; Rubenfeld, *supra* note 13, at 756-57.

275. *See supra* note 271.

276. The government might alternatively justify prohibitions on adultery as a rational condition on the benefit of marriage. However, it is unclear that such a condition is ever specified, and if it was, why the remedy should be criminal prosecution rather than the elimination of the benefit of marriage, i.e., divorce.

277. Rape, by definition, does not involve private consensual conduct.

278. *See* Carey v. Population Serv. Int'l, 431 U.S. 678, 692 (1977).

279. *See, e.g.*, Bellotti v. Baird, 443 U.S. 622, 631-32 (1979); Hodgson v. Minnesota, 497 U.S. 417 (1990). Admittedly, the abortion cases are not a perfect analogy. Statutory rape statutes do not impose the same infringement on the individual as a ban on abortion. Rather than significantly interfering with the mother's bodily integrity and imposing lifetime consequences, a statutory rape statute only requires lesser forms of sexual gratification for a finite period of time.

for when there is parental or judicial consent.[280] The need for such an exemption would be strongest in cases involving little disparity in the ages of the sexual partners. In such cases, intimacy is common and no longer has the same tradition of disapproval.

4. Prostitution

Laws against prostitution should be valid. The government can justify prostitution laws as commercial regulation necessary to preserve public morals and prevent harm to others. Ostensibly, prostitution laws interfere with private intimate relations between consenting adults. However, there is a question whether sex for a fee can be considered truly intimate. Even if such laws were analyzed under the sliding scale, the balance probably would favor the government. The interference with privacy rights are indirect and of limited significance for most.[281] Alternatives are available to individuals—sex with willing participants, masturbation, prostitutes in Nevada or foreign countries, and there is a tradition of laws banning prostitution. The government's concerns with public morality, crime, and health risks are real. Although the injuries are indirect and somewhat speculative, they can be severe. Regulation rather than prohibition can limit these effects. Nevertheless, the government's choice for a complete ban probably should be given deference given the relatively limited infringement of privacy rights and the tradition of criminalizing prostitution.

5. Incest

Laws prohibiting the paradigmatic cases of incest would be unchanged by this article's proposal. Minors are not capable of consent and the government has obvious interests in avoiding the psychological harm to the minor and the potential physical harms to any offspring produced by the incestuous relationship.[282] A more difficult question is how to treat incest involving adults. Here there is no question that there is interference with the individual's right to private intimate relations with consenting adults. The infringement is direct and for some is significant because it deprives the individual of intimate relations with one who may be the love of his or her life. For such a person, there is no substitute and the infringement lasts a lifetime, not merely until adulthood. Nonethe-

280. Under current law, failure to include a parental approval provision also might violate a parents' presumptive right to control their child's upbringing. *See* Troxel v. Granville, 530 U.S. 57, 65-69 (2000).

281. Perhaps a truly repulsive person might argue that prostitution is the only avenue for them to have sex with another. However, it is unclear how the individual could carry his burden of proof and doubtful that a court would want to suggest that someone is incapable of attracting a person of the opposite sex because of their appearance.

282. Obviously, an exemption in the case of parental consent is unnecessary when the parent is the party guilty of incest. Even if the adult is another relative, a parental consent provision should be unnecessary. Unlike in the case of statutory rape, conflicts of interests are involved when the accused is a relative and incest has a long and ongoing tradition of societal disapprobation.

less, the government has an interest in preventing offspring with increased likelihood of birth defects and decreased intelligence. Although such harm is speculative, the physical harm may be severe. The government could also argue that the consent is not valid when one of the incestuous parties is in a position of trust or if consent might otherwise be difficult to refuse.[283] There might be an applied challenge in a case where children are not possible or there is no blood relationship,[284] and there is no reason to question the validity of the consent. However, given the tradition against incest and the deference due to the government, the incest statutes should be facially valid.

E. Artificial Reproduction

The government can regulate the commercial aspects of artificial reproduction with nothing more than a rational basis unrelated to infringing privacy rights. However, a ban on artificial reproduction should be invalid. For many seeking birth artificially, there is no alternative method to have a child. Depriving someone of the parent-child relationship is a tremendous infringement of the individual's privacy rights. It is a direct interference that has lifetime consequences. There also is no tradition of outlawing artificial reproduction techniques. If it were shown that there was a much greater chance of abnormalities through artificial methods of birth or if a population control problem of much greater dimension than currently exists developed, the government might be able to justify a ban on artificial reproduction. However, given the significant infringement of privacy rights, the government would be required to develop a strong record to that effect.

F. Sale of Sex Toys

Several cases have upheld a ban on the sale of sex toys.[285] These cases would remain good law under this article's proposal. A state's ban on the sale of sex toys has a *de minimus* affect on an individual's privacy rights. For most, the use or non-use of sex toys does not significantly

---

283. *See supra* note 274.

284. For example, some incest statutes cover relatives by adoption. *See, e.g.*, ALA CODE § 13A-13-3 (2006); LA. REV. STAT. ANN. § 14:78.1 (2006); TENN. CODE ANN. § 39-15-302 (2005); TEX PENAL CODE ANN. § 25.02 (2005). There also is a question whether the possible harms justify prohibitions of incestuous relationship involving first or second cousins. *See* Robin L Bennett, et. al., *Genetic Counseling and Screening of Consanguinous Couples and Their Offspring: Recommendations of the National Society of Genetic Counselors*, 11 J. GENETIC COUNSELING 97 (2002) (first cousins are only 1.7-2.8% more likely than unrelated parents to have children with birth defects or mental retardation); Richard Coniff, *Go Ahead, Kiss Your Cousin: Heck, Marry Her If You Want To*, DISCOVER, Aug. 2002, *available at* http://www.discover.com/issues/aug-02/features/featkiss/ ("first-cousin marriages entail roughly the same increased risk of abnormality that a woman undertakes when she gives birth at 41 rather than at 30"); Denise Grady, *Few Risks Seen to the Children of 1st Cousins*, N.Y.TIMES, April 4, 2002, at A1 (stating that medical geneticists have known "for a long time that there was little or no harm in cousins marrying and having children").

285. *See, e.g.*, Williams v. Att'y Gen. of Ala., 378 F.3d 1232, 1250 (11th Cir. 2004); Pleasureland Museum, Inc. v. Beutter, 288 F.3d 988 (7th Cir. 2002).

affect their life. There are alternatives available to the individual, including traditional sex or manual stimulation. Sex toys generally even can be bought from other states without leaving one's home. The state easily can justify the sales ban as maintaining public morals and eliminating the secondary effects of the sale of such devices.[286] However, a statute prohibiting the private use or possession of sex toys should be problematic.[287] Obviously, private use cannot be directly justified by public morals or the secondary effect of sales. Nor could a state justify the ban on use as necessary to prevent sales. Unlike drugs, the sale of sex toys generally is not done surreptitiously on the streets. Direct enforcement of the sales ban is relatively easy. Although for most, the effect of the ban is not significant, for some, sex toys are used therapeutically to alleviate sexual dysfunction.[288] Absent any reasonable justification, this minor infringement of privacy rights should be sufficient to invalidate a prohibition on use.

CONCLUSION

If consistency has any value, the Court's fundamental rights/right to privacy jurisprudence is bankrupt. The Court fluctuates between alternative tests to determine whether a fundamental right exists. The dominant test in the lower courts, the "tradition test," lacks its promised objectivity, conflicts with the structure of the Constitution, and is incompatible with Court precedent. The standard of review for fundamental rights is even more problematic. The Court's language and practice are inconsistent. Hornbook law states that a law that infringes privacy rights must be narrowly tailored to further a compelling government interest. Yet, in many cases the Court applies an open-ended balancing test. This dichotomy between language and practice confuses lower courts and effectively paralyzes them from finding new fundamental rights.

This article has suggested a Lockean conceptualization of the right to privacy—a pact between individuals and the government to forego certain rights that are necessary to further society's interests, but with a reservation of rights in certain private areas where the government does not belong. The areas the government does not belong include one's thoughts, feelings, bodily integrity, and private intimate relationships. By contrast, the government acts in a proper area when it provides government benefits, regulates commercial activity and activity in public places, and seeks to prevent harm to non-consenting parties. A denial of government benefits is valid unless it has no rational basis other than to harm privacy interests. Government regulation infringing privacy inter-

286. *Pleasureland Museum*, 288 F.3d at 993 n.1.

287. *Cf.* Stanley v. Georgia, 394 U.S. 557, 567-68 (1969) (Prohibition of private possession of obscene material cannot be justified as "necessary incident to statutory schemes prohibiting distribution" of obscene material).

288. *See* Herald, *supra* note 91, at 23-26.

ests is judged by a balancing test which considers the importance of the right infringed, the extent of the infringement, the alternative available to the individual, the directness of the infringement, the government's justification and how closely the regulation fits the government's needs, the likelihood of defects in the democratic process, and tradition. Where the government regulates in an area it belongs and doesn't significantly infringe privacy rights the balance is presumed to allow the regulation unless irrational. Where the government infringes rights in an area it does not belong, the regulation is presumed invalid unless the government demonstrates that its regulation is narrowly tailored to further a compelling government interest. In the majority of cases, where government regulation is in an area it belongs but significantly infringes privacy rights, detailed balancing is required. Balancing is what people do; it is what the Court does.[289] It is time for the Court to openly and honestly acknowledge its practice and provide clear guidance to the lower courts.

The recommended approach was applied to resolve a number of privacy issues facing the courts. Admittedly, one can legitimately disagree with the conclusions reached. Nonetheless, what recommends the suggested methodology is its clarity. Competing interests are identified and the reasons for their weighting are specified. Litigators will know how to build an appropriate factual record, and lower courts will have a blueprint for analysis. Most importantly, a uniform approach would bring some measure of consistency to an area that for too long has lacked it.

289. *See supra* note 144 and accompanying text.

ASSESSING ALTERNATIVE COMPENSATION MODELS FOR ONLINE CONTENT CONSUMPTION

TAL Z. ZARSKY†

TABLE OF CONTENTS

---

† Lecturer, University of Haifa – Faculty of Law. LL.B./B.A., Hebrew University, LL.M., J.S.D., Columbia University. I would like to thank the organizers of the Intellectual Property and Digital Media conference at The Cable Center at the University of Denver for providing me with a platform to present these ideas. I also thank Neil Netanel, Oren Bracha, and Michael Birnhack for their important insights, and Ran Madjar and Ronen Fluss for their research. Research for this paper was partially funded by the generous help of the Infrastructure Fund – The Israeli Ministry of Science.

INTRODUCTION

The quick emergence of the Internet from a network that facilitates limited communications among academics and governmental agencies to a worldwide and extremely popular medium has changed and challenged the worlds of law and business in a variety of ways. Yet the content industry, especially its segments which involve the production and distribution of popular music, have perhaps been the most affected by the Internet's ability to allow for global, efficient, and cheap communications and data exchange. Over recent years, the leaders of the music industry have witnessed their existing business models come under attack. Rather than purchasing or licensing music from them directly, Internet users of all ages are accessing songs and other forms of content of their choosing online, without the consent or control of those holding the legal rights to such content ("copyright holders")[1] and without compensating them for such use.[2] In doing so, online users rely upon techno-

1. Throughout the analysis, I will refer to such entities as the "copyright holders," without addressing the specific intricacies of copyright law that set out the rights afforded to those that compose, author, and record the specific works. In the context of this paper, these distinctions are not crucial.

2. For recent information as to the extent of the file swapping phenomenon, see David W. Opderbeck, *Peer-to-Peer Networks, Technological Evolution, and Intellectual Property Reverse Private Attorney General Litigation*, 20 BERKELEY TECH. L.J. 1685, 1696-99 (2005); *see also* Neil Netanel, *Impose a Noncommercial Use Levy to Allow Free Peer-to-Peer File Sharing*, 17 HARV. J.L. & TECH. 1, 2-4 (2003).

logical innovations that provide access to vast amounts of content, and the quick transfer of data among users. Yet concerns regarding the strength and prospect of current business models are no longer limited to the music industry alone. Movie studios and other entities involved in the production of content[3] are quickly acknowledging the challenges they face in the twenty-first century given the ease of digital copying and the extent of content available online, and are constantly contemplating proper strategies to respond.

The troubles of the content industry, however, are not merely concerning the existence of destructive technological applications and communication networks. The technological innovations mentioned have caused substantial changes in the behavioral patterns of a (mostly) law-abiding segment of society. Citizens (especially, but not only, youngsters) are constantly and in many instances knowingly violating copyright law and infringing upon the rights of the copyright holders.[4] In other words, within certain social circles, social norms are now quite different than the actual copyright laws on the books.[5]

In view of these developments, the content industry's leaders are moving swiftly to secure and even improve their position in the online content markets. In many instances, they are making use of their influence and capital to guarantee the assistance of regulators in these efforts. As part of these efforts, they have suggested and begun to implement new models of content distribution that are premised upon secured networks and encrypted content.[6] With these measures in place, copyright holders should be able to capitalize on the Internet's broad communications and worldwide access, while assuring that they will be compensated for all instances in which their content is used. These models, usually referred to as digital rights management ("DRM") schemes, are facing a broad opposition of scholars, activists, and concerned citizens. This broad "coalition of dissent" forcefully argues that DRM schemes will impede on the public's rights to fair use, privacy and other fundamental rights.[7] Others argue in addition that there is no evidence that the content industry's existing business models are compromised and that the content providers' income is reduced in view of unauthorized online

---

3. Examples include eBooks, software and games; however, the analysis set out in this article will focus exclusively on music and video content.

4. For interesting insights as to the rationales for such file-sharing, see Gali Einav, *College Students: The Rationale for P2P Video File Sharing*, 2004 CITI Working Papers (on file with author); *see also* DAVID CALLAHAN, THE CHEATING CULTURE 185-87 (2004). For a discussion of file sharing and social norms, see Opderbeck, *supra* note 2, at 1700-01.

5. *See* WILLIAM W. FISHER, PROMISES TO KEEP 243-44 (2004); *see also* Tim Wu, *When Code Isn't Law*, 89 VA. L. REV. 679, 722-25 (2003).

6. *See* PAUL GOLDSTEIN, COPYRIGHT'S HIGHWAY 170 (2003) (discussing the early origins of this idea).

7. *See* discussion *infra* Part II.B.

content access.[8] The DRM schemes are also criticized by other powerful industry players, such as telecommunication providers and hardware manufacturers,[9] whose interests are not always aligned (and at times are even opposed) with those of the content industry.[10]

Beyond the overall attack on DRM, several scholars have been arguing for the complete or partial abolishment of copyright protection in the online context. Such protection, they argue, is no longer required as copyright holders can easily compensate for the income stream that the unauthorized online access to content is diminishing through the other advantages the online environment provides.[11] Artists could use the Internet to promote their offline products (such as live performances, CDs and other forms of merchandise) or even rely on the users' benevolence to compensate them for the enjoyment of their works (in the same manner the local artist is tipped at a street corner).[12] Clearly such arguments are not embraced by the content industry, and as such are not likely to be implemented on a broad scheme.

Within this spectrum of intellectual debate (which has many practical implications) concerning the future of copyright protection in the online realm, between strong online copyright protection and copyright abandonment, several new ideas and models have been recently discussed; ideas that offer sufficient incentives to generate content production in the digital era by well-compensated artists, while protecting the social interests involved. These are Alternative Compensation Schemes ("ACS") for the use of content online that are specially tailored to meet the specific challenges of the online world. These models, set forth by prominent legal scholars such as Terry Fisher[13] and Neil Netanel,[14] which rely in part on earlier scholarship ("The ACS scholars"),[15] provide for indirect compensation to the copyright holders of various works, which would be distributed by a governmental entity. The extent of the

8. *See generally* CALLAHAN, *supra* note 4.

9. *See generally* FISHER, *supra* note 5, at 242.

10. Interesting problems arise when a media conglomerate (such as Sony or GE) includes both a manufacturing and content division. Here, while one division might suffer from the ongoing file swapping, the other benefits by growth in the sales of relevant hardware. Here the conglomerate struggles in formulating its overall strategy in addressing these matters, and at times takes contradicting positions.

11. *See* Netanel, *supra* note 2, at 74-76.

12. Netanel refers to these forms of solutions as "digital abandon." *See* Netanel, *supra* note 2, at 74-76. This dynamic is at times referred to as the "Street Performer Protocol." *See* John Kelsey & Bruce Schneier, *The Street Performer Protocol and Digital Copyrights*, FIRST MONDAY (1999), http://www.firstmonday.dk/issues/issue4_6/kelsey/#k4.

13. *See generally* FISHER, *supra* note 5.

14. *See generally* Netanel, *supra* note 2.

15. *See, e.g.*, Raymond Shih Ray Ku, *The Creative Destruction of Copyright: Napster and the New Economics of Digital Technology*, 69 U. CHI. L. REV. 263, 263-70 (2002); Glynn S. Lunney, Jr., *The Death of Copyright: Digital Technology, Private Copying, and the Digital Millennium Copyright Act*, 87 VA. L. REV. 813, 813 (2001). For a recent summary of these models, see Jessica Litman, *Sharing and Stealing*, 27 HASTINGS COMM. & ENT. L.J. 1, 34-35 (2004).

copyright holders' compensation would depend on the relative uses of their content online, and would be distributed from a designated fund. The fund would be financed through levies set on specific services and equipment that are related to the online experience. As I will illustrate below, the recent scholarship addressing these alternative models for compensation in the Internet context is not engaged in merely floating abstract notions and legal concepts. The ACS Scholars go to great lengths to draw out, in extensive detail, the ways in which these schemes could and should be implemented.

According to the ACS scholars, implementing these models will meet several important objectives. These models will allow society to maintain a vibrant market of content production and online distribution, while changing today's reality in which millions of citizens are rendered infringers (and at times outlaws) by copyright laws.[16] In addition, the shift to the ACS model will hopefully mitigate several inefficiencies in today's business and legal frameworks: the existence of costly and extensive litigation that is required for resolving copyright disputes and sharpening the meaning of legal rules and terms;[17] the imbalance of power between the large media conglomerates and artists that have yet to transition into stars; and the fact that only a small portion of all artists are able to make a decent living off their talents.[18] Yet perhaps above all, the ACS scholars' objective in constructing these models is to allow for the enrichment of the public sphere with a great variety of easily accessible content for both the users' consumption and modification.[19] By doing so, they aim to promote important ideals related to free speech and democracy.

In this article I closely examine and constructively critique the ACS models and scholarship. In doing so, I part from the already growing base of literature addressing this issue that has been quick to reject ACS for a variety of reasons, without taking a close look at its internal dynamics.[20] I, however, choose to focus on the model's inner workings—its "nuts and bolts" that are the mechanisms aimed at transforming this model from abstract policy ideas into actual regulations and business practices. In doing so, I examine whether the model could be implemented as described and the possible outcomes of such implementation. This analysis leads to concrete suggestions for several changes in the

16. *See* FISHER, *supra* note 5 at 243.

17. *Id.* at 245-46.

18. *Id.*

19. *Id.* at 245.

20. *See, e.g.*, Mark A. Lemley & R. Anthony Reese, *Reducing Digital Copyright Infringement Without Restricting Innovation*, 56 STAN. L. REV. 1345, 1408-10 (2004); Rebecca Tushnet, *Copy This Essay: How Fair Use Doctrine Harms Free Speech and How Copying Serves It*, 114 YALE L.J. 535, 589-90 (2004); Peter K. Yu, *P2P and the Future of Private Copying*, 76 U. COLO. L. REV. 653, 708-12 (2005).

model to provide for its smooth, efficient, and fair implementation. In addition, it offers several points for future consideration regarding the balancing of benefits the scheme will bring against the unintended consequences of its implementation. In this article, therefore, I hope to lead the way to future scholarship and technological innovation that will respond to the challenges this paper draws out and addresses.

To thoroughly address this issue, the article is structured as follows: In Part I, I draw out the legal, business, and technological background leading to the current reality in which the ACS schemes might prove necessary. In view of the fact that a great deal of legal and other scholarship (including, of course, that of the ACS scholars) has addressed this background in length, I chose to address these facts in brief, while adding references to the most recent legal, technological, and business changes occurring in the online realm. In Part II, I address the DRM solutions promoted by the content industry, the legal and policy rationales that stand behind this model, as well as their shortcomings in terms of economic feasibility and technological sustainability. The importance of this part of the analysis is to establish a baseline for comparison which will serve us in the latter parts of this article. Only after understanding the DRM scheme can we later compare it to those offered by the ACS scholars while trying to establish which will lead to better results for artists, content owners, users, and society in general. In Part III, I address in further detail the ACS models, while focusing on the issues drawn out by Professor Fisher in his recent book, *Promises to Keep*,[21] and Professor Netanel in a recent article published with the Harvard Journal of Law and Technology.[22] In Part IV, I offer a constructive critique of the alternative compensation models, while addressing difficulties in their implementation and problematic results that might arise from their adoption.

Before going further, I must introduce several underlying assumptions that are needed to explain the somewhat limited scope of the discussion and analysis at hand. While the arguments presented throughout this article at times strongly oppose those promoted by the ACS scholars, I accept the notion that the implementation of these schemes is politically and legally feasible. Furthermore, I accept most of the legal and economic descriptions and analyses the ACS scholars provide as to the structure of today's content markets. In addition, I agree that unless the mentioned changes are made, courts and legislators will maintain today's legal status quo that embraces the rights of the copyright holders. For instance, I concede the fact that courts addressing today's copyright law in the near future will continue to find online file sharing as an infringement of copyright and will continue to allow the "breakdown" on such

21. *See generally* FISHER, *supra* note 5.

22. *See generally* Netanel, *supra* note 2.

infringers, which include many users that are usually law-abiding citizens. Within this framework and while applying these assumptions, I engage in an analysis that hopefully will promote this strand of scholarship that offers an interesting answer to a much discussed question.

I. THE CONTENT MARKET IN THE TWENTY-FIRST CENTURY: A PRIMER ON WHERE WE ARE TODAY

To understand the ACS scholarship, we must first acknowledge the motivations and concerns of the ACS scholars, as well as the technological, economic, and legal foundations of their analysis. We must do so to later examine how and whether these are met in the blueprint they provide for the ACS. Generally, from the perspective of the ACS scholars (which I share), today's innovations create an *opportunity* made possible by technology, and a *crossroads* for regulators. To meet important social objectives, the ACS scholars have a strong preference towards one path. However, in view of regulatory paths taken in the past and the balance of power at the present, they believe that in the future, unfortunately, another path will be taken. Thus, they draw out an ambitious plan that would allow for the maximization of the social benefits of technology, while allowing content creators to maintain some of the rights and benefits they have today.[23] As we are quietly approaching a crucial crossroads at which regulatory decisions concerning the future of copyright policy must be made and the other dominant options currently debated are in their opinion extremely unattractive, the authors strongly promote ACS, with its many shortcomings and compromises they are the first to admit. I devote the following paragraphs to draw out the opportunity, the crossroads, and the pressures at this juncture to better understand the background for the emergence of the ACS models and the urgency in addressing the model at this time.

A. Opportunity

Before addressing the opportunities made possible by technology, a few words about the technology itself. When referring to "technology," commentators in this field usually mean the software, hardware, knowledge, and communications infrastructure made available to a growing portion of the American public, and to almost every college student[24]: a connection (broadband, in most cases) to the Internet, a computer, or

23. *See* FISHER, *supra* note 5, at 6-10; Netanel, *supra* note 2, at 45-46.

24. This part of the discussion might call for an analysis of the "digital divide;" the fact that not all segments of the population have equal access to these technological riches and opportunities. I decide not to address this issue within this analysis, both because I do not wish to broaden the scope of this article and analysis, and because I believe this is a matter that will be resolved over time—or at least severely mitigated. For interesting data and perspective on this matter, see Amey Stone, *The Digital Divide That Wasn't,* BUSINESS WEEK ONLINE, Aug. 19, 2003, http://www.businessweek.com/technology/content/aug2003/tc20030819_4285_tc126.htm?chan=search.

other devices that allow for data storage and content use, and software that allows for quickly searching and downloading such content from the Internet, as well as transforming various forms of content available off-line into formats that could be easily subjected to the online processes mentioned.[25]

While many associate technology which promotes "digital copying" with threats to content holders or even the "death of copyright,"[26] it is essential to point out that the technology itself creates many opportunities for large media firms, copyright holders, and the general public. The large media firms, at first, can make use of such technologies to cut the costs of packaging, manufacturing, and marketing. Rather than print and burn CDs, ship them across the country, and incur other costs related to the physical manifestation of digital content, such firms can provide their content directly to consumers online.[27] The dollars saved from these improvements will not only find their way to the firms' shareholders and executives, but also to consumers who would benefit from reduced prices for media content and artists that are funded by these media firms. For users, the new media technologies present additional benefits, as the new technologies create an extensive and varied media market in which users could easily find whatever form of content they might desire, and at all times. Furthermore, the digital medium transforms users from passive content recipients to active speakers who "rip, tear and burn"[28] text, music, and video on their way to creating new and improved works, while "glomming on"[29] their own statements to existing works using the wonders of modern technology. The importance of this benefit is not confined to the commercial context. As commentators point out, the existence of the technologies mentioned and the opportunities they create enrich the public domain with new forms of expression from many new outputs which address an array of topics.[30] This outcome is extremely important to our social fabric and can promote a democratic culture with a variety of speakers and ideas available to all.[31]

---

25. See FISHER, *supra* note 5, at 13, for an additional description of this technological background.

26. For a brief demonstration of articles that carry this title, see for example Eben Moglen, *Anarchism Triumphant: Free Software and the Death of Copyright*, FIRST MONDAY (1999), http://emoglen.law.columbia.edu/my_pubs/anarchism.html; *see also* Lunney, *supra* note 15, at 813.

27. *See* FISHER, *supra* note 5, at 19.

28. This famous Apple slogan has become a term now commonly quoted by scholars aiming to demonstrate the potentials of the new technology. *See, e.g.*, Netanel, *supra* note 2, at 5-6.

29. This was a term coined by Jack Balkin to illustrate today's ability to make use of content to generate new ideas. *See* Jack M. Balkin, *Digital Speech and Democratic Culture: A Theory of Freedom of Expression for the Information Society*, 79 N.Y.U. L. REV. 1, 9-12 (2004).

30. Fisher refers to these benefits as "Semiotic Democracy." *See* FISHER, *supra* note 5, at 28-31.

31. See Balkin, *supra* note 29, at 1-2, for an additional discussion regarding this issue.

B. Threats and Crossroads

The flip side of the various benefits mentioned above, is (as many content providers are acknowledging) that the new technologies and the dynamics they make possible generate a substantial threat to the content providers' existing business models. The online exchange and distribution of content takes place at a dear price to copyright holders (or so they argue) who strongly object to these practices. They object because of the lost revenue such sharing causes and the loss of control over the uses of their content.

It should be noted that addressing the threat to content providers as a whole is somewhat misleading, as every segment within the content market is affected differently by the emergence of the Internet and digital copying. Within the music industry, it is argued that illegal file swapping causes a dip in the sales of CDs, as users will not purchase content they can now get for free online. However, there is only limited empirical evidence to support this intuitive assertion.[32] Within the various video markets, on the other hand, identifying the threat to existing business plans is somewhat trickier. Motion pictures, for instance, present an interesting test case. For many years this medium was considered an "experience good,"[33] "consumed" as part of a larger experience that involved going to the cinema with others.[34] Therefore, merely sharing such content online should not seriously affect the revenue stream from the box office. However, over the recent decade, Hollywood has discovered an additional stream of revenue in DVD sales that are proving to be extremely lucrative.[35] Therefore, it is argued that file-sharing compromises this revenue stream as well, by causing a dip in such sales.[36] File swapping might also affect additional revenue streams, such as movie rentals and the fees for broadcasting these films on various television channels at a later time.

Television shows present an even more complicated issue. Generally, in this medium, revenue is generated through advertising and financed by those paying for advertisements slotted throughout the programs. Thus, sharing such content online after "stripping" it from these

---

32. *See supra* note 8; FISHER, *supra* note 5, at 31-34; Stan J. Liebowitz, *Pitfalls in Measuring the Impact of File-sharing on the Sound Recording Market*, 51 CESIFO ECON. STUDIES 439, 440 (2005), *available at* http://www.utdallas.edu/~liebowit/intprop/pitfalls.pdf.

33. For more on this term in this context, see YOCHAI BENKLER, THE WEALTH OF NETWORKS 427 (2006).

34. Socializing with friends, eating popcorn, etc.

35. *See DVD Sales Reshaping Film Industry*, CBSNEWS.COM, http://www.cbsnews.com/stories/2003/10/20/eveningnews/main579020.shtml (last visited Oct. 20, 2006).

36. Note, however, that it is more difficult to download entire motion pictures, and is less appealing to view them on computer screens. That is why it is assumed that the damage to the motion picture is less severe than in the music context. For more on this issue, see FISHER, *supra* note 5, at 5.

advertisements, adversely affects this business model as well. Other TV-based business models generate revenues directly from viewers in the form of subscription fees.[37] Here file sharing online directly undermines the subscription business model, while allowing non-subscribers to enjoy content that was solely intended for subscribers. Yet it should be noted that both TV-based business models are currently in flux as viewers make use of TiVO and other PVRs to skip advertisements during viewing shows and even send these shows to others.[38] In view of the above, television stations are now reconsidering and revising their business models. Some are doing so by providing content online for free or at low prices through online vending sites.

Though the differences among the various media regarding the ways in which they are affected by the online realm are an intriguing topic, I will leave the analysis of such differences for future discussions. For the purposes of this article, I will address the content industry as a whole (unless indicated otherwise) while referring to the basic argument that online unauthorized file sharing adversely affects the business model and revenue stream of content providers.

The brief description above clearly points out that the threats and benefits of the Internet age lead to crucial crossroads, at which policy makers and courts must confront the fears of the content providers, and draw out rules that address the future uses of digital content online. In addition to the problems mentioned above, today's status quo presents a serious educational and legal challenge as well: many law abiding citizens are infringing on the copyrights of others and at times are subject to criminal punishment. Furthermore, in many cases the acts constituting infringement are carried out intentionally and with full understanding of their legal ramifications.[39] This is indeed an unwanted turn of events and might have serious effects on the ethical behavior of individuals in other social contexts.[40] The status quo is also allowing content companies to selectively sue users for extensive damages. These steps are frowned upon by many even within the content industry and is far from contributing to these firms' goodwill.[41]

At this juncture, the content industry is strongly pushing for a protective legal scheme that would allow them to maximize control over their content—schemes that promote the use of DRM systems. I address the DRM option, its advantages, shortcomings, and the reasons it is approached with disdain by many legal scholars and public activists, in Part

37. For example, premium channels such as HBO and Showtime generate such revenues.

38. See FISHER, *supra* note 5, at 131.

39. *See generally* CALLAHAN, *supra* note 4, at 185-88.

40. *Id.*

41. *See* Netanel, *supra* note 2, at 3; FISHER, *supra* note 5, at 126-27.

II below. However, the urgency in applying a new policy paradigm to the use of content online is also shared by those opposed to DRM, who understand as well that society is reaching a crucial crossroad concerning the regulation of online content consumption.[42]

After establishing the opportunities and the problematic crossroads we now face, I move on to address the options at hand or the paths to be taken. I start with the DRM model, move on to additional models contemplated today, and finally reach the ACS. In doing so, I describe how business and policy makers faced and dealt with similar challenges at junctures in the past, how those were resolved, and what that might tell us of the paths most likely to be taken in the future.

II. EXAMINING THE COMPETITION

In this part, I address several models set forth to suggest a path of action at the crucial crossroads we are facing. I describe some ideas briefly, while emphasizing DRM (which is the main contender at this time), and address several elements, advantages, and shortcomings of this model that will prove essential for the later segments of our analysis. Readers who are well versed in the intricacies of the recent "copyright wars" and the models set forth to resolve the problems at hand should therefore feel free to skip ahead to Part III (or Part IV if well versed in the dynamics and mechanisms of the ACS models).

A. The Foundations and Advantages of DRM

The DRM business model is premised on technological, business, and legal assumptions and requirements. I now address them in turn, while mentioning the potential benefits this model has in store.

1. Business Model

In an ideal setting, the DRM system will allow large content firms, or even individual artists, to set up a portal or website in which they would provide users with their content online for a price to be paid prior to such usage. These systems will allow consumers to select from a wide range of possible products and transactions. For instance, users could choose to pay for one or multiple uses, for uses at one outlet or at several possible locations and through various applications. In addition, users could choose to purchase the right to pass such content on to others or to modify it if they choose to do so.[43] Needless to say that without paying for the service, access would not only be prohibited and unauthorized, but almost impossible (at least for the lay computer user), as the system would be secured and encrypted. Ideally, future DRM tools will facilitate online stores that offer all forms of content, be it audio, video, or any

42. *See generally* FISHER, *supra* note 5.

43. *See* GOLDSTEIN, *supra* note 6, at 201.

reading material at a low price and a guaranteed high quality.[44] In this way, many argue, the DRM model will fulfill the promise of today's technologies and bring us to an outcome that is welcomed by both artists and consumers. Limited versions of these DRM models already exist—for instance Apple's online music store iTunes, that relies on trusted systems installed in the music player on the user's computer, and on that of the Apple's iPod (both of which comply with Apple's overall protocol).[45]

2. Technology

The business model drawn out above calls for the creation of a challenging technological infrastructure; it requires the development of software and hardware that would work efficiently, seamlessly, and with minimal malfunctions (which lead to consumer frustration and loss of revenue to the content provider). Therefore, the system would require a "trusted systems" infrastructure, which must efficiently attend to the management of the many forms of content available. When doing so, it must correctly link between the relevant content and the individuals who were provided with authorization to use it, while limiting such authorized use to the exact usage for which the consumer has paid. In addition, it must include a reliable payment system that could accurately and securely account for the purchases of content use. Finally, the DRM system must be "secured" to ensure that users cannot exceed the authorized use they purchased. Meeting this final challenging objective includes difficult tasks: assuring users cannot "hack" into the DRM systems and access content without authorization, assuring users cannot make copies of content in which they were granted limited access, or pass such content on to other individuals or applications when denied the right to do so. For meeting this objective, DRM engineers apply several forms of encryption and rely on changes and modifications to today's existing hardware and software. I will address specific technological challenges while discussing several shortcomings of this system below.

3. Law & Policy

While the implementation of DRM seems to amount to a technological and business matter, it raises difficult legal and policy questions on several conceptual layers. First, on the most basic level, there are the system's legal "nuts and bolts." Although the security of the system would be guarded by technology, it must be bolstered by specific legal protection that would allow content holders to sanction (with help from the government and the criminal system) those attempting to tamper with the system or the business model as a whole. Such legal rules will prohibit tampering with the DRM infrastructure and construction of pro-

44. *See* FISHER, *supra* note 5, at 155.

45. *Id.* at 156. For more on this, visit www.itunes.com.

grams that can do so. In other words, these rules will resemble the anti-circumvention provisions included in the DMCA.[46] Also, for the DRM scheme to work, an additional set of rules must require hardware manufacturers to comply with specific standards that would facilitate the setting of DRM in place. Such standards are needed to assure that no applications used for the consumption of content would allow for the "leaking" of such content outside the trusted environment.[47] Finally, those advocating the implementation of DRM at times argue for stricter prohibitions against unauthorized users themselves and even for "self help" remedies to be placed in the hand of the copyright holders, which resemble those existing in the "general" realm of property law[48] (such as allowing content providers to aggressively attack the P2P networks and even install viruses within the computers of "heavy" file swappers).[49]

On a more abstract level, the endorsement of DRM systems represents adherence to a specific jurisprudential perspective as to the role and strength of the protection amounted to intellectual property in general and copyright specifically. According to Fisher, DRM models (as well as other models that provide extensive protection to the copyright holder) are premised on a simple, yet dangerous, policy assertion: that copyright in musical and video works must be protected almost to the same extent as other "strong" property rights (such as rights in real property), and therefore include (among others) the rights to exclude all unauthorized actions and take aggressive steps to assert these property rights (which include the self help measures mentioned above).[50]

However, as many point out, this policy assertion is problematic. Musical and video content has indeed been afforded property protection to promote the continuing creation of new content.[51] Yet the accepted theory behind this legal rule is that absent such protection, individuals would "free ride" and use the newly created content extensively, without the consent or compensation of the author (a phenomenon broadly associated with most "public goods"). This would lead content creators to apply their talents elsewhere, where they could fully reap the fruits of their hardship and labor,[52] and thus would lead to a sharp decline in the

46. For more on this issue, see Lunney, *supra* note 15, at 823-45; FISHER, *supra* note 5, at 87-98.

47. Recently, the FCC attempted to put in place regulation requiring hardware manufactures to comply with specific standards that would allow for the tracking of content use. This attempt, referred to as the "Broadcast Flag," has been struck down by courts that found the FCC to exceed its authority by setting these regulations in place. For more on this issue, see generally Susan P. Crawford, *The Biology of the Broadcast Flag*, 25 HASTINGS COMM. & ENT. L. J. 603 (2003).

48. *See* FISHER, *supra* note 5, at 150.

49. *Id.*; *see also* Netanel, *supra* note 2, at 18-19.

50. *See* FISHER, *supra* note 5, at 143.

51. *Id.* at 199-203.

52. *See* Litman, *supra* note 15, at 30 (noting somewhat cynically that with lower levels of compensation, artists might opt to become investment bankers).

quality and quantity of new content in the public sphere.[53] This rationale is not identical to those upon which other property rights are premised. Indeed, copyright differs from other "classic" property rights, such as real property or chattel, in several crucial ways. First, it is non-rivalrous; it allows users to make use of such "property" simultaneously without one individual's use degrading the ability of others to make use as well. Furthermore, in the digital age, the marginal cost for making perfect copies is close to zero. Therefore, the intangible goods protected by copyright have always been provided with a narrower realm of protection than holders of other forms of property.[54] For instance, the copyright holders' right to exclude others is limited by the right of others to engage in the "fair use" of such rights (even without the right holders' consent). In addition, copyright is limited by scope (mere ideas are not protected) and by time.

The tension between content firms' perspective, that strive to command full excludability and other rights usually paired with "property," and critics of this position that stress differences between the underlying rationales for legal copyright protection (which includes severe limitations on this property right) and those of other property rights lead to several critiques of DRM we will examine below.

B. DRM: Disadvantages and Shortcomings

Though the DRM business model has several appealing traits and is embraced by many content holders, it has been subject to a great deal of scrutiny and criticism.[55] These critiques have led the ACS scholars to advocate their somewhat radical solutions as an alternative so to avoid the problematic consequences of DRM. I will briefly address these critiques, while referring to the technological, business, and legal assumptions and requirements mentioned above.

1. Technology & Business

From the technological perspective, many technologists argue that DRM systems cannot and will not meet the ambitious objectives drawn out above (a secure and sustainable trusted system). This is because it is not feasible to implement such systems, as the security challenges DRM presents are too great, and the risks and vulnerabilities are too varied.[56] Here, technologists commonly refer to the experience the industry has had with limited DRM schemes, such as the DVD player, and the speed

53. Another rationale (that is not as salient in U.S. jurisprudence) is the protection of the "moral rights" of the author, who has a right to maintain the integrity of the work she authored, and receive credit for its use. This rationale does not dominate US copyright law and policy, and therefore is not discussed in this article. For more on this issue, see GOLDSTEIN, *supra* note 6, at 137-40.

54. *See* Netanel, *supra* note 2, at 30.

55. *See generally* FISHER, *supra* note 5.

56. *Id.*

and ease with which these schemes have been "hacked."[57] Furthermore, a successful DRM system must be flawless and would only be as strong as the weakest security system on any application connected to the network; a flaw at one point within the system would lead to a "leak" of high quality content into the illegal file-swapping networks.[58] Since the DRM system must be implemented in every application used to consume content, achieving such a high and reliable standard seems somewhat far-fetched.

However, these arguments have not gone without a response. The common counter argument states that although the system cannot be entirely secure against various breaches, the existence of the security measures and the legal sanctions for their breach would sufficiently deter most attempts to hack and infringe. As a sufficient number of consumers would make exclusive use of the legal DRM applications to access content, this business model would still prove profitable and successful.[59]

Additional critiques of DRM from the technological perspective argue that the implementation of DRM systems in all media players would cripple these applications (both in terms of hardware and software), slowing them down and blocking the use of the full potential of the technology. Others argue that setting a DRM standard would adversely affect competition in various hardware and software markets and allow those setting the standards to box out competitors.[60] As it is quite difficult to assess these arguments at this point of time, I will leave them for future inquiries.[61]

2. Law & Policy

DRM has come under heavy fire from legal scholars and social advocates concerning a variety of topics: the fact that these systems (and the legal infrastructure they require) can potentially impinge on the public's right to engage in fair use of copyrighted materials, lead to price discrimination, change the Internet's open architectural structure, and

---

57. For a discussion of the vulnerabilities of the SDMI technology (especially with regard to the work of Professor Felten in proving the systems weakness), see GOLDSTEIN, *supra* note 6, at 177-81.

58. Note that the existence of content of high quality is a somewhat rare commodity within these networks, as many files are corrupted or partial. This is why such leaks will be extremely harmful as they guarantee access to quality content.

59. *See* FISHER, *supra* note 5, at 156-58.

60. Regarding the anticompetitive elements of DRM in the way it could stall competitors, see Timothy B. Lee, *Circumventing Competition: The Perverse Consequences of the Digital Millennium Copyright Act*, March 21, 2006, http://www.cato.org/pubs/pas/pa564.pdf.

61. For a most recent analysis of DRM and its vulnerabilities (especially in terms of its crippling effects on the systems it uses), see J. Alex Halderman & Edward W. Felten, *Lessons from the Sony CD DRM Episode*, CENTER FOR INFORMATION TECHNOLOGY POLICY, DEPARTMENT OF COMPUTER SCIENCE PRINCETON UNIVERSITY (2006), http://itpolicy.princeton.edu/pub/sonydrm-ext.pdf.

intrude on the privacy of individuals. I briefly address these arguments, in turn.

a. Fair Use and Means of Self Expression

The DRM models allow media firms to control the exact use of the content they include in their repertoire, while providing a variety of specific usage rights upon request. However, usage beyond the specific authorization provided to every user would be practically impossible, as it would be blocked off by the system's security measures. This reality sharply differs from the one we have today in which copyright holders usually cannot rely upon technological protection, but are confined to the protection and enforcement of the law to uphold their rights. The law, as mentioned above, does not provide for protection against unauthorized uses at all times, but includes important exceptions limiting the time and scope of copyright as well as the exception for fair use—the ability of users to access and modify content without authorization when meeting specific criteria set out by the law and established by the courts. DRM systems, therefore, can provide content owners with *de facto* rights that exceed those provided to them by law.[62]

When contemplating this potential future outcome of DRM implementation, several IP scholars assert that it is extremely problematic and therefore DRM schemes should be rejected, or at least changed.[63] They argue that copyright law, as it stands, sets a delicate equilibrium between sufficient property rights to the authors (as well as performers, etc.) and protection of the basic right and liberties of other users, creators and the broader public.[64] Specifically, they assert that the "fair use" exception provided to users promotes the distribution of ideas and allows individuals to stand on the shoulders of giants when constructing their arguments and convey their message more effectively. Thus, the fair use exception is closely tied to the fundamental concept of freedom of speech and expression. With DRM systems in place, the "fair use" exception would be effectively eliminated, thus harming important social interests of users, artists, and society in general.[65] This inability to make fair uses of digital

62. Also note that while making use of these technologies, content providers are able to control and limit many forms of personal uses, which were not considered legal, yet were outside the realm of copyright law enforcement for practical reasons. For more on this issue, see Jessica Litman, *Lawful Personal Use*, JOHN M. OLIN CENTER FOR LAW & ECONOMICS (August 2006), *available at* http://papers.ssrn.com/sol3/papers.cfm?abstract_id=926575#PaperDownload.

63. One might argue that the mentioned limitations to the fair use exception for copyright protection are not a result of legislation, but of technological steps taken by the content holders, and therefore are irrelevant to this policy discussion. The clear response to this argument is that these new extended rights are indeed protected by other laws that prohibit users to make unauthorized uses—such as the DMCA, state and federal anti-hacking laws and additional laws and regulations that might be put in place to facilitate the DRM scheme. *See* 17 U.S.C.A. § 1201 (West 2006).

64. For a discussion of this issue, see generally Tushnet, *supra* note 20.

65. A common example in this context is of a student preparing a school project about the Holocaust. The student wishes to use a graphic excerpt from the film "Schindler's List" but cannot

content and the new technological tools that allow for creating, modifying, and editing content with great ease is a key example of opportunities missed and therefore a road that must be avoided.

At first blush, this concern regarding DRM and fair use need not lead to the overall rejection of these models, but rather calls for rules to assure that the equilibrium mentioned would be maintained even with the adoption of DRM. For instance, such rules might state that for DRM systems to benefit from the legal protections addressed above, they must incorporate a "fair use" exception within the system. They would do so by providing users with free access to content given the fulfillment of specific factors that reflect today's legal understanding of the "fair use" doctrine.

However, as Burk & Cohen explain,[66] it is almost impossible to establish *ex ante* (namely, at the time the DRM systems are structured), what would constitute a "fair use" in practice. In other words, it is extremely difficult to decide at this early juncture when and to what extent such uses should be permitted. The criteria as to what is and should be considered as "fair use" are abstract and ever-changing, and it would be nearly impossible to translate them into a set mathematical algorithm. Reality will continue to produce instances that call for recognition of the "fair use" exception but were not preconfigured into the system. Therefore, the *ex ante* setting of the "fair use" exception in "code"[67] would almost always be applied too narrowly and thus impede on the important interests of the public.[68]

DRM advocates, however, offer an additional response to the "fair use" challenge. They argue that existing loopholes in the DRM system which allow for unauthorized uses and cannot be controlled by the DRM system would in fact allow users to exercise their right to fair use. One famous loophole is the "Analog Hole," which refers to the assumed inability of DRM systems to block users from making copies of protected content through various analog means.[69] Through this "hole," users

---

do so because the content is locked by DRM protection (assuming she owns a DVD copy, for instance). For more on this example, see EFF Post-Hearing Comments Requesting Exemption of DVDs from Section 1201(a) (June 2000), *available at* http://www.eff.org/IP/DMCA/20000623_eff_dmca_dvd_comments.html.

66. *See generally* Dan L. Burk & Julie E. Cohen, *Fair Use Infrastructure for Rights Management Systems*, 15 HARV. J.L. & TECH. 41 (2001).

67. For a famous explanation regarding the differences and problems in setting legal rules in digital code, see LAWRENCE LESSIG, CODE: AND OTHER LAWS OF CYBERSPACE 6, 89 (1999).

68. Burk & Cohen, *supra* note 66, at 65. Burk & Cohen go even further to suggest that should a DRM system be put in place with insufficient "fair use" embedded exceptions, individuals should be provided with a "right to hack" the systems and use the protected content without authorization, if such uses amount to "fair use" according to the legal standards. *Id.*

69. In this context, it should be noted that there might be a disparity between audio and digital content (which I have addressed almost throughout the analysis). While there are many ways to "capture" audio content using analog means without losing quality, video might pose more of a problem.

should be able to meet the important social objectives mentioned. Therefore, the objection to DRM on the basis of the inability to provide for a robust "fair use" right would be resolved. Yet the "Analog Hole" argument must face several challenges. First, DRM architects are striving to shut this "hole," or at least limit it to content of very low quality (use of which would not allow consumers to meet the "fair use" objectives mentioned).[70] Second, one could argue that for meeting the important objectives the "fair use" doctrine promotes, merely allowing for copying through the analog hole is insufficient. This is because the tools for making digital copies of analog outputs are too costly, they require a high level of sophistication, and above all they might provide the relevant content in low quality.[71] In view of the breadth of this matter, I will not resolve the question as to the extent of the analog hole and its relevance to this debate and leave that for future analysis. At this point, I merely conclude that the analog hole is far from being a "silver bullet" which will mitigate all "fair use" concerns in a DRM environment.

In summation, it is important to note that the fear of the extensive control over content, which is premised on the arguments stated above, is one of the leading reasons quoted for preferring the ACS solutions over the DRM ones.[72]

b. Price Discrimination

An additional objection to applying DRM concerns is the pricing schemes this model enables. These models allow for the construction of elaborate mechanisms for pricing different services differently. Yet the DRM systems potentially allow content providers to go even further and charge different users different prices for the *same* product or service. Here, the differentiating factor would be the individual and not the service at hand. Content providers might charge higher prices from consumers that have the ability to pay more. They might also try to overcharge when they believe a consumer has a special need for their content and would therefore be willing to pay a higher price at that time. When engaging in such price discrimination schemes, content providers will tailor their prices using personal information they previously collected

70. For more on this issue, see Crawford, *supra* note 47, at 618.

71. This would occur due to the transfer of the content file from analog to digital. *Id.* I thank Phil Weiser for his insights regarding this issue.

72. In addition to these arguments, on the jurisprudential level, proponents of the DRM models can argue that the "fair use" exception should not be considered as a right – but is merely a defense against infringement claims in specific instances. In other words, the fair use doctrine in copyright does not provide individuals with a "right" to use content – and therefore is irrelevant in this context. Furthermore, they argue that free speech in the First Amendment context is an irrelevant argument to the issue of fair use. The issue at hand does not involve a state action, but one of private parties. A discussion of this issue exceeds the limited confines of this article.

about their users, such as data concerning the users' place of residence, past content consumption patterns, overall financial standing, etc.[73]

Even though "price discrimination" has a sinister sound to it, such a pricing dynamic features several benefits addressed in the literature.[74] Mainly, it allows for pricing in a range that is closer to the specific user's actual ability to pay. Therefore, although some users would be charged with higher prices, others would possibly[75] be charged with lower prices they could now afford.[76] However, many scholars argue that such benefits are outweighed by the model's potential detriments.[77] They argue that these models will lead to a transfer of consumer surplus to the content firms running the DRM applications. There is no guarantee such funds would be shared with consumers or other artists, but would be plainly shared with the firms' shareholders and executives. Others argue that these schemes will create consumer concern and unease given the omnipresent surveillance and the ongoing analysis of personal information that is required to facilitate this model (an issue that ties into the privacy concerns addressed below).[78] Knowledge of such ongoing surveillance and its effects on pricing might also change the way in which users conduct themselves online, leading consumers towards cautious and restrained behavior—another undesired social outcome of implementing DRM.

c. The Internet Architecture and the End-to-End Principle

Other concerns regarding the implementation of DRM systems focus on their effect on the Internet and future innovations within its realm. The Internet's tremendous growth over a short period of time has been commonly attributed to the openness of its infrastructure, which allowed anyone to contribute and develop new applications without the need to request permission or receive source codes from controlling entities. In other words, the Internet has thrived thanks to having the "intelligence" of the system at the end users' side, while leaving the pipes "dumb."[79]

---

73. *See* FISHER, *supra* note 5, at 165.

74. CARL SHAPIRO & HAL R. VARIAN, INFORMATION RULES: A STRATEGIC GUIDE TO THE NETWORK ECONOMY 39 (1999).

75. The text cautiously indicates this would only possibly happen, as there will be instances in which market forces will allow vendors to pocket the entire surplus from such pricing schemes without lowering prices for specific consumers. *Id.*

76. On the issue of data mining and the use of personal information to facilitate this business model, see FISHER, *supra* note 5, at 167-68; *see generally* Tal Z. Zarsky, *Mine Your Own Business: Making the Case for the Implications of the Data Mining of Personal Information in the Forum of Public Opinion*, 5 YALE J.L. & TECH. 1, 24-25 (2002-2003).

77. FISHER, *supra* note 5, at 168-69.

78. *See generally* Julie E. Cohen, *DRM and Privacy*, 18 BERKELEY TECH. L.J. 575, 576-77 (2003) (explaining privacy interests, perceptions of privacy, and DRM intrusion).

79. FISHER, *supra* note 5, at 171 (citing Mark A. Lemley & Lawrence Lessig, *The End of End-to-End: Preserving the Architecture of the Internet in the Broadband Era*, 48 UCLA L. REV. 925, 930-31 (2000)).

These attributes are commonly referred to as formulating the end-to-end (E2E) principle.[80]

Broad implementation of DRM systems potentially threatens future compliance with these attributes. The successful implementation of a DRM scheme will feature a trusted system embedded within every piece of hardware connected to the web. Therefore, developers of new applications (be they of software or hardware) must comply with the standards the DRM systems use, which would be proprietary and at times unavailable, and would not be free (as they are today) to develop applications in an open environment. Therefore, with DRM, the central system running this scheme would violate the end-to-end (E2E) principle, as the existence of a central system takes the intelligence of the network out of the hands of the end users, and thus stalls Internet innovation. Therefore, the violation of the E2E principle and its effects serves as another reason to oppose the broad implementation of DRM.[81]

d. Privacy

Finally, DRM systems take their toll on users in terms of their ability to maintain their privacy. As Julie Cohen points out, the implementation and operation of a trusted system that coordinates the DRM schemes requires the collection, storage and use of vast amounts of personal information.[82] This personal information includes data concerning the content consumed by individuals, and the times and places they did so. This data could be later compiled to form a revealing profile for every user.

Given the sensitive information they include, the existence of such profiles and databases create several privacy-related concerns: fears that they would be passed on to the government, used improperly by the content providers for marketing (or other commercial objectives), or sold on the active secondary database market. In addition, several scholars mention fears that such data would possibly fall into the hands of unwanted parties in view of improper security measures taken by the database holders.[83] All these reasons and concerns add to the overall discontent with the DRM solution.

---

80. The importance of maintaining the E2E principle and the question as to what extent regulators should intervene in maintaining the Internet's open architecture is now hotly debated in the context of the "Net Neutrality" debate—whether ISP's should be permitted to discriminate among content providers when delivering the Internet connection to the end users (and thus violate the E2E principle). On this issue, see Christopher S. Yoo, *Would Mandating Broadband Network Neutrality Help or Hurt Competition? A Comment on the End-to-End Debate*, 3 J. ON TELECOMM. & HIGH TECH. L. 23, 38 (2004).

81. FISHER, *supra* note 5, at 172.

82. Cohen, *supra* note 78, at 584-85.

83. *See* Halderman & Felten, *supra* note 61, at 1-2, 26 (discussing security risks in trusted systems).

C. Summing Up DRM—Learning of the Future from the Past

Thus far, I have demonstrated how DRM models permit copyright holders, upcoming artists, and even consumers to capitalize on the benefits of new technological innovations. In addition, I presented the many shortcomings that have led policy makers, scholars, and even business entrepreneurs to oppose DRM and therefore search for other solutions, which I address in the following chapters. Yet even in view of the business-related, technological and legal difficulties and shortcomings DRM portrays, there is a good chance the media industry leaders will remain unconvinced by the above arguments, and that the DRM model will still prevail. This concern is fueled by a glance at the recent history of choices made by legislators and courts on the one hand, and content firms' executives on the other when dealing with new technologies and the threats to the existing business models of copyright content distribution. The history of society's dealings with such challenges is one of capture (in the hands of content industry) and misjudgment (again by the content industry). In the next few paragraphs I will address the concern that DRM will prevail after all and the historical background and milestones confirming it, which include instances in which regulators and courts accepted the positions of content companies regarding the protection of their content. In addition, I will mention several instances in which content firms moved to block innovative technology and business models that required they partially concede their control over content, even when these models would have proven beneficial to consumers, artists, and at times the content firms themselves.

I start with legal responses to the technological innovations which came with the emergence of digital content, communications, and, thereafter, copying. Here, the ACS scholars (as well as many others addressing these issues) point out that the content industry has been using its influence over legislators to strengthen their hold over their assets by expanding the legal protection afforded to copyright (for instance by extending the time limitations for the lapse of copyright protection), and the creation of para-copyrights (additional ancillary rights which protect the core copyright).[84] A frequently mentioned example of this latter phenomenon is the adoption of the DMCA's anti-circumvention provisions that provide legal protection from attempts to tamper with technological systems put in place by media firms to protect their content from unauthorized uses.[85] In addition, content firms, by way of expensive lawyering, were able to persuade courts to expand the rights they may exercise regarding their content, thus blocking unauthorized activities (and presumably passing legal expenses from their legal battles on to their customers). Here, courts accepted theories of secondary liability for

84. Balkin, *supra* note 29, at 17-18.

85. *Id.*

copyright infringement to block the actions of entities that facilitate the unauthorized exchange of content files. For instance, the content industry has argued successfully that ventures facilitating (in the case of *A&M Records v. Napster*[86]) or providing the software tools for file sharing are indeed subject to secondary liability (or inducement in the case of *Metro-Goldwyn-Mayer Studios v. Grokster*[87]). On the basis of this history, it is quite likely that the content industry would be able to successfully influence legislators, regulators, and courts to accept the legal rules required to set the DRM models in place.

In addition and as mentioned above, many scholars fear that DRM would be implemented regardless of its many shortcomings which are harmful to users, artists, and the media firms themselves. Yet these shortcomings would be overlooked by the content firms who will opt for DRM while aiming to sustain control over their content. In the past, large media conglomerates have often opted against new strategies that would provide less control, but reap other benefits. For instance, scholars often refer to the MyMp3.com venture, which was forced out of business by the large media conglomerates after offering a service that allowed consumers who purchased CDs to access the songs it included at any location through the website's database.[88] This service, arguably, would have provided users with additional convenience and perhaps even promoted CD sales (as they could now be enjoyed with greater ease). Yet the potential benefit to all parties involved did not deter content firms from burying this venture. In addition, the ongoing attacks of the media industry against file sharing networks are arguably another example within this broader pattern of behavior. Several artists have been arguing that the availability of content has led to many benefits, especially for those artists who have difficulty in self promotion and in gaining access to a broader audience. These artists have welcomed the file swapping phenomenon[89] and argued that it led to increases in ticket sales to live performances, and in some instances even to increased CD sales.[90] Again, the content industry has ignored these benefits and voices, and the industry moved to silence this form of uncontrolled distribution.

In view of these historical trends, the ACS scholars argue that the content industry would move to implement DRM without hesitation and

---

86. 239 F.3d 1004, 1013 (9th Cir. 2001). For a description of this case, and an analysis that shows that the final outcomes of this case and others were not clear cut, see FISHER, *supra* note 5, at 116-23.

87. 125 S. Ct. 2764 (2005).

88. FISHER, *supra* note 5, at 99-101.

89. For example, see John Borland, *Musicians Launch National Anti-Napster Campaign*, C|NET, July 11, 2000, http://news.com.com/2100-1023-243021.html?legacy=cnet (regarding Courtney Love).

90. *Id.*

consideration of benefits both to them and others. For this reason, they believe alternative models must be forcefully promoted to counter the historical force of the content industry.

D. Other Business Models and Suggestions

In addition to the DRM and ACS models (which I address at length below), there have been other suggestions for resolving the challenge of compensating copyright holders while promoting various social objectives, in the digital age. I now briefly address some of these proposals and models. However, for the balance of this article, I examine and consider the DRM model only, as it is considered the most serious contender and had been openly embraced by many parts of the content industry.[91]

One radical and therefore somewhat theoretical solution calls for the elimination of copyright protection online. In the online realm, several scholars argue, there is no need for copyright law protection to promote content creation and meet other social and individual objectives. The Internet's ability to facilitate worldwide distribution at nearly zero marginal cost, should allow content creators to rely upon other forms of compensation. These would more than substitute for the "lost" compensation they would have received from online users. Such sources of income might come from the benevolence and gifts of consumers,[92] profits derived from live performances, merchandising, or the sale of CDs offline. However, given the dominance of the media firms and their influence over regulators, it is hard to believe such a policy would be accepted.[93]

Another option to resolve the issues at hand (as addressed by Fisher)[94] calls for a "regulatory solution," where the government would use its authority to directly intervene and set the compensation for the authors of works consumed online. Although this solution seems awkward at first, Fisher points out that industries and markets that bear some resemblance to content markets are or were closely regulated.[95] Fisher also notes that several segments of content markets are already subject to heavy regulation.[96] However, it is doubtful that this solution would be acceptable as it allows for government to intervene in society's choices

---

91. For a summary of the broad array of solutions currently contemplated, see Yu, *supra* note 20, at 698.

92. A famous example and experiment of using benevolence to generate compensation was conducted by Mr. Stephen King. For an analysis of this incident, see Kylie J. Veale, *Internet Gift Economies: Voluntary Payment Schemes as Tangible Reciprocity*, FIRST MONDAY (2003), http://www.firstmonday.org/issues/issue8_12/veale/.

93. *See* Burk & Cohen, *supra* note 66, at 48-49.

94. FISHER, *supra* note 5, at 183, 186-95.

95. *Id.* at 181-84.

96. *Id.* at 184-85.

regarding speech and content—an interaction that would be broadly criticized (rightly so) as unhealthy and unwanted.[97]

Another solution coming from the business realm provides users with access to a vast repertoire for their own personal use, at a fixed fee paid on a monthly or annual basis. The business models these companies (such as iMesh and the "new" Napster)[98] utilize present some of the shortcomings and challenges of the DRM model (such as the need for a secured system to block leaking) while limiting others (such as the fear of price discrimination and to a certain extent, privacy). It remains to be seen whether consumers would accept these business models, which offer a smaller repertoire than the one available within the illegal peer-to-peer (P2P) networks, and for obviously a higher price.[99] However, when using this model, consumers are assured of the quality of the content and the legality of their actions.

Finally, other business models appear to have conceded to the fact that today's distribution networks are uncontrollable. Therefore, rather than battling them, they choose to take advantage of these networks and benefit from the broad distribution they facilitate. For instance, advertisers develop prime content which is also intended to promote a brand (by hidden or even blatant commercial content), and release it within the network, while hoping it will generate interest and traffic.[100] This dynamic would be interesting to track during the next few years and might be indeed well suited for some works. However, artists whose content will not mesh well with sponsorships or embedded promotions cannot rely upon this model for proper incentives, and will be looking to other options for compensation.[101]

In summary, a great deal of academic and other writing addresses the digital market's promises, threats and some solutions. At this time, DRM seems to be the industry's favorite but a nightmare for many others. We will now address an additional option which competes with DRM—the ACS models.

---

97. Litman, *supra* note 15, at 41-42.

98. For more information, see iMesh, http://www.imesh.com (last visited Oct. 30, 2006), or Napster, http://www.napster.com (last visited Oct. 30, 2006).

99. While the abovementioned services are not expensive, the cost of illegal file sharing is still near zero, with the additional cost of the risk of downloading a low quality copy and the slim chance of being sued by the media firms. The cost of the illegal service is of course zero, and the risk of being prosecuted is very slim.

100. A famous example is the American Express campaign featuring Superman and Jerry Seinfeld that was released with great success throughout the file sharing networks, *see* Maria Mandel, Partner, Executive Dir. of Digital Innovation, OglivyInteractive, General Session on Consumer Behavior in a Digital World at the Summit on Intellectual Property and Digital Media, The Cable Center, University of Denver (May 22, 2006).

101. Netanel, *supra* note 2, at 76.

III. THE ALTERNATIVE COMPENSATION MODELS

A. Elements of the ACS Models

Although the ACS models' specific elements are quite complex, the models' overall objectives are simple: They strive to fairly compensate copyright holders for the use of their works online, without directly charging for the use and enjoyment of such works. In addition, these models strive to do so while legalizing today's illegal yet widespread file swapping activities (the models differ as to exactly what elements should be legalized, as explained below). To meet these objectives, the ACS models require careful accounting for the actual uses of content online and thereafter distributing funds that were specifically collected for this reason to the relevant copyright holders. This elaborate task would be conducted by a specific governmental agency—preferably within the Copyright Office (but for this analysis I will refer to it as the "administrating agency"). The scheme has three main components: registration, collection of funds, and the distribution of funds to the right holders (a component which includes the process of assessing the relative usage of works online).[102] I will briefly explain what each component entails, in turn. It should be noted that the ACS models have been recently suggested in several variations by various scholars,[103] although the general theme is mostly the same. For this analysis, unless indicated otherwise, I refer to the model presented by Fisher, which is perhaps the most detailed and comprehensive.

1. Registration

The starting point for implementing this model (as well as for the flow of content and information within it) is the registration process. At this point, copyright holders interested in participating in the ACS register a specific work as their own online, and receive a specific code. This code is to be "watermarked" into the relevant work in its digital format (be it an audio or video file). From that point, the work could be released online, and its subsequent online uses would not require consent.[104] However, thanks to the registration and watermarking process, the copyright holder would be accredited for subsequent uses of the work, and compensated accordingly.

2. Collection of Funds

The next point would be the collection stage at which the administrating agency must extract sufficient funds from the public so as to properly compensate the creators for the online use of their content. This

---

102. For a description of the various ACS models, see Litman, *supra* note 15, at 32-33.

103. *Id.* at 32-34.

104. *See* FISHER, *supra* note 5, at 203.

stage presents several key questions in its implementation, especially as to "how much" and "how." When addressing the "how much" question, the ACS scholars strive to maintain the status quo. In other words, the collection process is structured to assure that the overall level of compensation through the model would reflect the losses content owners would incur due to the legalizing of online content file sharing and streaming.[105] To correctly estimate the status quo and the extent of the losses incurred, the ACS scholars engage in extensive calculations to draw out the overall revenue content markets generate, the actual and predicted harm from the illegal online activity and the percentage of the overall revenue that is and would be lost from such online actions.[106] While the actual sums and percentage rates differ among scholars and sub-markets,[107] the results the ACS scholars present lead to a very large sum of over two billion U.S. dollars per year.[108]

After establishing that substantial resources are required to create and maintain a fund for the full compensation of right holders, the models turn to the question as to "how" these sums should be gathered. The simplest response calls for collecting such sums as part of the general federal tax system. While this solution has several advantages,[109] its overall and overwhelming shortcoming is that it seems politically implausible. No administration would raise taxes to meet this objective and risk the public backlash usually associated with tax hikes. In addition, such changes in general tax policy would lead to a public outcry stating the fact that the "tax dollars" of individuals who do not use the Internet at all (or only rarely do so) are cross-subsidizing the increased (not to mention obsessive) content consumption of others.[110] In view of these anticipated difficulties, the ACS scholars suggest that the funds should be raised by a levy to be set on selected products and services that are closely associated with the consumption of online content. Here, they

105. *Id.* at 208; Netanel, *supra* note 2, at 46-47.

106. For instance, see Netanel, *supra* note 2, at 60-67.

107. According to Fisher's calculation, we should account for a loss of 20% of revenue in the music market (which pertains mainly to CD sales) and 5% of the video market, whose most dominant component is DVD sales (but also DVD rentals, premium channels and the growing demand for V.O.D and pay-per-view). FISHER, *supra* note 5, at 209-14. Netanel's calculations lead to somewhat different results (25% loss of revenue in the audio market; 7% in the video market). Netanel, *supra* note 2, at 61.

108. Clearly, it might appear that this segment of the analysis is extremely shaky and might even appear to some as mere guess work. The authors here attempt to assess a future market reaction based on information on which economists cannot agree upon even today (regarding the question as to the effect of online file sharing on the content markets). The authors are aware of this line of criticism—and respond that this is merely a starting point, and the model as well as the sums that must be collected (and thereafter distributed) will be updated on a continuing basis, in accordance to updated information from surveys and the industry. *See* FISHER, *supra* note 5, at 209-15; Netanel, *supra* note 2, at 65-67.

109. Fisher points out that adding this amount to the general tax burden is unlikely to cause any radical distortions, and would be relatively simple to implement. FISHER, *supra* note 5, at 216.

110. *See id.* at 217.

convincingly argue that public opinion would be more likely to accept this limited taxation scheme in the form of a levy, as its impact and effects on individuals that are removed from online content consumption would be minimal.

To apply the levy, policymakers must establish the tax base (namely, which products and services would be subject to the levy), and the actual level of taxation on products and services that are part of this base. Both tasks create difficult policy and empirical questions, and the ACS scholars provide several models to resolve them. The specific ways in which the base and the level of taxation are formulated need not concern us at this time[111]—not because they are uncontroversial but because they would be subject to change in view of updated information streaming to the administrating agency from the industry and from timely reports examining the ways in which content is used and consumed online. In his book, Fisher draws out an initial taxation base (which would be subject to change)[112] and sets the levy at about 11.4% (as opposed to about 4% according to Netanel).[113] In other words, consumers would be charged an additional fee every time they purchase products and services that are part of the tax base. In "return," users would be permitted to make use of an extensive library of content that is available online, in any way they might desire. They would be permitted to listen or view the content, and even include it in digital forms of content they produce so long as they register their work and include a reference to the content they made use of.[114]

3. Distribution of Funds to Right Holders

The final component of the ACS scheme is the *distribution* of the funds to the right holders of the relevant works.[115] Again, the initial (and modest) objective of the ACS scholars is to maintain the status quo when shifting to the ACS model. In other words, the objective is to provide rights holders with proper compensation for the revenue lost when legal-

111. Generally, Fisher draws out four categories that would include the tax base: (1) Equipment that facilitates digital copying—such as CD and DVD burners; (2) Equipment for digital storage—such as blank CDs and flash memory; (3) Internet access providers (although Fisher believes the levy should be limited to broadband only); and (4) Systems and software that facilitate file sharing. *See* FISHER, *supra* note 5, at 217. Netanel suggests adding "dial up" Internet connections to the levy as well, and would include a levy on the purchase of computers themselves. Netanel, *supra* note 2, at 60-62. For the analysis of the tax basis, see FISHER, *supra* note 5, at 217; Netanel, *supra* note 2, at 60-62.

112. Especially in view of recent changes in the ways individuals connect to the internet – i.e. through the use of WiFi technology. *See* FISHER, *supra* note 5, at 251.

113. The disparity in these figures stems from the differences in defining the tax base, as mentioned above.

114. This issue leads to the complicated "derivative works" issue and the problem of accounting for several authors of a single work. The ACS Scholars attend to this matter at length. *See* FISHER, *supra* note 5, at 234-35; Netanel, *supra* note 2, at 57.

115. *See* FISHER, *supra* note 5, at 202.

izing online file sharing and downloading. Here, the ACS differ from more ambitious schemes, that attempt to restructure the way in which artists should be compensated for content production.[116] Yet even the ACS scholars' limited goal presents serious challenges. First, how would the administrating agency know what part of the overall fund every copyright holder should receive? In today's content markets, the public signals its content (or discontent) with various works by paying for them. Since the ACS models involve indirect compensation there is no such direct payment per use to rely upon. To resolve this difficulty, the models turn to a substitute: information about another scarce human resource—attention. Namely, the models call for allocating the funds in accordance to the way in which consumers allocate their attention towards specific works—while providing greater compensation to authors of works that were "experienced" more times.[117] To achieve this, the models call for the construction of elaborate "counting" mechanisms that would allow the administrating agency to count the uses of content online, sum them up and by the end of every year provide a full report as to the number of times each work was used. After taking into account the overall number of works used and the size of the fund for every given year, the copyright holders receive a check from the government that constitutes their "share" of the overall fund collected through the levy.[118] The construction of these counting mechanisms presents many technical and policy challenges, and as these issues stand in the core of my critique, I will address them in greater length in a subsequent part of this article.

Beyond the three components addressed, the model requires several adjustments in the existing legal regime. First, it requires changes in current copyright laws so that the downloading and streaming of content online will not constitute copyright infringement.[119] It should be noted that the ACS scholars disagree on this point. While Fisher believes all

116. Fisher mentions several theories according to which compensation should be distributed (such as models premised on voting), but concludes that in the first stage, it is best to simply maintain the status quo and thus base compensation on usage. FISHER, *supra* note 5, at 234. For more on competing voting schemes, see Peter Eckersley, *Virtual Markets For Virtual Goods: The Mirror Image Of Digital Copyright?* 18 HARV. J.L. & TECH. 85, 111 (2004).

117. At this point, the model somewhat differs from the "brick and mortar" reality, in which the "signaling" usually takes place only once—at the time the content is purchased. However, the online realm is somewhat inappropriate for this form of measurement, and therefore the authors' decision to "count" actual uses of the content, as opposed to its mere "download" is indeed correct. Online, users tend to download a vast amount of content, yet use a minuscule portion of it. Therefore, compensation per download will provide a biased result and will not reflect actual trends of content usage and appreciation.

118. FISHER, *supra* note 5, at 202.

119. In addition, Fisher addresses the possibility that content owners would argue that the shift to the ACS model constitutes a "taking." FISHER, *supra* note 5, at 248-49. Fisher explains why these arguments would probably be rejected, or would not lead to any meaningful compensation for the content owners. *Id.* In addition, Fisher mentions international treaties that might conflict with the ACS model. *Id.* I will not address the international aspect of the ACS model in this article.

online use should be permitted, Netanel argues that only non-commercial uses should be allowed.[120] Of course such changes should also address (and forbid) the self help measures that content companies are trying to apply at this time, and have no place in a regime in which the model has been facilitated.[121] Second, some ACS scholars argue that with the model in place, the legal protection of DRM-like secured systems should be repealed, so as to encourage content providers to participate in the model's registration and compensation schemes (by making the alternatives seem less attractive).[122] Lastly, the actual implementation of the model would require regulatory intervention concerning the various components addressed above. Regulation must address the registration process and the role of the governmental administrative agency. Moreover, it must set rules regarding the mandatory levy, how it would be collected and ways to limit its evasion. Finally, regulation must address the "counting" process, which (as I explain below) would probably include rules concerning the mandatory installment of counting systems meeting a specific standard on all machines.

B. Similar Past Experiences

These models in general and the legislative changes required for their implementation in particular might seem radical to some readers, as they require copyright holders to concede their full control over subsequent uses of their works, and for indirect compensation for such uses. However, the ACS scholars are quick to point out that while their initiative is indeed innovative, it has deep roots in the existing laws and in business models governing the consumption of content, where similar schemes have already been implemented (with varied levels of success) for quite some time. To make this point, they refer to three instances within the realm of the content industry: compulsory licensing schemes, private copying levies and performing rights organizations.

1. Compulsory Licenses

Compulsory licenses have often been set in response to technological changes that made the arms-length negotiations for the use of copyrighted materials costly, impossible, or unwanted for various reasons.[123] Examples go as far back to the early twentieth century and the regulation of piano rollers,[124] with recent examples pertaining to the use of content

120. FISHER, *supra* note 5, at 246-47; Netanel, *supra* note 2, at 37.

121. *See* discussion *infra* Part III.C.

122. Fisher and Netanel disagree regarding this issue as well, as Netanel calls for repealing the anti-circumvention provisions. FISHER, *supra* note 5, at 248; Netanel, *supra* note 2, at 40.

123. Such as to limit the ability of copyright holders to use their rights to exercise an unfair and anti-competitive advantage. *See generally* Netanel, *supra* note 2, at 31.

124. *See* GOLDSTEIN, *supra* note 6, at 64-66.

over cable and satellite television[125] and most recently, by webcasters.[126] In these instances, copyright holders cannot block the use of their content, yet are compensated by users in accordance to a rate set by a neutral (more or less)[127] entity.

2. Private Copying Levies

In addition to these licensing schemes, in various instances, legislators chose to compensate copyright holders indirectly for private copying, while acknowledging that policing the copyright owners' rights against such actions is close to impossible. These schemes exist on a limited basis in the United States (concerning the regulation of the failed Digital Audiotape Recorder ("DAT")),[128] and on a much broader scale in other legal regimes (such as Canada, Germany, and France).[129] In most of these instances, rights holders are compensated through a governmental fund that is financed by a levy set on various applications relevant to the use of such private copies.[130]

3. Performing Rights Organizations

Finally, in several instances, the copyright holders themselves opt for a business model in which the rights to use their content are not negotiated at arms length with the end user. Instead, these users negotiate with intermediaries or collectives, whom at a later time compensate the right holders from the fees they collect.[131] These intermediaries, such as ASCAP and other performing rights organizations have been put in place voluntarily by the rights holders to collect compensation for public performance rights in an attempt to mitigate transaction costs. The intermediaries stand in for the copyright holders, and directly interact with places of business, such as bars, music halls, and barber shops that pay them a set fee for a "blanket license" for the right to publicly perform. This solution is by far preferable to requiring these businesses to locate

125. FISHER, *supra* note 5, at 41-42.

126. FISHER, *supra* note 5, at 103-05.

127. *Id.* (explaining how the webcasting licensing scheme led to uncompetitive results).

128. *See* Audio Home Recording Act, 17 U.S.C.A. §§ 1001-1010 (West 2006); FISHER, *supra* note 5, at 84-87.

129. For Canada, see Jeremy F. deBeer, *The Role of Levies in Canada's Digital Music Marketplace*, 4 CANADIAN JOURNAL OF LAW AND TECHNOLOGY, 153, 153 (2005). For Germany and France, see P. BRENT HUGENHOLTZ, LUCIE GUIBAULT & SJOERD VAN GEFFEN, THE FUTURE OF LEVIES IN A DIGITAL ENVIRONMENT 24-25 (2003), *available at* http://www.ivir.nl/publications/other/DRM&levies-report.pdf.

130. Be they blank disks, or even computers in Germany's case. *See* HUGENHOLTZ, GUIBAULT & VAN GEFFEN, *supra* note 129, at 25-26.

131. Michael A. Einhorn, *Intellectual Property and Antitrust: Music Performing Rights in Broadcasting*, 24 COLUM. J.L. & ARTS 349, 350 (2001).

the specific rights holders and obtain consent for the use of their content.[132]

However, it should be noted that although these examples demonstrate similar existing dynamics that were successfully implemented, the project outlined by the ACS scholars is far more ambitious. With ACS the market scope and forms of usage are much broader and might deter copyright holders from accepting these proposed models, even though they resemble those agreed upon in the past. The ACS scholars are well aware of such possible hesitation on behalf of the content providers, and offer to ease the way into the full mandatory ACS model by first adopting a voluntary model for content sharing.[133] This model includes the same components mentioned above, but instead of setting a mandatory levy, relies upon voluntary contributions by users interested in enjoying and using the repertoire the model provides. This latter model, which Fisher refers to as "the coop," has already been set in place in some countries,[134] and resembles an interesting business model addressed above.[135] However, as I explain above, voluntary "coops" face several shortcomings and I therefore choose not to further address this option, and focus the analysis on the mandatory ACS model.

C. Presumed Effects and Model Outcomes

At the end of the day, the ACS scholars envision a model that will achieve several important objectives, which would justify the vast framework and radical regulatory changes the model requires. They argue that benefits from implementing this model span across users, content creators and society in general, as well as weaken the hold of today's large media conglomerates which exercise extensive power in today's market setting. *Users and consumers* will enjoy access to a vast library of content at a very limited marginal cost, and would not be subject to manipulative pricing schemes (or price discrimination, as addressed above).

In addition, consumers would be free to make use of this content to express their thoughts in an extremely effective and creative manner.[136] *Content creators* and artists, according to the ACS scholars, would benefit from the shift to the ACS model as well. They would greatly benefit from the availability of vast amounts of content for them to "glum on" to and make use of, thus leading to better and richer outcomes. In addition,

---

132. For the history of such organizations, see GOLDSTEIN, *supra* note 6, at 68-75.

133. *See* FISHER, *supra* note 5, at 252; *see also* Daniel J. Gervais, *The Price of Social Norms: Towards a Liability Regime for File-Sharing*, 12 J. INTELL. PROP. L. 39, 71-72 (2004). Gervais believes that such a voluntary scheme could suffice as a long term solution. *Id.*

134. *See* FISHER, *supra* note 5, at 258 (regarding Brazil).

135. *See* discussion *supra* Part III.

136. FISHER, *supra* note 5, at 238.

the scholars point out that these models allow creators to distribute their works throughout the market (by using the net) without relying upon today's intermediaries. These intermediaries, such as the large record or motion picture companies would therefore be unable to leverage their market position to extract high rents and draconian contractual terms from starting artists. With time, the ACS scholars predict these intermediaries will even give way to other firms that will assist users in choosing content. Thus, this model will weaken the hold of these few powerful entities over the forms of content the public consumes,[137] leading to another important outcome from the models' implementation.

Finally, in terms of *society* as a whole, the ACS scholars mention several overwhelming benefits stemming from the adoption of this scheme.[138] First, implementing the model will end the current shameful state of affairs according to which a large segment of the population (and an even larger segment of our youth) are deemed "copyright infringers," "pirates," and even "criminals." Since all online content sharing will be deemed legal, this serious social concern would evaporate almost immediately. In addition, the system allows for limiting many undue "transaction costs" that result from today's legal setting. For instance, the ACS allows for reducing legal costs which arise from the need to resolve complicated doctrinal questions concerning copyright protection in the online setting. In addition, it allows for reducing costs related to the enforcement of copyright online.[139]

Can the ACS model indeed meet these objectives? Could it be successfully implemented as described? I now move to my critical examination to find out.

IV. TAKING ACS SERIOUSLY: EXAMINING AND CRITIQUING

The ACS scholars go to great lengths to assure that the models they construct should not be deemed a mere intellectual exercise, but a feasible solution with fair chances of actual implementation and success. Indeed, the implementation of these models would lead to many beneficial outcomes, as drawn out above, and their structure is based on a deep understanding of the legal and business background of today's content markets. In view of these elements as well as the breadth of the analysis and the stature of the scholars involved, I see importance in addressing these models. In doing so, I choose to examine their "nuts and bolts"

137. *Id.* at 238.
138. *Id.* at 243.
139. *Id.* at 243-44.

specifically, leaving others[140] to address and critique the underlying doctrinal and economic assumptions on which the model is premised.[141]

An overall critique of the ACS models is an extensive task. Due to the breadth of the proposals, such a critique calls not only for an in-depth analysis of the law and business of copyright, but that of property law, regulatory law, and taxation law and policy to mention a few. Therefore, within the confines of this article, my contribution is quite specific—I closely examine two specific, yet important, issues and questions arising from the ACS scholarship:

(1) Could a system constructed in accordance to the blueprint provided by the ACS scholars, fairly (accurately, as we will soon see, is too much to ask for) measure the uses of content?[142] If so, what would the implementation of measures assuring such fairness entail? By framing the questions narrowly, I set aside (for now) the difficult questions pertaining to the way in which such funds should be raised, the extent of the overall level of compensation to be divided among the right holders, and the way such sums should be updated along the way. Instead, I address the components of ACS charged with measuring consumption and examine their problematic aspects. I then offer ways in which these problems might be resolved and draw out open questions for future inquiry and technological development.

(2) What would the long term effects of implementing this scheme be? The scholars promoting this model present high hopes that it would benefit artists and creators, weaken the dominant position and standing of today's content intermediaries, and enrich society in several ways. At this point, I assume the model will be implemented as described and thereafter examine whether the high hopes and extensive objectives of the ACS scholars would be met, while pointing out where my projected outcome parts from the ACS scholar's rosy predictions. In some instances throughout the analysis I suggest changes to the model and offer external mechanisms to meet the important objectives mentioned, while focusing on modifications to the models' content distribution mechanisms.

A. Measuring Fairly (?): Internal Challenges

To provide copyright holders with fair compensation, the model must present an extensive and accurate mapping as to how content is

---

140. *See supra* note 20 and accompanying text.

141. *But see* Litman, *supra* note 15, at 31 (arguing that setting the details could come later). I disagree, and believe that the many policy issues at hand must be concluded at this early stage in view of the various choices which must be made (many of which have serious policy ramifications).

142. It should be noted that Fisher addresses other solutions for accounting for the users' preferences, such as allowing users to vote for works rather than receive compensation. *See* FISHER, *supra* note 5, at 230, 233. However, this system creates several key difficulties and therefore Fisher rightfully objects to its implementation. *Id.* at 232-33.

used and consumed in the online realm. The model's ability to fairly and accurately measure usage rates of various works is crucial to its overall success, for two obvious reasons (1) without such fairness and accuracy, copyright holders would strongly oppose the model's implementation, and the entire scheme would lose its legitimacy in the eyes of the public; and (2) systematic biases within the measuring process will affect the forms of content society as a whole would generate. In a market operating according to the ACS model, the measurement of content usage is the primary way for consumers to signal their content or discontent with an artist or specific work. Such signaling must be correct to assure the progression and evolution of content markets, as by interpreting and reacting to these signals artists "learn" what form of content is desired by the public, and change their production process accordingly.

The measurement task at hand is colossal. When taking into account the number of Internet users in the United States alone (roughly 100 million) and the number of different "pieces" of content every user will "consume" a day (this of course greatly varies among users, but I believe and average of three would be a modest estimate), the number of factors that must be accounted for in every fiscal year might exceed 100 billion. Thus, when carrying through the measuring task, the administrative agency must overcome both internal challenges (that entail dealing with an extensive dataset and collecting the information in an effective and precise, yet non-intrusive manner) and external challenges (from those who have an interest in intentionally tampering with the data and tilting it in their favor) as well. I address these challenges in turn, while examining what steps must be taken to meet them. These steps, however, create severe side effects, in terms of the system's openness and privacy—which I address below.

1. Internal Challenges & Sampling

The ACS scholars were well aware of the internal challenges, and offer several suggestions. Generally, they suggest that to meet the "counting" objective, the model must introduce a sampling system, which will include several elements.[143] It must include a piece of software to be installed on the end users' machines, that would count their content uses and "report" to a central registry the total amount of uses of different forms of content (the "Counting Software"). In addition, there

143. *Id.* at 225-29; Netanel, *supra* note 2, at 53-54. Here, Netanel also mentions existing technologies which engage in similar sampling tasks. Netanel, *supra* note 2, at 54. A firm that is currently engaged in measuring of content usage through peer-to-peer networks is BigChampagne. *See* BigChamapgne Online Media Measurement, The Data, http://www.bigchampagne.com/thedata.html (last visited Oct. 20, 2006) (describing the way the firm gathers information). However, these practices have met some criticism regarding the accuracy of their results. I also doubt whether BigChampagne could provide sufficient information regarding consumption patterns of content of limited distribution. *See* Jeff Howe, *Big Champagne is Watching You*, WIRED, Oct. 2003, *available at* http://www.wired.com/wired/archive/11.10/fileshare.html.

must be a central system that would sum up, on an annual basis, all the uses streaming in from the individual users (the "Central System"). The Central System will have additional tasks in assuring external fairness, as I mention below. Finally, to efficiently deal with the enormous amounts of data these tasks entail, the ACS scholars suggest that the systems randomly sample a large number of users at any given time, and only include them in the overall database. In other words, while the Counting Software will be working at all times on the users' machines, the Central System will randomly select a specific set of users at set intervals (for instance, every month), and only account for the information streaming in from these specific users at that time. In that way, the model would be able to effectively overcome the massive amounts of data the "collection" and "distribution" stages entail. When addressing sampling, the ACS scholars point out that similar practices have been applied for many decades to establish the rating of the programming on various broadcast stations for the benefit of advertisers (ratings carried out by Nielsen and Arbitron for the television and radio markets, respectively).[144] However, the scholars conceded that the task at hand differs from those mentioned above (regarding radio and TV) as the sample size must be substantially larger than those used in the broadcast context. Yet they do not offer concrete examples as to the sample's size.

2. Sample Size

A closer analysis of the issue of sampling leads to some interesting outcomes. At first, with regard to the actual sample size, I believe that referring to the sampling carried out in the broadcasting context, such as the Nielsen rating model, is a problematic comparison.[145] In the broadcast context, a sample of mere thousands is used to represent the content preferences of many millions. Yet the sample required for the ACS models must be several magnitudes larger. I devote the following paragraphs to the actual size the model must employ, and thereafter move on to examine the implications of using a sample of such magnitude. It is interesting to note that the ACS scholars have neglected to address the actual size of the sample—either nominally, or in terms of the required percentile of the overall sampled population. As I make apparent in my

144. FISHER, *supra* note 5, at 226. For more information as to how these firms engage in sampling see Nielsen Media Research, *Inside TV Ratings: How the Numbers Come to Life*, http://www.nielsenmedia.com/nc/portal/site/Public (follow "Inside TV Ratings" hyperlink) (last visited Oct. 20, 2006); Aribitron, *About Arbitron: What We Do*, http://www.arbitron.com/about/home.htm (last visited Oct. 20, 2006).

145. The Nielsen sampling model has created several controversies of its own regarding its presumed ability to correctly sample preferences in the broadcast context. For instance, it has been argued that the "ratings" are biased against minorities (this was explained by these groups' aversion to fill in the logs they were presented with) and other internal errors in the measuring process. However, many of these problems will not occur within the ACS models in which the sampling is carried out automatically, and users will not always be aware of the specific instance during which they are chosen to be sampled. *See* FISHER, *supra* note 5, at 227-28.

analysis below, setting this parameter is not a technical statistical task which might be left for a later time, but one that required several judgment calls and policy decisions given its potential implications.

Setting the size of the sample involves reaching a compromise between the models' overall efficiency and cost (which are elements advancing the use of a narrow sample), and the fear of unfairness and harm to the motivation of artists whose works are left outside the sample, or who are under-compensated if applied too narrowly (clearly elements advancing a broader sample). Yet striking a balance between these polarizing elements is far from simple. The elements mentioned seem somewhat abstract, while the task calls for identifying concrete parameters for the sampling process. Therefore, to establish a suitable sample size, I move to strike a balance between these elements while taking into account an important and concrete element neglected thus far: the actual level of compensation copyright holders would receive from the administrating agency through the ACS dynamics. It is clearly unrealistic to construct a model that would count every instance of content usage online and provide for full compensation for every such event. Even setting the counting issue aside, it would prove an unbearable administrative burden to send out checks for mere tens or even hundreds of dollars to specific users out of the enormous two billion dollar pot every single year. Yet even to assure that compensation for the sum of $5,000—which is a non-negligible sum for many Americans and especially young artists (and therefore would serve as the baseline for the rest of the analysis)—would not be often neglected, overlooked or under-compensated, the sample must be of considerable size. Using the $5,000 sum as a benchmark leads to an important insight; $5,000 amounts to a mere 0.0025% of the overall yearly fund, yet represents 250,000 separate uses of the specific work every year. Therefore, setting the $5,000 benchmark implies that the sampling process must be sensitive enough so to identify 0.0025% trends within a dataset of 100 billion bits of data pertaining to content preferences.

When taking into account this level of sensitivity, an initial statistical analysis concludes that the size of the sample must be about 0.1% of the overall population (which in this case, as indicated above, would be about 100 million users, and growing, in the United States alone). In other words, this calls for a sample of about 100,000 users! Only with such a sizeable sample, could artists who are entitled to receive annual compensation of about $5,000 be relatively assured there is a reasonable chance[146] that the use of their works would be accounted for and their compensation would not be lost to a statistical error. Any smaller sample, in my opinion, would be unacceptable to these copyright holders, and rightly so. It is therefore apparent that the sampling tasks at hand

146. *See* Statistical Appendix, *infra* note 147.

sharply differ from that of the Nielsen rating system mentioned by the ACS scholars. Yet this is to be expected, as the Nielsen ratings pertain to viewers' choices among merely tens (in the most extreme case that takes into account the various cable channels) of options. In the situation at hand, the model attempts to sample a selection of millions of different forms of content, which display a multitude of consumption patterns.[147]

The assertion that such an extensive sample is required to accommodate artists who are deemed to meet the $5,000 annual threshold can come under several forms of attack. First, on the statistical level, one could argue that even with a much smaller sample, artists whose content is consumed around the $5,000 threshold should not object. Applying a smaller sample would not necessarily mean these artists are to be neglected and left uncompensated. To the contrary, in many cases, the exact opposite would occur; not only would the usage of their content be accounted for, but due to a statistical error acting in their favor, they would receive a payout that is double or triple the size of the sum that would reflect the actual consumption of their respective content. Moreover, in the long run, after several years and samples, the chances for statistical errors of measuring a specific "piece" of content are minimized, and a year of over-compensation would be followed by a year of under-compensation, and vice versa.

In response to this critique, I return to the important objective of achieving fairness in the counting process and reasons for such fairness. I believe that should artists who deserve compensation at the $5,000 level (who stand at the core of those which the model sets to promote and protect)[148] be confronted with the risk of losing substantial compensation in a given year due to a statistical error, they would strongly object to this model, and deem the model unfair, even when facing a similar chance to "double" their income. Furthermore, I believe such an objection would have substantial merit, as it indeed seems unfair that a large portion of the population would not receive their fair share of the overall fund due to an unlucky sample.[149] I also find the "long run" argument stated above unconvincing. Many forms of content have a very limited "life span" on the virtual shelves. This does not result from the lack of space on such shelves, but because of the limited appeal they might have

147. Statistical Appendix (on file with author), *available at* http://law.haifa.ac.il/faculty/lec_papers/zarsky/denver.pdf.

148. As mentioned, the model strives to protect and promote artists with limited market appeal and distribution that are served unfairly by today's market. *See* discussion *supra* Introduction.

149. These assertions might sound merely speculative. They are based on conversations with artists, and the understanding of the alternative options (both those discussed below, and those that provide for a sufficient sample size) that could allow for accurate compensation to a broader array of artists. Clearly, establishing whether these assertions are correct will require surveying public opinion. However, should the ACS model be seriously contemplated by regulators as a viable option, I would assume the court of public opinion will bring the actual opinions and voices of artists on these issues into play.

given new materials that are constantly brought into the market. Therefore there would not be any substantial subsequent sampling "rounds" to potentially offset the unfair errors of one given year, as the public's attention and taste would have already wandered elsewhere. Furthermore, even if a specific copyright holder would receive her fair share several years later, she would still be required to "survive" for several years without receiving sufficient funds—an outcome that might prove unbearable for many starting artists.

The next critiques of my assertion regarding the breadth of the sampling method required (and the goal of protecting the prospective income of the artists with limited market share) return to the main objective of the ACS model—maintaining the status quo. Here, the cautious and careful reader may raise two critiques. First, it could be argued that forms of content with such a meager usage percentile (0.0025%!) which the current analysis addresses are usually disregarded in today's media markets. Therefore these authors need not complain, as the shift to the ACS model does them no harm. For instance, in the broadcast context, there are several examples of programs which attracted a small but devoted audience, yet received a "0.0" rating score according to Nielsen.[150] However, given the fact that advertisers have no interest in shows with extremely low ratings, this sampling error had no real implications. In the content-retail context, works (such as books, DVDs, or music CDs) that are consumed in such a limited number which puts them at risk of being unduly ignored in the overall sampling process are also deemed to be quickly removed from the shelves of the relevant retailers and thus destined for oblivion. Therefore, the results of using a limited sample would, in the worst case, leave these copyright holders at the same point they are today—which as mentioned is the overall objective of the model at this time.

Furthermore, a critical reader may add that when taking a realistic look at today's content business structure and practices, especially with regard to music industry, artists whose works are consumed at such limited scales (such as those mentioned above), rarely receive substantial compensation at all. As Fisher explains in great detail,[151] artists receive mere pennies on every dollar of CD sales revenue. Yet more importantly, starting artists rarely receive any compensation, as the funds they might incur are first applied to cover the advances they received (advances that in many cases were used for promotion purposes).[152] Therefore, when structuring a model to maintain the status quo, the interests of

150. For example, such an instance occurred regarding John McEnroe's short-lived show on CNBC. For more information, see Wikipedia, *Nielsen Ratings*, http://en.wikipedia.org/wiki/Nielsen_ratings (last visited Oct. 20, 2006).

151. FISHER, *supra* note 5, at 55-58.

152. *Id.* at 58.

artists with such limited circulation need not be taken into account, as in today's world their overall situation is grim as it is.

My response to both of these critiques is that they are flawed, as they are premised on a comparison between the legal and business setting we have today, and the one created by the ACS models. However, the ACS scholars' aspiration to achieve a status quo must also take into consideration the outcomes of other, competing solutions to the challenges of the Internet and digital media. In other words, conducting such a comparison must take into account the outcome of the use of DRM systems (that as mentioned above are the leading contenders in today's policy debates) which are backed by appropriate legal and regulatory steps. Such a comparison brings a very different result.

Generally, the implementation of DRM leads to several beneficial outcomes for artists whose works achieve only limited exposure and circulation. At first, DRM systems allow content creators to receive full compensation for all uses of their content, as limited as they may be. These models face no difficulties in capturing all traffic and uses of content (in fact, as mentioned above, that is one of the major critiques of this design) and directly charge per use, regardless of the relative percentage of such use within the overall social consumption pattern. In addition, once the DRM systems are set in place, artists would be able to present their content to a vast crowd with limited expenses associated with the manufacturing and distribution process. Therefore, copyright holders would receive more pennies on the dollar, thus weakening the above mentioned argument that artists with a very limited market share will rarely receive any revenue after the content industry gets their cut. And finally, DRM systems will not be limited by shelf-space, as today's brick-and-mortar stores are, and therefore would allow for the "long tail" phenomenon to take place.[153]

In view of the above, authors of works that are consumed in "small portions" in a DRM architecture over an extended period of time would witness a stable flow of revenue (as they are starting to see today), which they would hardly want to replace with the fluctuating, luck-driven revenue stream the ACS model would provide when using a small sample. In view of all the above, an ACS model using a small sample, which guarantees compensation only for those who produce content with a broad appeal,[154] would seem to be an unwanted option for authors with small- and medium-size audiences, who would probably opt for a DRM-based

153. The newly coined term "the long tail" refers to the fact that thanks to the endless shelf space the Internet e-commerce websites provide, we are witnessing a new and interesting phenomenon – a much greater variety of works are being consumed at non-trivial levels. For more on this issue, see generally, CHRIS ANDERSON, THE LONG TAIL 16 (2006).

154. For an explanation as to why works with a broader appeal face a lesser risk for a statistical error, see Statistical Appendix, *supra* note 147.

solution given its beneficial traits mentioned above. To avoid this result and reach an outcome that would be attractive and fair to this important constituency (which might draw sympathy of large segments of the public), a broad sample must be applied to the ACS collection practices.

In addition to arguments for the use of a broad sample premised on achieving fairness for artists, other arguments for such broad samples could be premised on a different interest—the fact that the use of a limited sample would generate an unwanted incentive structure for content creation. As mentioned, with small samples, the risk of error substantially differs between popular and not-so-popular works. Therefore, with a limited sample, the model might generate strong incentives for authors to develop "instant hits," which lead to great exposure over a short period of time during which such hits are "consumed" time and time again, and quickly thereafter disappear.[155] These hits, of course, will generate a sure revenue stream, as the chances they would be missed by the sampling process are very low. As some scholars argue in other contexts,[156] a hit-driven content market, in which artists strive to deliver "hits" rather than works that might be cherished by a limited audience but have no broad and instant appeal, leads to low quality content—hardly an outcome we would strive for the ACS models.[157]

In summation, in this section I am not arguing that the ACS models are inherently flawed, but that they call for the use of an extensive sample base. However, the relevance of this discussion does not end here. Recognizing that the model calls for extensive sampling has several important implications. First, implementing ACS will call for the construction of a new and unprecedented sampling model that is very different from the ones we have today. Therefore, policymakers must establish whether the task of dealing with such an extensive data set is a feasible

155. Concerning this final argument, it could be stated that these results do not create unwanted changes in the incentive structure, but merely point authors in the direction they were heading in the first place—creating music that would generate the greatest possible revenue! This however, is not always true: First, the statistical analysis I conducted shows that there is a much greater chance for the model to account for the use of a "work" that is consumed *many* times by *few* users than for works that are consumed *few* times by *many* users. *See* Statistical Appendix, *supra* note 147. Therefore, the sampling structure creates incentives for content that is used many times by the same users—a pattern of behavior which resembles those of today's "instant hits." Second, some artists might have a preference in producing several works every year, with every work aiming at a specific crowd, setting or state of mind. This pattern of creation, which might lead to works of high quality, may become unpopular in view of the risks of not being included in the sampling model.

156. FISHER, *supra* note 5, at 79-81.

157. In addition to the reasons stated in the text, it could be assumed that in some cultures, not receiving exact compensation for the use the authors' works, would cause aggravation of all artists. In Israel, for example, AKUM the local equivalent of ASCAP goes to many lengths to provide for a full account of public performances of works (and have even implemented a costly and sophisticated system that aims to account for the use of all works broadcasted on various stations rather than make use of sampling). Interview with Ramat Gan, CEO & COO, AKUM, in Isr. (2005). While the wisdom of such policy (as well as whether it services the interests of its members) can be debated, it still indicates the motivations and state of mind of artists to have a full picture of consumption patters, even at a very high cost.

one. Second (and assuming that applying this sample is indeed feasible, which is not far fetched given the extensive datasets today's corporations manage),[158] applying such a vast sample will complicate the process and create additional costs which must be added to the overall cost-benefit analysis carried out before applying the model. Third, and most importantly, the use of such an extensive sample affects the way in which the models' architects can deal with other challenges, as I will show below. I now move on to point out other problems arising from the implementation of this model that in part result from the necessity of a large sample base. I start with external challenges and the fear of gaming.

B. Measuring Fairly (?): External Challenges and Gaming

1. Introduction to Gaming

Beyond internal challenges to the ACS architects' efforts to provide an accurate picture of the patterns of content consumption, we must now confront challenges arising from attempts of various interested parties to taint the results of the sampling process. In other words, even after resolving the problems of fairly assessing online consumption patterns, the ACS model faces serious external challenge that might undermine its sustainability and render it extremely unfair and unwanted—the challenge of gaming.

By gaming, I (and the ACS literature in general) refer to actions of online users who strive to artificially inflate the number of registered uses of content in the administrative agency's final annual report. The overall reason for engaging in such conduct is clear—to increase the payout to the individual to whom the relevant work is registered. The identity of the gamer and his or her specific motivation and sophistication might vary; gaming might result from actions of professional criminals trying to manipulate and abuse the ACS; newly-founded business ventures that will specialize in "promoting" artists and their works within the ACS collection process; and even the actions of devoted fans that strive to promote their beloved artist and in that way prove their loyalty and affection (possibly after being encouraged to do so by the artist himself).[159] As I will explain and illustrate below, such gaming might be carried out by use of various means, but generally would constitute an attempt to simulate the "use" of a specific form of content, a great number of times. This could be done manually, or through the use of auto-

158. *See, e.g.*, Charles Babcock, *Data, Data, Everywhere*, INFORMATION WEEK, Jan. 9, 2006, *available at* http://www.informationweek.com/news/showArticle.jhtml?articleID=175801775.

159. Artists commonly ask of their fans to "check out their website." Therefore, it is easy to imagine artists encouraging in various ways their fans to access their songs multiple times (it is also easy to imagine that fans of certain forms of music might be more willing to comply—yet I leave the discussion of these different trends of fan behavior for future research).

mated applications, such as "bots," that would constantly "use" the specific form of content.[160]

It is fair to assume that given these various reasons and incentives to game, gaming would indeed occur in the ACS model. The temptation to game would be too great to resist, even though some of these actions are fraudulent and illegal according to today's law, and should surely be rendered illegal by specific rules as part of implementing the model. To prove this point, it should be noted that various gaming practices are creating an overall problem in the Internet media market. In several contexts, commercial entities have an interest to artificially inflate the popularity of certain online products (especially content), as this would lead to a lucrative payout to an interested party. Several examples concerning Google come to mind. For instance, website owners try to game Google's PageRank algorithm and system to assure their website would be prominently displayed as a response to various keywords submitted by searchers. They do so by (among others) artificially linking to and from the relevant site. These efforts have created an entire industry (Search Engine Optimizing—or SEOs) and generate an ongoing cat-and-mouse game between Google and those attempting to "game" its rankings.[161] In addition, website owners try at times to game Google's AdSense system, which posts advertisements on websites and compensates the website owner per clicks on these ads. Here, these website owners attempt to increase their payout by inflating the number of ad-clicks on their webpage, thus threatening the creditability of Google's business plan.[162] These examples show, that when business models compensate (directly or indirectly) for mere attention, and such attention could be artificially simulated online via technological means, then gaming practices would surely be quick to follow.

The existence of successful gaming opportunities and initiatives in a content market operating in accordance to the ACS model would be extremely problematic. It would threaten the stability of the model, and lead to discontent and frustration with its overall structure. Not only will gaming lead to compensation of the undeserving, it will adversely affect other artists that refrain from these practices. Since the sum to be divided among the right holders is set and limited for every given year, the distribution of funds amounts to a "zero-sum-game" in which any additional compensation to one claimant directly diminishes or even elimi-

160. Some ACS Scholars refer to these practices as "ballot stuffing." *See* FISHER, *supra* note 5, at 226; Netanel, *supra* note 2, at 55.

161. For more on this dynamic, see JOHN BATTELLE, THE SEARCH 161 (2005). For an explanation as to how the SEO firms work, see Search Engine Optimization: Information from Answers.com, http://www.answers.com/topic/search-engine-optimization (last visited Oct. 20, 2006). For an explanation of one dynamic, "Spamdexing," see Spamdexing: Information from Answers.com, http://www.answers.com/topic/spamdexing (last visited Oct. 20, 2006).

162. *See* BATTELLE, *supra* note 161, at 187 (regarding "click fraud" of the AdSense system).

nates the payout to the other. For these reasons, minimizing gaming should be seen as an important objective both in planning the model's structure, and throughout its use.

Yet before going further, I offer the following concrete example, which involves two fictitious individuals, Angela and Bruce, who interact in a content market governed by the ACS model. I believe it might somewhat illuminate the abstract gaming concerns mentioned above:

Angela is a gifted violist making her first independent steps in the music business. She has made several tapings of pieces she composed and preformed, and registered them online. Thereafter, she sets up a personal website, where she makes her works available for streaming and downloading. By tracking the usage rates of her website and information as to the trends of the popular file-sharing networks, she learns that there is an interest in her work by a growing number of avid fans.

Bruce, on the other hand, is a terrible musician yet a shrewd businessman. When the ACS model is implemented, he registers several works in his name, all of himself banging away on the drums with no sense of rhythm. He too sets up a website at which his works can be streamed and downloaded. Immediately thereafter, his works are downloaded and used an extensive number of times.[163] *This results from the fact that Bruce, who moonlights as an IT expert at several schools and businesses, has "planted" a small and undetectable piece of software on all of these computers' mainframes. This program causes all the computers within these networks do download and endlessly "play" the pieces registered in Bruce's name.*

At the end of the year, the administrating agency divided the annual fund and sent out "royalties" to the relevant registered right holders. Bruce (and other entrepreneurs like himself) received a hefty sum, which reflects constant usage and a great amount of interest in his "works." Angela received nothing (or close to nothing), as the threshold for receiving funding through the ACS model has been heightened by the actions of Bruce and others, leaving those with works that led to limited, yet genuine interest, with no compensation whatsoever.

The ACS scholars have been quick to identify the risks of gaming and have addressed several strategies in which this concern could be confronted and mitigated.[164] In the following paragraphs, I offer a taxonomy for examining the gaming risks and proper responses, while critically examining the response strategies offered thus far, and suggesting additional insights as to how to approach the concerns of gaming. I also

---

163. This notion of gaming through the use of "bogus content" is not discussed by the ACS scholars, who focus on the promotion of existing works. I, however, believe that this form of gaming would be a major threat and concern.

164. Netanel, *supra* note 2, at 55-57.

examine what the effects of such anti-gaming measures would be on important objectives the ACS scholars point out elsewhere, such as the E2E principle and maintaining information privacy.

2. Confronting Gaming by Sampling

As gaming might turn out to be a serious and strategic threat to the existence of the ACS model, it is wise to examine at this early stage what steps might be taken to mitigate these practices. These steps would be part of the overall "distribution" stage in which the administrating agency determines trends of consumption and allocates funds to the deserving right holders. As mentioned, this stage includes: (1) the task of assessing content usage while relying upon sampling, using (2) Counting Software installed on every system individuals use for content consumption online, and (3) a Central System run by the administrating agency for summing up all streaming results.[165] I will examine how every one of these elements might contribute to mitigating gaming concerns, and what the implications of changing these elements to confront this challenge might be.

The first element which the ACS scholars argue would mitigate "gaming" is the sampling process that is used to create the database, according to which the funds are later distributed to the relevant authors.[166] Since only a portion of the overall population at any given time is sampled and only the information collected in the sample would impact the distribution of funds, it would be extremely difficult for a "gamer" to affect the overall outcome of the fund distribution. This is because there is only a remote chance that his or her gaming attempts would be accounted for—a fact that would discourage potential gamers from engaging in these practices. Gaming practices of course carry some costs (of hardware, software, computer power and time) and risks (of getting caught), and after carrying out a cost/benefit analysis, potential gamers would choose to focus their time and attention elsewhere. Thus, "sampling" not only provides for a more efficient process, but a safeguard against external threats to the accuracy and fairness of the distribution process.

Yet in my opinion and in view of the analysis presented above, relying on sampling alone to battle the threat of gaming is insufficient. This is because of the previous conclusions reached concerning the size of the sample the ACS model would require (in order to adhere to "internal fairness"). As explained, the sample must amount to around 0.1% of the overall population.[167] A sample of such magnitude, would not deter prospective gamers from engaging in gaming practices. When conducting

165. *See infra* Part IV.A.1.

166. Netanel, *supra* note 2, at 56.

167. Statistical Appendix, *supra* note 147.

their cost/benefit analysis as to whether to engage in such gaming, they might presume that given this sample size, it would be feasible to penetrate the sample on a regular basis and assure an increased payout. This presumption would probably be true.

Clearly, unsophisticated gamers, who will try to achieve this by endlessly "playing" their works over their own computer, would surely be disappointed and unsuccessful. However, sophisticated users would surely apply other means to increase the chances of inclusion in the sample. They would simultaneously use several "identities" from the same computer; abuse their access to a network of computers, and even distribute computer viruses and "Trojan horse" programs which will cause other computers to "play" the content of their choice without the actual knowledge of these computers' owners (as Bruce did in the example above).[168]

A cautious critic at this juncture, might question the logic of this last argument, as follows[169]: the sample size should not have an impact on the gamers' decision whether to engage in gaming or not. What the gamers would be looking at is the expected return on their business venture (which is their engagement in gaming practices with their related expenses and risks). This expected return is calculated by multiplying the probability of their success to place "their" content within the sample, by the payoff in the case of such success. It is true that the smaller the sample, the smaller the probability of placing within the sample. However, there is a flip side here as well: the smaller the sample, the larger the payoff in case of inclusion within the sample.[170] Therefore, whatever the sample size, the expected value remains the same, and the motivation of a potential gamer to engage in such gaming should not change. Thus, the argument goes, sample size is an irrelevant factor.

My response to this argument is, that merely examining the expected return in both instances (the one in which the sample was extremely small and somewhat larger) is insufficient. Even though the expected return is equal, the risk involved in both investments is very different: the smaller the sample, the greater the risk that the venture would fail at every given attempt to game. It is true that the expected value is the same, and given an unlimited opportunity to engage in gaming at zero marginal cost, attempts to game a small or large sample should lead to the same economic results. However, marginal costs will

168. It is fair to assume that in a world operating in accordance with the ACS models that it would be relatively easy to distribute these forms of viruses and it would happen frequently.

169. I thank Neil Netanel for engaging me in a discussion regarding this point.

170. This is because the smaller the sample, the more every sampled piece of content would be valued in the final process in which the funds are distributed to the right holders. In other words, the smaller the sample, the more dollars every right holder would receive for every single instance of usage.

never be zero and the gaming practices involve risks. In addition, as any other business venture, this one as well requires some assurance that there is a relatively high chance that the investors would be able to reap the fruits of their labor after a reasonable period of time. An ACS model using small samples will not allow gamers to have such assurance (which is why small samples would in fact be an affective tool to battle gaming). Yet as explained above, the sample would by no means be small. Thus, gaming ventures would be economically viable, and additional measures must be taken to block them.

Another critique, coming from a very different direction, would argue that the actions described above seem far fetched or too pessimistic, as there is no real reason to believe that individuals would go to such lengths in an attempt to squeeze extra dollars out of the ACS model, especially if there is over a 99% chance that any specific gaming effort would not be accounted for at all. Therefore, sampling will be a sufficient deterrent against gaming. It is of course extremely difficult to predict future behavior and outcome in the ACS model at this early juncture. However, I believe that given the relative low costs of computer power, and the way the ACS model is structured, gaming would be sure to become a lucrative business to some and a massive headache to others (namely, the model's administrators, other artists and regulators) even when sampling is applied. To prove this point, I refer to another contemporary online dynamic, which is somewhat similar: *spam*. Here entrepreneurs engage in business practices (that in terms of their legality could be described as varying from grey to the completely illegal) that generate an easy profit by multiplying their voices online at a very low marginal cost.[171] Spammers, and the firms paying for their services, are not deterred by the low rate of success and response these messages have. Because of the extremely low marginal cost of sending multiple messages, merely splinters of one percent in responses to the spam solicitations is sufficient for them to break even.[172]

Continuing this analogy somewhat further, I believe that comparing the potential risks of gaming to the very real problems of spam, teaches an important lesson: mistakes and lack of vision at the early stages of planning systems lead to serious problems at a later stage. At later stages, opportunities to make easy profits through abusing the system are extremely difficult to defeat. Rather, they lead to an extensive "arms race" between those trying to protect the system and those trying to con-

171. Similarities aside, it should be noted that the premise of the spammers business plan is very different. Rather than benefit from a governmental fund, they strive to capitalize on a very limited number of gullible consumers that would purchase the services offered through spam, and in that way render the entire process profitable. *See* Spam: Information from Answers.com, http://www.answers.com/spam (last visited Oct. 20, 2006).

172. For one description of the spamming business model, see Spam: Information from Answers.com, *supra* note 171.

tinue to contaminate it. In the process, they create an overall waste of technological innovation, and additional social costs in terms of burdening courts and other law enforcement agencies. Therefore, if the implementation of an ACS model (on a broad or limited scheme) is to be taken seriously, its planners and administrators must take steps to mitigate the gaming problem in advance. I now address what steps might be taken, while explaining how they could be implemented in the other two elements of "the distribution" stage (that attends to monitoring and assessing content usage)—the "Counting Software" and the "Central System."

3. Fighting Gaming—Beyond Sampling

Clearly, additional measures are needed to deter gamers and mitigate the effects of their actions. I now address several solutions which strive to meet this objective. Here, I refer both to solutions mentioned by the ACS scholars, and my own proposals. The latter are based on proposals mentioned in the context of battling spam.[173] Generally, such strategies will strive to (1) undermine the gamers' business model, and the outcome of their cost/benefit analysis concerning their decision whether to engage in gaming, and (2) Block content usage that is clearly artificial and a result of gaming attempts.[174] The former would mostly be applied through the users' local "Counting Software," while the latter through both the "Counting Software" and the "Central System." The following analysis will address these two components in turn, starting with the "Counting Software."

At this point, one might ask (as I have been asked several times): Is this discussion indeed suited for a legal and policy crowd, as these are not legal nor policy issues but mere technical ones? Shouldn't these questions be left for technologists, system architects, and computer engineers, who are supposed to identify such risks and move to resolve them at the time the system would be implemented? Perhaps. Yet I believe that decisions as to what actions should be taken against gamers are far from merely technical. They involve policy decisions as to the way the ACS model would be structured, which, in turn, have important implications regarding several issues policymakers and scholars found important in the past. For these reasons, I not only believe this discussion is timely, but that it must involve and be of interest to policymakers and lawyers as well.

---

173. As mentioned above the problem at hand somewhat resembles that of spam, and therefore the some of the solutions selected are ones that are applied to battling spam as well.

174. Clearly there would be a problem with definitions here—how should we define, for example, the actions of fans mentioned above (replaying the works of their favorite artist). I assume these actions would and should be rendered legitimate.

a. The End Users' End

As mentioned, the gamers' business model could be undermined by forcing gamers to incur costs when attempting to place content within the sample. With such costs, it would not prove worthwhile to engage in these practices so to receive the mere pennies for every time the gaming attempt proves successful. Possible measures for achieving this objective are structuring the "Counting Software" to register a work as "used" (or "consumed") only if viewed or heard in its entirety,[175] or that additional computational processes must be carried out before the work would counted.[176] Of course, for this scheme to work, it must be applied on every machine and application that might be used to "consume" content (whether they are computers, PDAs, or other portable devices such as iPods or MP3 players). Yet I believe that these measures alone are far from sufficient. Gamers would clearly try to defeat them by attempting to penetrate the Counting Software and shut down these anti-gaming measures. Moreover, they will also try to game the Counting Software itself, so that it sends out indications that specific forms of content have been consumed numerous times, even though that was not the case.

Some of these concerns could be dealt with through other measures I will mention shortly. However, to properly block the gamers' efforts, steps must be taken to protect the Counting Software from tampering. It is of course difficult to establish today what steps must be taken, but it is fair to assume that the industry must establish a standard for "safe" counting software, and that regulation must be put in place to assure that all manufacturers comply. Furthermore, to assure that the system would be secured from tampering, the protocols for carrying out these tasks might have to be kept secret.[177]

Walking through the steps required to mitigate gaming by blocking artificial content usage through measures installed on the users' end (i.e. the Counting Software) leads to similar outcomes and conclusions. Here, to block suspicious trends of usage, the ACS planners must establish a limited number of daily (or monthly) legitimate uses of every form of

175. FISHER, *supra* note 5, at 228.

176. This is a solution that has been suggested in the Spam context. *See* Jo Twist, *Microsoft Aims to Make Spammers Pay*, BBC NEWS (Dec. 26, 2003), http://news.bbc.co.uk/2/hi/technology/3324883.stm.
Other solutions have been suggested in the "spam" context might fit as well—such as Microsoft's initiative to charge a miniscule sum for every email used after a very high number. Here the model might choose to charge users an additional sum (in addition to the levy) if they consumer over a specific number of works in a set period. *See* Microsoft, Q&A: Microsoft's Anti-Spam Technology Roadmap (Feb. 24, 2004), http://www.microsoft.com/presspass/features/2004/Feb04/02-24CallerID.mspx.

177. In other words, these applications would be required to use "closed source" code as opposed to "open source" code that provides for many benefits in terms of allowing for other developers and innovators to add on additional and complementary applications and programs to the existing infrastructure.

content from a single user.[178] Every additional form of use to be registered with the Counting Software would not be accounted (out of suspicion it is merely a result of gaming) and will not increase the payment to the relevant right holder. Again, gamers would try to interfere with these measures by overriding this application, or even figure out ways in which every "machine" could unnoticeably run several pieces of Counting Software simultaneously, and in that way defeat this defensive measure.[179] Therefore, again the model's engineers must preempt this threat by standardizing, securing, and even locking the Counting Software application.

b. ACS vs. E2E

This ongoing circle of action and reaction described above leads to an interesting final outcome: At first it illuminates additional required adjustments to the legal system when implementing the ACS—adjustments that would address the standardizing of the Counting Software and possibly render illegal any attempts to tamper with its inner workings as part of an attempt to game. Yet beyond that, it is apparent that for the ACS model to work smoothly (and battle gaming efficiently), several elements featured and heavily criticized in the DRM systems, must be included in this model as well! For instance, standardization of the Counting Software could be used as means to engage in anti-competitive practices. Furthermore, the ACS model will include elements that would interfere with the Internet's E2E principle.

As mentioned,[180] the E2E principle states the importance of allowing any developer to easily add new applications to the network without requesting the consent of others. However, with Counting Software that includes the secured elements mentioned set in place, developers will be limited to complying with the Counting Software's specifications. This might prove a problem. As discussed in the DRM context, these developers might not be able to exercise their full potential to innovate when forced to comply with external constraints. This would diminish the overall innovation that characterized the Internet, and thus lead to an unwanted social outcome.[181]

The ACS scholars do not address the tension between the aim to achieve external fairness in the counting process and maintaining an open network that adheres to the E2E principle (though in several places they discuss the importance of maintaining the latter, as a reason to object to DRM solutions). However, as this analysis indicates, a conflict

178. FISHER, *supra* note 5, at 229.

179. *Id.* at 226 (addressing this threat).

180. *See supra* note 80.

181. I concede to the fact that the harm to the E2E principle will be less severe than that caused by DRM, yet effects this principle nonetheless.

with the E2E principle might be inevitable. This is not to say that the ACS model should be rejected on this basis. I believe that much to the contrary, the other benefits ACS brings into play should justify the limited use of locked and closed components within the Counting Software. The ACS model promotes creativity by opening up many forms of content to the general public for their unrestricted use—and paying a price in terms of somewhat limiting innovation in the development of web applications is acceptable.[182] Clearly, however, others might not share this view, and therefore this matter must be acknowledged, discussed, and resolved (even on a temporary basis) prior to implementing an ACS model.

c. Battling Gaming, the Central System and Privacy

Additional measures that would be surely required to effectively battle gaming must be implemented through the Central System. These measures will again strive to locate and thereafter disregard content usage that is artificial and therefore an attempt to game. To do so, these Central Systems would be structured to limit the number of times a specific work would be counted from a specific destination within a specific timeframe. A "destination" could be defined as a specific IP address, a specific "machine," or a Counting Software application.[183]

Yet this might not be enough. As gamers would apply dynamic IP addresses and shift from one machine to another,[184] the system must have the ability to detect normal trends of content consumption, and disregard action patterns that sharply differ from these trends (that indicate gaming and distortion might be afoot). However, as I will now explain, meeting this task again conflicts with an important principle the ACS scholars strive to adhere to—maintaining the privacy of the content users (as opposed to the DRM systems, which have been criticized for compromising the users' privacy).[185]

At first, a few words about the ACS model and privacy. On its face, the contemplated ACS models create serious risks for privacy harms. The models call for frequent reporting of the content consumed by individuals to a data inventory controlled by the administrating agency that in turn is part of the government.[186] Clearly, the information the ACS model involves is extremely delicate, as it could provide a great deal of insight into the individual's personality and most inner thoughts that are

182. Of course efforts should be made to construct Counting Systems that allow for both the blocking of distortions and the use of open applications, and in that way enjoy the benefits of both worlds. This is a point worth explaining to technologists and policy makers upon constructing the ACS model.

183. *See* FISHER, *supra* note 5, at 228 (alluding to this solution).

184. *See id.*

185. *See* Netanel, *supra* note 2, at 55.

186. *See* discussion *supra* Part III.

reflected in her decisions as to what forms of content to use and consume.[187] In addition, an individual's knowledge that her entire pattern of content consumption is constantly being viewed and stored would, have an adverse effect on the users' online behavior. Individuals will fear that such data could be passed onto other entities within the government, commercial entities, or abused by individuals with access to the database. Not only would such knowledge and fear cause users to feel intimidated and perhaps a loss of autonomy,[188] but it would affect the content users' choices in selecting content to listen to and view online. Users will conduct themselves in a conforming manner; namely, they would refrain from listening to content that could be viewed as outside the mainstream in fear of what others might think.

Clearly these are unwanted results that would lead to the quick failure of the ACS. For this reason, the ACS scholars have specifically addressed this matter,[189] while setting a very high threshold of privacy protection. They state that the model must include rules prohibiting any subsequent use of the data collected, and requirements that such data be immediately purged after being summed up to formulate the overall sum of works consumed at a specific time (a process carried out by the Central System). These rules, which would reflect similar restrictions existing in some media,[190] will assure users that there is no need for concern regarding their privacy, and that such fear need not impact their content preferences and selections.

Although I strongly agree with the ACS scholars' privacy concerns, I believe these rules set a privacy threshold that is far too high. The personal information pertaining to the content preferences of many individuals is indeed sensitive and raises serious privacy concerns. However, this same information would probably prove crucial in attempts to mitigate gaming. To effectively battle gaming, the administrating

187. *See* Stan Karas, *Privacy, Identity, Databases*, 52 AM. U. L. REV. 393, 438-39 (2002) (discussing the privacy concerns arising with regard to the collection of "mere" consumer data). For a glimpse of the ways in which such concerns generated public outrage in a much more limited context, see various stories concerning the collection and use of personal data by TiVo. *See, e.g.*, Jeffrey Zaslow, *If TiVo Thinks You Are Gay, Here's How to Set It Straight*, WALL ST. J., Nov. 26, 2002, at A1.

188. For a discussion of privacy concerns stemming from the fear that one's actions are constantly being viewed, see Tal Z. Zarksy, *Desperately Seeking Solutions: Using Implementation-Based Solutions for the Troubles of Information Privacy in the Age of Data Mining and the Internet Society*, 56 ME. L. REV. 13, 32 (2004). For the view that the monitoring must be limited in order to limit misuse and embarrassment, see Jerry Kang, *Information Privacy in Cyberspace Transactions*, 50 STAN. L. REV. 1193, 1212-17 (1998). For the view that such monitoring might harm autonomy, see Julie Cohen, *Examined Lives: Informational Privacy and the Subject as Object*, 52 STAN. L. REV. 1373, 1425 (2000). For additional philosophical background on the fear and privacy concerns associated with the creation of vast databases that include personal information, see Daniel Solove, *Privacy and Power: Computer Databases and Metaphors for Information Privacy*, 53 STAN. L. REV. 1393 (2001). *See generally* DANIEL SOLOVE, THE DIGITAL PERSON (2004).

189. *See* FISHER, *supra* note 5, at 228; Netanel, *supra* note 2, at 55.

190. Netanel, *supra* note 2, at 55.

agency must constantly track and analyze the databases of online content consumption, in order to establish a baseline of normal and abnormal consumption patterns and in that way identify attempts to distort the model. At this early stage, it is of course difficult to establish what forms of data would be required to quickly and efficiently detect these patterns. However, I will assume that the analysis cannot rely upon the aggregated data of samples taken throughout the year. Rather, I believe that for the first few years such analysis would require information concerning the origins the data, in terms of an IP address, or a specific Counting Software—in other words, information that might compromise information privacy and would not be available if the privacy measures stated above are taken.[191]

An interesting example of another gaming concern, and the way in which it is confronted, illuminates the nexus between battling gaming and the use and analysis of Internet traffic. Here I refer to concerns regarding click fraud and the threat to Google's AdSense model. As mentioned,[192] Google is currently battling attempts to "game" their lucrative business model, according to which website publishers are compensated per click on advertisements set on this page. In a lawsuit about to be settled,[193] it has been argued that these practices cause advertisers massive losses, and therefore these practices might indeed threaten to undermine Google's business model.[194] As a renowned security expert recently noted,[195] these gamers at times use sophisticated strategies while "attacking" from multiple IP addresses and at times using "Trojan horses" that take over the machines of unsuspicious users and apply them towards these causes. The settlement mentioned, and the documents published by experts involved in the case provide us with some insight as to how Google confronts this challenge. Here too, Google is responding in several ways. First, its experts automatically block repeated clicks that are clearly fraudulent. Yet to block more sophisticated gamers, Google employs teams of experts as well as sophisticated algorithms that examine the overall database of clicks, which include a data trail about every click, with information regarding its originating IP address and the

---

191. Here I disagree with Ku who holds that privacy would not be a problem as information beyond the aggregated sums of usage will not be required. This assertion is incorrect given the risks of gaming. Ku, *supra* note 15, at 314-15.

192. *See supra* Part IV.B.

193. Plaintiffs' Second Amended Class Complaint, Lane's Gifts and Collectibles LLC v. Yahoo! Inc., Case No. CV-2005-52-1 (Ark. Cir. Ct. Feb. 17, 2005). On this issue, see Eric Goldman, Technology & Marketing Law Blog: *Lane's Gifts Click Fraud Lawsuit Near Settlement*, http://blog.ericgoldman.org/archives/2006/03/lanes_gifts_cli_1.htm (Mar. 8, 2006, 16:25 EST); Nicole Wong, Official Google Blog: *Update: Lane's Gifts v. Google*, http://googleblog.blogspot.com/2006/03/update-lanes-gifts-v-google.html (Mar. 8, 2006, 13:58 EST).

194. *See* BATTELLE, *supra* note 161, at 186-88.

195. Bruce Schneier, *Wired News:Google's Click-Fraud Crackdown*, WIRED NEWS, July 13, 2006, http://www.wired.com/news/columns/0,71370-0.html.

time it took place. After examining these databases, they attempt to establish what constitutes a normal and abnormal form of ad-related clicking. Thereafter, they move to disregard abnormal clicks, and amend their automatic filters to disregard such clicking patterns from there on.[196]

Returning to the ACS model, I believe the analysis and example above provide us with some important insights as to future planning and the way in which the ACS model must deal with personal information. Clearly, this analysis should not lead to the conclusion that within the ACS model, all privacy protection must be abandoned (an outcome that might undermine the ACS model in its entirety). However, understanding this potential conflict between the need to respond to gaming and privacy concerns requires us to realign the means of privacy protection the model will employ. First, the ACS model must abandon the very high threshold of privacy protection mentioned above that would not allow for the meaningful analysis required for fraud detection. Clearly, data regarding content consumption should not be put to subsequent commercial uses, or passed on to third parties (either commercial or governmental), yet it cannot be purged immediately as well. Rather, the ACS's privacy policy must be structured to allow the administrative agency to probe the dataset of information pertaining to the samples gathered, which also includes data as to the sources of the sample (in terms of IP addresses or even an identification number for every Counting Software).

Allowing such practices to take place will, of course, generate some privacy concerns. Individuals may fear that security will be breached and the data regarding content consumption will leak, or that someone within the administrative agency will misuse the data. They might also fear that the government will subpoena such information if it deems it necessary for an investigation (a realistic option given recent events). Therefore, steps should be taken to preserve privacy, while maintaining the ability to battle gaming. For instance, strict security requirements could be set in place regarding these databases, with harsh punishments for those who breach them.

Should the steps mentioned prove insufficient to confront fears of the government systematically misusing this database, the model's planners might consider solutions recently examined by security agencies in the context of the war on terror. Here, the government is faced with the challenge of examining vast commercial databases (such as credit card

196. For an in-depth discussion as to how these practices take place, see Alexander Tuzhilin, *The Lane's Gifts v. Google Report*, http://googleblog.blogspot.com/pdf/Tuzhilin_Report.pdf (last visited Oct. 20, 2006); Click Quality Team, Google, Inc., *How Fictitious Clicks Occur in Third-Party Click Fraud Audit Reports* (Aug. 8, 2006),http://www.google.com/adwords/ReportonThirdPartyClickFraudAuditing.pdf#search=%22How%20Fictitious%20Clicks%20Occur%20in%20Third-Party%20Click%20Fraud%20Audit%20Reports%2C%20Click%20Quality%20Team%2C%22 (last visited Oct. 20, 2006).

and airline databases) that might hold important clues as to future terrorist attacks. Obviously, allowing the government to access these databases without restrictions, while conducting massive "fishing expeditions" creates serious privacy concerns.[197] To meet this challenge, researchers are trying to formulate ways in which the government could engage in data analysis and data mining, while searching and detecting data patterns of dangerous anomalies, without having the ability to "see" the data itself.[198] Only after such anomalies that indicate the existence of a security risk are detected, are the law enforcement agents permitted to receive information as to the actual data within these datasets (as opposed to overall trends).

Shifting back to the ACS model, applying these new technologies (should they indeed prove workable) would allow for creating a database of information relating to the users' content consumption to be held in confidence by a trusted third party that would not be permitted to make any use whatsoever of the data. Thereafter, the governmental administrative agency would analyze and "mine" this database, without having access to the data itself. Only after establishing the existence of gaming patterns, will the agency move to block similar actions in the future. Clearly this issue requires additional research as to whether it is even feasible to blindly recognize such patterns in an effective manner and these options still need to be discussed and weighed before decided upon. Yet the importance of maintaining the ability to battle gaming must be borne in mind when addressing privacy questions in the ACS context. The abovementioned examples show that based on other contexts, balancing gaming and privacy might be possible, with proper planning and understanding of the interests involved.

I conclude with two final points regarding privacy. First, I note an additional issue that would require future discussion and analysis—the ability of the administrative agency to bring action against "serial gamers" (by way of existing or special laws). To do so, the administrative agency must not only block the gaming practices, but establish their existence, locate the "gamer" and link him or her to the gaming activities—all tasks that require access to personal information. Clearly the extent of the gaming problem will set the tone as to what forms of actions would be taken. Enabling the administrating agency to engage in these actions will create privacy concerns, as well as concerns of selective enforcement, which will somewhat echo concerns voiced today regarding

197. For a recent account of this issue, see JEFFREY ROSEN, THE NAKED CROWD 148-49, 196 (2005).

198. For additional details, see K. A. Taipale, *Data Mining and Domestic Security: Connecting the Dots to Make Sense of Data*, 5 COLUM. SCI. & TECH. L. REV. 2, 74-81 (2003). For a discussion of the legal implications of these tools, see ROSEN, *supra* note 197, at 148, 196.

the ongoing lawsuit against individuals engaged in file swapping. I leave these discussions for a later time.[199]

Second, privacy in terms of the ACS model raises an interesting question regarding the ability of the copyright holder to capitalize on their right without exposing their identity. In today's content markets, and especially in one enabled by DRM, creators could sell their works without the need to establish their identity, and can even rely upon the use of pseudonyms.[200] With ACS, however, the administrative agency must have proper identification and contact information about the copyright holders so as to provide them with proper compensation by the end of the year. This requirement might compromise the ability to capitalize on anonymous or pseudonymous works.[201] These concerns could be substantially mitigated by structuring the ACS model in order to provide for the pseudonymous registration of works. However, allowing such registration would (again) lead to problems concerning the ability to track gamers, who would constantly attempt to register bogus works in their name. I leave the balancing between these objective (anonymous creation and defeating gaming) for a later time.

To conclude our discussion of achieving fairness when measuring usage in the ACS model, I point out that the key to this task is balancing. After correctly establishing the benefits and detriments of every policy choice and steps, regulators must balance privacy, network openness, accuracy in measurement and, vulnerability to gaming. I hope this discussion will assist in this complicated task, by drawing out the elements involved, and the possible tools available for constructing the proper balance.

C. The Outcome of ACS Implementation

At this point of the analysis, I put aside my examination of the "nuts and bolts," and move to examine the impact of implementing the ACS model. I do so while accepting that the implementation of the ACS is politically and legally feasible. As mentioned above, when Fisher sums up his description of the ACS model, he mentions the benefits of such implementation to users, artists, and society on the one hand, and the much welcomed weakening of today's overpowering media conglomer-

---

199. FISHER, *supra* note 5, at 225-26 (addressing this matter briefly).

200. Note, however, that right holders can rarely enforce their rights without revealing their identity.

201. This is of course not to say that the model compromises the important right to speak anonymously. On this issue, see Tal Z. Zarsky, *Thinking Outside the Box: Considering Transparency, Anonymity and Pseudonymity as Overall Solutions to the Problems of Information Privacy in the Internet Society*, 58 U. MIAMI L. REV. 991, 1024 (2004). The question as to whether there is also a right to financially capitalize on content distributed anonymously should be addressed at a later time.

ates on the other hand.[202] In this sub-part, I will confront these predictions while closely examining three issues: (1) Which segments within the content industry might choose to voluntarily exclude themselves from the model, and whether such exclusion will cause a problem to the overall implementation of the ACS model; (2) How the model's implementation would affect the balance of power between the various players in the media market (3) Unplanned and unwanted effects the model's implementation might have on the creation, development and distribution of content in the post-ACS digital environment. Throughout my analysis I examine how these issues compare to the objectives the ACS scholars set out to achieve, and suggest changes that might be required in the model's structure to overcome instances in which these objectives will not be met.

1. Forms of Content—The Limits of the Model (Or: Why Pornography and ACS Don't Mix—and Why We Need Not Worry About That)

As mentioned above, even though the ACS model consists of many mandatory elements (such as a levy on services and applications related to online usage and possibly the use of Counting Software), it still requires voluntary participation of one important group—the copyright holders who must agree to provide their content online through the open ACS model. The other option such content owners might exercise is refraining from providing and distributing their content online altogether, or limiting such distribution to locked DRM systems that would directly control the ways in which users access their content. As the ACS scholars argue correctly, this option will be quite unattractive and artists have strong incentives to participate in ACS. For upcoming artists, the model provides for vast exposure to a worldwide audience. For already renowned artists, the model provides compensation for online uses that are already taking place and will probably continue to take place online in any event (in spite of the industry's attempt to block them) given the fact that DRM technology cannot provide foolproof locks against the leaking of content to illegal sharing networks.[203] In addition, opting for limited distribution through a DRM model should be undesirable,[204] as this re-

202. *See supra* Part III.

203. Artists are usually unable to block the migration of their works online, as users upload versions of the famous works of these artists to the web within the file-swapping networks, while at times "cracking" various locks installed on these forms of content offline (such as the cracking of DVD encryption and uploading full length motion pictures to the file swapping network). *See* Netanel, *supra* note 2, at 9-10.

204. As mentioned, according to Netanel, the adoption of an overall ACS scheme also calls for the repeal of the legal protection amounted to the trusted systems enabling the DRM infrastructure. *See id.* at 40-41. In other words, the implementation of ACS models will call for canceling or limiting DMCA-like provisions that prohibit and criminalize the circumvention of copyright protection mechanisms. Thus, the authors would have an even greater incentive against opting for the DRM option. *See id.* at 59. Note that Fisher objects to this notion—while arguing that individuals should

quires substantial setup expenses, as opposed to the mere registration the ACS model entails.

However, even after taking these benefits and detriments into account, some copyright holders might still refuse to participate in the ACS model. It is, of course, extremely difficult to predict how many and whom will chose to exclude themselves, yet I believe, it is safe to assume that a specific segment of the online content market will refrain from participating in ACS—those creating and producing pornographic materials.[205] These copyright holders will probably continue to make use of trusted systems for distributing their materials, and lead the way in the development of new applications with greater security.

I draw this conclusion while relying upon several arguments. First, I believe these copyright holders would refrain from participating in the ACS in view of the registration process. This process will explicitly link their names with the production of this form of content in a public register, thus making this information available both to the public and to the government.[206] Second, these content providers will fairly assume that individuals interested in "consuming" such content online, would be uncomfortable with reporting their consumption histories through the use of the counting systems mentioned above,[207] and will therefore try to hide any traces of such consumption.[208] This is quite the opposite of many other settings, where users would be happy to indicate their interest in specific forms of content, as it would lead to additional compensation for the artists whose content they now enjoy. The users' reluctance to participate in the counting process would lead to a drop in the compensation such content providers would reap through the dynamics of the ACS model (as opposed to the compensation they might reap through other compensation models).[209] In view of these arguments, I believe the por-

---

be allowed to make use of both models, and that they would no doubt flock to ACS that is preferable by far. *See* FISHER, *supra* note 5, at 108-10.

205. In this segment, I only refer to works whose distribution is permitted according to relevant laws. The online distribution of content deemed illegal requires an extensive analysis that is beyond the reach of this article.

206. The motivations here might be mixed. Some might fear the public eye and social backlash of being associated with this form of content. Others might not want the government to have easy access to such lists. This argument is not without flaws—as the current system requires these content providers to provide information to both government and the public concerning their activities (when suing to enforce copyright, registering websites or even filing tax returns concerning their operations). However, I argue that the ACS model would require a great deal of exposure and an easily accessible central repository.

207. Note that these concerns will be exacerbated given my previous analysis of the required balance of privacy measures and concerns with anti-gaming activities that will inevitably broaden the concerns users will have regarding the governments ability to track, save, and see what content they are consuming online. *See supra* notes 186-91 and accompanying text.

208. *See* FISHER, *supra* note 5, at 227 (describing a similar dynamic regarding the tracking of content usage by Nielsen Media (in the context of television)).

209. Note that I need not argue that the consumption of pornography will decline because of the tracking devices put in place—an argument that is somewhat problematic to prove given the fact

nography content industry will opt for the DRM model, as opposed to the ACS model addressed above.[210]

Yet the reluctance of the creators and distributors of pornographic content to make use of the ACS model need not indicate the model's weakness. Much to the contrary, I believe such reluctance would strengthen the model and assist in its implementation, by somewhat silencing two powerful critiques against its adoption. One such critique would argue that a model that does not charge for marginal uses of content will potentially allow users to access endless amounts of pornographic materials at no additional costs. This, in turn, may lead to an array of problems, such as unhealthy addictions of online users to such materials. Another critique (with somewhat of a populist flavor) will argue that the ACS scheme leads to the cross-subsidizing of the "consumption" of pornographic content online. In plain terms, it argues that individuals who barely use the Internet in general and file sharing applications in particular, would be indirectly funding (through their contribution to the governmental fund, via the levy) the production of pornography and enabling excessive use of such content.[211] Clearly, this argument could be made with regard to many other elements and forms of content.[212] However, placing this critique in the "pornography" context is sure to generate additional support and may threaten the implementation of the model, as it will find its way to the hearts of many citizens.

In summation, pornography has been a driving force in the development of online technologies, and generates a vast amount of online traffic. According to this analysis, this industry will remain outside the model. However, in view of the benefits stemming from this outcome, I believe this should not be a reason to reconsider the way the model is constructed.

2. The Role and Power of Intermediaries in the ACS Model

a. General

An important objective the ACS scholars aim to achieve in the shift to the ACS model concerns the realignment of power in the media con-

---

that the consumption of free pornography is extremely popular online. I however argue that users will take actions to avoid being tracked by the various means mentioned above when consuming pornography—which would lead to substantial losses to these copyright holders.

210. Note that DRM creates privacy concerns of its own. *See supra* notes 82-83 and accompanying text. However, these concerns can be mitigated by sophisticated consumers making use of e-cash and similar measures that will not allow for tying their payment method to their real-world identity. Such measures will not be helpful in the ACS world, which must track (for reasons mentioned above) the actual IP address the consumer is using.

211. FISHER, *supra* note 5, at 217.

212. For instance, conservatives, who use their computers for word processing and email only, will argue that they are cross-subsidizing the consumption of music and video content which advocates ideas they strongly disagree with (and vice versa). *Id.*

tent markets. These markets feature large media conglomerates, which serve as intermediaries and deliver the works of artists to their prospective consumers. In today's media markets, such intermediaries are vested with a great deal of power, which according to several scholars leads to unwanted results to artists, consumers and society as a whole.[213] The ACS model, so it is argued, empowers both artists and consumers and thus mitigates the unwanted results stemming from the existence of the overpowering media intermediaries in the current market (a result stemming from today's market structure).[214] I hereby examine the assertion that a shift to the ACS model will indeed realign this market balance and cure the many problems this balance (or rather, imbalance) creates. In doing so, I provide both an analytical and comparative analysis that might prove otherwise. Thereafter, I discuss steps that might be taken to allow the ACS model to achieve this objective. I also address points for future research to sharpen the understanding of the role of intermediaries in an ACS content market.

To start out, a few words regarding the role of media conglomerates as intermediaries in today's content markets.[215] First, in terms of their relation with artists, these firms provide them with funding, connections, knowledge and expertise, and in this way promote them from anonymity to stardom. The firms make use of their massive distribution and promotion mechanisms and deliver the relevant forms of content to the actual and virtual doorsteps of the masses.[216] Before launching this process, however, in the music context, most artists assign their rights in the sound recording over to these intermediaries and in return receive mere pennies for every dollar to be made in sales of their works.[217] Artists are forced to do so because they lack any other meaningful option to promote their content and deliver it to interested consumers. Second, the intermediaries provide an important service to the audience (the consumers) as well. They choose specific works from a nearly limitless selection and advise consumers that such content is worthy of their limited attention.

---

213. For one description of market concentration, see ROBERT MCCHESNEY, THE PROBLEM OF THE MEDIA 177-83 (2004). For an opposing view, see BENJAMIN M. COMPAINE & DOUGLAS GOMERY, WHO OWNS THE MEDIA? (2000).

214. FISHER, *supra* note 5, at 242. In all fairness, it should be noted that according to Fisher, the future role of today's powerful intermediaries is unclear—they might be able to capitalize on their expertise and power to remain vital and profitable in the new realm, but might also be outperformed by newer players. Elsewhere, however, when addressing the effects on artists, Fisher mentions that the model will allow them to be less dependant on a few intermediaries. *Id.* at 240. Below I examine this key assertion in depth.

215. This segment of the analysis is structured in terms of the music industry. The arguments could be rephrased to meet the structure of the television and film industry, which are probably far more concentrated.

216. For example, see FISHER, *supra* note 5, at 21-22 for the roles of music intermediaries in today's markets.

217. *Id* at 54-55.

Given the economics of scale and scope content markets involve, these intermediaries have grown in size, and a limited number of them dominate a vast portion of the market.[218] Beyond several advantages such integration provides, this phenomenon leads to problematic outcomes on both sides of the equation (i.e. vis-à-vis artists and consumers). The market power and dominance these firms enjoy allows them to obtain draconian terms when negotiating with artists, thus limiting the artists' actual benefits from the fruits of their talent and labor.[219] With regard to consumers, it is argued that the content selection these concentrated intermediaries provide is dull and mainstream, as well as limited given the almost endless array of options.[220] This results from the intermediaries' policy of maximizing profits, which at times conflicts with other social objectives.[221] In addition to this critique as to the actions of the intermediaries, it has been argued that permitting a limited number of corporate entities to control the content consumption patterns of a vast segment of society is problematic *per se*, as these entities will control what the public knows and therefore how it thinks and acts.[222]

The ACS scholars are well aware of these concerns regarding the powerful position of content intermediaries in the media market, and advocate ACS as a way to mitigate these problems. Indeed, at first glance, the ACS seems to provide a reasonable response to these failings of today's content markets, with regard to the troubles of both artists and consumers. First, in terms of artists, the ACS model creates a media market with an extremely low barrier to entry. Any artist could easily upload her work to the Internet, where it could be accessed and used by a very large audience, and receive indirect compensation for the content's consumption (after going through a quick, cheap, and simple registration process). Therefore, these artists would not be forced to rely upon the assistance of the mentioned intermediaries, while making use of the Internet's infrastructure and features for content distribution and promotion. Also, they need not rely upon the intermediaries for compensation, which they receive directly from the administrative agency.[223] Consumers, too, would benefit from the shift to ACS. They will not be limited to the content the intermediaries choose to promote and distribute, but could access a broad array of content directly online, while interacting directly with the artists themselves.[224]

---

218. *See* MCCHESNEY, *supra* note 213, at 177-83. *But see generally* COMPAINE & GOMERY, *supra* note 213.

219. *See, e.g.*, FISHER, *supra* note 5, at 54-55.

220. *Id.* at 80-81, 238.

221. *Id.*

222. BENKLER, *supra* note 33, at 202.

223. FISHER, *supra* note 5, at 238.

224. *Id.* at 239.

b. The Critical View

Critically reviewing these rosy predictions of the realignment of forces in the ACS content market leads to some skepticism of their accuracy. The shift to the ACS model will indeed lower several barriers to entry for artists to the content market, and will allow them to easily upload their content, as well as distribute it directly to consumers. Therefore, the importance of these aspects of the content intermediary's role will quickly diminish. However, in a market operating in accordance to the ACS, the role of intermediaries, vis-à-vis consumers, will still remain, and gain importance. Here, consumers will reach out to intermediaries for guidance in selecting content which might meet their specific interests, be of the highest of quality and thus worthy of their attention. Indeed, in the ACS model, human attention is a scarce commodity (which eventually leads to the artists' compensation) and one which consumers will try to guard when facing the abundance of content the model provides.

The ACS scholars acknowledge and even welcome the prospect of important intermediaries in the ACS content market. However, they argue that these intermediaries need not be so powerful as to allow them to abuse artists.[225] In addition, they need not be the same conglomerates we have today (although these entities have been known to leverage their market power in one medium towards another). Rather, the ACS model will lead to the appearance of an abundance of experts, media critics, or simply music or movie lovers that will comb through the Internet searching for notable materials, and will list and link to them at their respective home pages.[226] These will be the new intermediaries of the ACS age.

Though this description may seem convincing, I find it too optimistic. It is missing a crucial element I address below—that of the power which would be amounted to intermediaries in the ACS content market. Furthermore, I believe there is a good chance that all the ailments that inflict the general media market will manifest in the ACS content market as well, thus leading to the reappearance of today's concerns of concentration and imbalance. In the next few paragraphs I will explain why.

As mentioned,[227] the ACS content market will feature many intermediaries, which will all offer content "consumers" lists of recommended forms of content. The key question, however, is which intermediaries will the public choose to trust and entrust with their valuable attention span, and what will the trends of "intermediary selection" resem-

225. *Id.* at 238.

226. This notion was mentioned by Volokh with regard to the broader Internet context. *See* Eugene Volokh, *Cheap Speech and What It Will Do*, 104 YALE L.J. 1805, 1815-16 (1995).

227. *See supra* note 226 and accompanying text.

ble. Based on similar instances occurring in other media markets,[228] I assume that most of the public will focus on and flock to a limited number of intermediaries. The public will demonstrate a trend of concentration on a limited number of resources. These trends, in turn, will provide the "popular" intermediaries with a great deal of power. Moreover, I will argue below that there is a good chance that these "popular" intermediaries will be an extension of the same powerful conglomerates that dominate today's content industry. Thus again leading to the reemergence of various concerns.

Let us begin with the future trends of "intermediary selection." Clearly, in the ACS model, almost anyone could become a self appointed intermediary. However, as recent work in the fields of sociology and network theory indicates, human attention tends to be concentrated, and masses tend to focus most of their attention on a very limited number of resources, for a variety of reasons I need not address here.[229] In other words, human attention tends to be distributed according to a "power law," rather than equality among various outputs available. This is best demonstrated by recent studies concerning the Internet. While the Internet allows almost anyone to set up a website that is accessible world wide, various analyses of online content consumption, market structure, and even link structure lead to the somewhat surprising result—the Internet is turning out to be as concentrated as other forms of media (even though the physical barriers to entry are considerably lower). This concentration is expressed in the overwhelming share of a limited number of entities in the revenues the online market generates,[230] the attention users pay to websites,[231] and the number of incoming links other websites post on their pages.[232] These studies show that the Internet content market is demonstrating interesting trends of concentration that lead to high barriers to entry and new hubs of power. They also show that in today's online realm it is quite difficult for an independent website to gain a dominant market position, even though the barriers to entry were assumed to be very low. While we are constantly confronted with anecdotal stories of blogs, video clips and songs that start out in the author's garage and reach a very broad audience, these are still exceptions to the

228. *See infra* notes 229-33 and accompanying text.

229. On these issues, see DUNCAN J. WATTS, SIX DEGREES (2003); ALBERT-LÁSZLÓ BARABÁSI, LINKED (2002).

230. Eli M. Noam, *The Internet: Still Wide Open and Competitive?*, TPRC (2003), http://tprc.org/papers/2003/200/noam_TPRC2003.pdf.

231. *See* BENKLER, *supra* note 33, at 238. For a study proving this assertion in the limited context of blogging, see Shirky: Power Laws, Weblogs, and Inequality, http://www.shirky.com/writings/powerlaw_weblog.html (last visited Oct. 20, 2006).

232. *See* Matthew Hindman, Kostas Tsioutsiouliklis, Judy A. Johnson, *Googlearchy: How a Few Heavily-Linked Sites Dominate Politics on the Web* (2003), http://www.cs.princeton.edu/~kt/mpsa03.pdf. Benkler sums up these empirical studies. *See* BENKLER, *supra* note 33, at 238-40. However, Benkler (in the context of the mass media in general) does not believe that the Internet displays or leads to over-concentration, but is just right. *Id.*

overall trend and pattern of content consumption online.[233] The market is mostly dominated by a selected few (the identities of which we will soon address).

Superimposing these theories and concrete findings on the issues at hand leads to the conclusion that the world of content intermediaries in the ACS content market will probably prove to be no exception. Most of the users (especially those lacking sophistication) will flock in great numbers to a limited set of intermediaries, who will dominate the "attention" market. Once the dominance of these intermediaries would be established, it would be quite difficult to penetrate this closed circle. At this point, these intermediaries will command a great deal of power over artists (and to a lesser degree, consumers).[234]

Next, let us give some thought as to who these intermediaries might be (although the actual identity is of only secondary importance for forwarding this argument). Arguably, these intermediaries could be anyone who wins the public's trust in this ever-changing medium. However, here again a view of the Internet's trends of content consumption proves instructive. Reviewing the lists of this medium's most popular destinations leads to a limited number of websites, which include several websites which were founded early on and therefore enjoyed a "first mover" advantage.[235] However it mostly includes websites affiliated with renowned brands of the offline media world (such as Time Warner, Disney and Microsoft).[236] The success of these websites could be explained by these firms' ability to leverage their success and position in other media markets towards domination in the Internet medium as well. Such leveraging is achieved while making use of their capital, brand, and goodwill as well as their ability to divert the attention of their audiences in other media towards their online presence.[237]

Again, let us return to the ACS model. Here, it is fair to assume, these dominant intermediaries would be able to leverage their dominance in other media as well. However, an important caveat is in order: in the "general" online context, the media conglomerates were able to assure their online dominance by capitalizing on their vast content inventory which they control through intellectual property laws and are already

---

233. C. Edwin Baker, *Media Concentration: Giving Up on Democracy*, 54 FLA. L. REV. 839, 895-97 (2002).

234. For instance, see BATTELLE, *supra* note 161, at 153-59, for examples of the power of Google as an intermediary. Battelle demonstartes that if excluded from Google's results this might cause a devastating outcome for the excluded party.

235. Ebay, Yahoo! and Amazon.com are examples of such websites. *See* BENKLER, *supra* note 33, at 245-46, for conflicting studies regarding the role and dominance of first-mover websites online.

236. *See* MCCHESNEY, *supra* note 213, at 221 (relying upon a study by the *Columbia Journalism Review*).

237. *Id.* at 177-83, 221-27.

known to the public. By presenting such content exclusively on their new websites, they were able to attract Internet traffic and attention. In the ACS model, such leverage would not be possible, as any other intermediary would be permitted to recommend, present, and link to the content of others. Therefore, it remains to be seen whether these media conglomerates could gain dominance in the ACS media market while relying on their offline goodwill and trademarked brand alone.

The notion that the Internet would become a concentrated medium that will allow (and according to the above mentioned studies, indeed allows) many offline media conglomerates to maintain their strong market position is far from novel. Already in 2000, in an insightful article, Professor Netanel pointed out that the Internet medium will not lead to an overall restructuring of the media market power balance, but would facilitate the continued dominance of the major media conglomerates in this new medium, for the reasons mentioned above.[238] I believe this analysis should be applied to the narrower context of the ACS model as well, which will share many of the attributes of the broader online context, and lead to similar forms of concentration.

Finally, we reach the third tier of the critique as to the role of intermediaries in the ACS content market, which stems from the previous two; if merely few intermediaries will command the access to the majority of consumers, then the actual barrier of entry to this new content market will remain extremely high for upcoming artists. These artists can post their materials online, or integrate them into the file-swapping networks. Yet if the majority of public attention is focused on the content specific intermediaries recommend (which might be early movers, or the "good-old" media conglomerates), artists will only reach true fame and compensation if selected and endorsed by these powerful intermediaries. Therefore, the actual change in the balance between the media conglomerates and artists might not occur. It is quite possible that upcoming artists would still be forced to sign one-sided agreements with the dominate ACS intermediaries in order to gain name recognition to the extent that would lead to substantial compensation (in what would resemble the infamous payola scheme which often exists between artists and radio stations).[239]

c. Possible Solutions

For ACS to indeed weaken the position of dominate intermediaries (which, as I explained, might be the same ones we have today), the

238. Neil Weinstock Netanel, *Cyberspace Self-Governance: A Skeptical View from Liberal Democratic Theory*, 88 CAL. L. REV. 395, 440-41, 463-65 (2000).

239. As explained above, a substantial level of usage is required to reach high levels of compensation. *See supra* Part IV.A. Therefore, given the immense competition these markets will demonstrate, the role of these intermediaries will be as important as ever.

model must include additional elements. One somewhat aggressive solution might call for regulatory intervention requiring dominant content intermediaries to carry all forms of content equally and without discrimination, as well as regulatory steps to assure that dominant firms in other media would not move to take over the online distribution market.[240] However, the chances such solutions would be accepted are slim, as they will meet fierce objections. They would especially meet the objection of the media conglomerates' representatives (that have demonstrated their ability to influence legislators and assure their interests remain secure) while arguing that such regulation impedes upon the rights of these content firms to engage in free and unregulated speech.

Beyond the regulatory solution, I suggest that the strengthening of the artists' position in the ACS realm is achievable by researching, developing, funding, promoting, and maintaining alternative means to distribute content when shifting to the ACS model (in addition to the measures put in place to achieve proper compensation). One possible option mentioned calls for reliance upon various sites that provide ranking and sorting that are formulated in a "bottom-up" process; in other words, users (working in collaboration) would both classify the many forms of content available online, and rank them according to their subjective liking.[241] The advantage of these forms of recommendations mechanisms, also currently referred to as "folksonomies,"[242] is that they do not reflect the preferences of one central intermediary (that might have specific interests), but thousands of individuals.[243] In other words, this is a "many-to-many" process.

Folksonomies are coming into existence through several websites that allow users to sort and rank various forms of information, including content. This new concept is currently being closely examined by academics and businesspeople. At this time, however, I am somewhat skep-

---

240. Fisher addresses a similar option in drawing out a possible broad alternative solution to the challenges of digital copyright—which includes an extensive regulatory framework to promote content creation. *See* FISHER, *supra* note 5, at 186-98. However, such proposals are usually struck down because of the overall aversion to governmental intervention in the regulation of content markets (which at times could be understood as impediments on the free speech rights of various market actors). *Id.*

241. *See* Thomas Vander Wal, *Understanding Folksonomy (Tagging that Works)* (2006), http://s3.amazonaws.com/2006presentations/dconstruct/Tagging_in_RW.pdf. Vander Wal coined the term "folksonomy" to describe this "bottom-up" process. *See id.*

242. For more on this term that is used for bottom-up processes used for sorting and ranking, see http://www.answers.com/topic/folksonomy. For a current critique of this model's problems by Clay Shirky, see Clay Shirky, *Folksonomy,* MANY 2 MANY, Aug. 25, 2004, *available at* http://many.corante.com/archives/2004/08/25/folksonomy.php. As mentioned, this dynamic has been addressed by Benkler. BENKLER, *supra* note 33, at 76-80. For a somewhat critical view of these dynamics, see STEVEN JOHNSON, EMERGENCE 159-62 (2002).

243. Fisher refers to the use of such distribution methods, not in the context of the mandatory ACS model, but of the voluntary "coop" one. *See* FISHER, *supra* note 5, at 254-55. Although he praises this model (while referring to Benkler's work regarding Slashdot.com), he does not see its endorsement as part of the ACS, nor discuss its shortcomings or ways it could be promoted. *Id.*

tical whether these dynamics, will counter the concerns voiced above. Although this dynamic seems to reflect a grass roots movement of sorting and ranking, it could be manipulated by powerful interest groups, which would promote specific forms of content, while again leaving independent and unaffiliated artists outside the loop. This would lead to the resurfacing of the problems addressed above, and the creation of a high barrier to entry on the one hand, and an unattractive intermediary on the other.[244]

This skepticism regarding the role of folksonomies stems from my belief that they could be tainted and manipulated in various ways. First, they could be subjected to gaming by external entities, which will rely on many of the dynamics addressed above to generate results that are favorable to their clients. Artists with financial or other backing would be able to apply various technological means to simulate broad satisfaction with their content that would lead to a high ranking. Thus, the fact that this is a "many-to-many" medium can turn out to be a weakness. Folksonomies could be gamed by internal entities as well; in other words, the apparently-neutral entities running the ranking and sorting websites might be overtaken by a large media conglomerate that would secretly or actively promote "their" artists throughout the various rankings, regardless of the "bottom-up" process. To those who believe these predictions are somewhat pessimistic, I merely mention the growing interest among today's large media conglomerates and moguls in social network websites which generate folksonomies of their own. Recently, News Corporation (News Corp.) has purchased the extremely popular MySpace.com website, which has become a successful platform for launching and distributing new forms of music through a sophisticated recommendation and accreditation system. While News Corp.'s plans and intentions for MySpace are unclear, the potential risk for the "contamination" of the bottom-up process to meet the objectives of the media moguls is apparent.[245]

Could the problems and threats to this form of distribution be resolved? Possibly. But to do so will require additional research regarding these issues—research that should be funded by the ACS fund (that is funded by the levy described above) should the model be implemented. In addition, the fund should finance non-affiliated folksonomy sites,

---

244. Benkler frames this concern as the fear that "money" would still allow specific entities to buy their way into a dominant market position in the connected world (note that Benkler concludes that the end of the day this problem is substantially mitigated in the Internet medium). BENKLER, *supra* note 33, at 234.

245. *See* Steve Rosenbush, *News Corp.'s Place in MySpace*, BUSINESSWEEK ONLINE, July 19, 2005, *available at* http://www.businessweek.com/technology/content/jul2005/tc20050719_5427_tc119.htm. Note that News Corp. is still cautious about the ways in which it would use this new addition to its group, yet already mentions the use of this tool to promote its own content. *Id.*

which will provide limited incentives to the sorters and rankers, while assuring that these mechanisms remain untainted.

Another possible option for content distribution in the ACS content market, which would not lead to the unnecessary empowerment of intermediaries (be them new or old), is of content distribution among smaller circles of users who belong to virtual communities. Here, as opposed to the dynamic mentioned above, I refer to one that could be defined as "few-to-few" and therefore somewhat insulated from the disadvantages inflicting the broader folksonomies.[246] Within these communities, members could inform others of various forms of content they have "stumbled upon" online, and which they could recommend to other community members. Such a recommendation will carry merit as it is both made by a community member whom has earned the other members' trust in the past, and who they know has preferences and tastes that are similar to their own. As Eben Moglen has pointed out long ago, this distribution model will allow for the very quick spreading of content and ideas, while taking full advantage of the Internet's robust and worldwide network.[247] Clearly this distribution model sharply differs from content distribution in the offline world that has been mostly premised on the "broadcast" or "one to many" model, according to which one central source sets out to meet the preferences and tastes of a very broad audience. This model could supplement or even substitute other models that rely on "central" intermediaries that provide general recommendations to the broad public (that could result from both a top-down and a bottom-up process).

The Internet allows for this alternative model for content distribution to transpire while making use of communities that are created online and convene in several possible settings (such as chat rooms, mail-lists, forums and others)[248] that are referred to as "social software."[249] These communities, of various sizes, are formulated to address or discuss specific, yet mutual topics, which could be related to hobbies, work, neighborhood, and past experiences[250] or are premised on a common trait

246. Clearly the empirical question that lurks in the midst concerns the line between a mere "community" to an overall "many-to-many" folksonomy. I will not address this matter here and leave it for future research, yet mention that many of the benefits of the "community" come from both a feeling of intimacy and familiarity with the other community members (notions that are absent on the broader scale). The question, as to the point at which such intimacy and familiarity disappear is an extremely difficult one. For this issue, I would use the definition adopted by Benkler—"larger than a dozen, smaller than a few hundred."

247. *See* Eben Moglen, *Comment: Liberation Musicology*, THE NATION, Feb. 22, 2001, *available at* http://www.thenation.com/doc/20010312/moglen.

248. *See* BENKLER, *supra* note 33, at 357 (explaining that the Internet is creating many new looser social networks.

249. *See id.* at 373.

250. *Id.* at 368.

or attribute of all the participants.[251] The dynamics of these virtual communities have led to several astounding accomplishments, such as the creation of elaborate software tools and detailed content repositories—and all without a "classic" top-down structure.[252] In addition, the Internet is filled with anecdotal examples of various works that gained worldwide exposure and fame after being passed on through word of mouth.[253]

This form of distribution carries numerous benefits. Because of its diffused nature, it does not support the creation of a small yet powerful group of intermediaries which have an overall grip over a large portion of society, and as such could leverage their power towards the artists and consumers. Therefore the existence of content distribution within these communities could mitigate concerns of overpowering intermediaries in the ACS model.[254] In addition, the diffused nature of this model makes it considerably harder to game. Within these communities members "know" the others by their specific reputation, and therefore are less prone to manipulation by external or internal forces.[255] For these reasons, I believe this form of content distribution is preferable to the use of the folksonomies mentioned above.

However, content distribution through the use of such virtual communities is not a concept without challenges and problems. This field as well has generated an enormous amount of recent scholarship, which addresses these issues. A problem that is constantly mentioned when addressing these dynamics is that of motivation[256]: How can society motivate individuals to partake in the community dynamic, and in that way both contribute recommendations and receive feedback within these circles? Clearly participation and motivation to participate are key elements, as without them, consumers will revert to the "customary" modes of content consumption (and collecting information about such content)

---

251. For a recent survey as to these various realms, see James Scott & Thomas Johnson, *Bowling Alone But Online Together: Social Capital in E-Communities*, 36 J. COMMUNITY DEVELOPMENT SOCIETY 9 (2005).

252. For instance the free software movement that led to the development of Linux. On this issue and for additional examples, see BENKLER, *supra* note 33, at 59-74.

253. For example the famous "JibJab" cartoons. *See* Funny Videos, Pictures & Jokes at JibJab.com, http://www.jibjab.com (last visited October 25, 2006).

254. Benkler makes a similar argument with regard to the broader, Internet context. *See* BENKLER, *supra* note 33, at 255. In other words, he argues that a "thin tail" of user traffic in peer-to-peer and other social networks mitigates many of the troubles of media concentration online. *Id.*

255. Recent scholarship indeed indicates that "successful communities" include users who provide personal information about themselves, and in that way contribute to their reputation and the accreditation of the content they convey. *See* Chris Forman, Anindya Ghose & Batia Wiesenfeld, *A Multi-Level Examination of the Impact of Social Identities on Economic Transactions in Electronic Markets* (July 2006), *available at* http://ssrn.com/abstract=918978.

256. On these issues, see the work of Paul Resnick which presents several projects and papers on the issue of motivation in this context. *See generally* Paul Resnick's Home Page, http://www.si.umich.edu/~presnick/ (last visited Oct. 20, 2006).

with all their shortcomings. Another problem arising in this context is gaming within these circles. Namely, the fear that interested parties would penetrate these communities, and provide recommendations that appear trustworthy, yet reflect the actions of interested parties and are financed by well-to-do artists and their intermediaries. Regarding this last issue, there is some hope, as recent experiences with recommendation systems in the e-commerce context show great progress in overcoming this difficulty. These sites, as well as other virtual communities, have been struggling with the challenge of identifying fraudulent recommendations and have begun to learn to neutralize them.[257]

In view of the advantages of content distribution through virtual communities, I believe the implementation of the ACS model must include measures to strengthen this mode of distribution. One way to achieve this is by subsidizing (again, from the ACS fund collected through the abovementioned levy) computer equipment, support, and other related expenses for community centers and other not-for-profit organizations, to promote the formation of online mechanisms which will facilitate these social networks. Such a subsidy will assure these social networks will not be connected to any commercial entity that might taint the content distribution process. Other funds could be used to motivate participants in these communities by providing limited prizes. Yet clearly additional research is required to establish other ways to achieve this objective.

In conclusion of this article's analysis as to the role and power of intermediaries, the shift to the ACS model must also include an examination as to how content would be distributed in a market operating in accordance to this model. This analysis must look into ways to promote distribution through alternative platforms and networks. It must also examine whether these platforms and networks will prove to be broad and robust enough to effectively compete and even replace the distribution mechanisms controlled by today's media conglomerates. Only by specifically addressing distribution, could the ACS model meet its important objective of realigning the power balance between artists, intermediaries and consumers.

3. The Outcome of the Model—Content and Content Producers

Beyond the model's effects on the media market and its intermediaries, the shift to the ACS model might profoundly change the consumption patterns of content online. These changes in consumption patterns will be followed by changes in the compensation authors, performers, and artists receive for creating such content. These last changes, in turn,

257. For example, Ebay has enhanced its actions against those manipulating vendors' feedback. *See* EBAY, *Frequently Asked Questions: Feedback Manipulation Policy*, *available at* http://pages.ebay.com/help/announcement/22.html (last visited Nov. 21, 2006).

will presumably lead to changes in the content which is produced by the market.[258] In the following paragraphs I examine these changes and their possible adverse effects on society. Thereafter, I suggest several amendments to the model to avoid these problematic effects, some of which resemble the alternative distribution mechanism that were mentioned above in a different context.

A key element to this part of the analysis is an assumption that not only will the ACS model be accepted for governing the compensation for content use online, but that this model's grasp will reach beyond this limited realm and pertain to a significant amount of *all* content consumption. This assumption is required, as should this not be the case, authors and artists will continue to receive compensation through today's conventional channels. These other media channels (such as retail, TV, etc.) will not be affected by the new ACS and dynamics it creates, and would offset the specific market and social forces of the online model. This underlying assumption regarding the breadth of ACS is not far fetched; the Internet is hardly a confined universe of content use and consumption. Today users download music, and through burned CDs or other portable devices enjoy this content when they are away from their computers. Clearly, in the very near future, the shift of content from the online world to the offline world (as well as vice versa) would be seamless, and the technical challenges of shifting and streaming video content from computers to TV sets will be resolved.[259] With the abundance of free and high quality content available online (through the ACS model), it is fair to assume that with time, this realm would become a hub of content exchange and a prime source of compensation for artists. It is at this point of time where this segment of the analysis will turn relevant. However, as identifying this point of time would be difficult and applying changes to the ACS model at a late stage costly and complex, I believe these matters are best discussed and addressed at the early stage of planning the model, as I do now.

In the following paragraphs, I argue that the switch to the ACS model will generate changes in the consumption pattern of consumers.[260] The reason for such changes will be additional limitations and pressures to be set on the users' attention span upon consuming content. These

---

258. This argument is premised on the notion that artists, when deciding what form of content to create, take into account the amount of profit they might reap from its subsequent sale. Not all take this notion as a given. *See* Litman, *supra* note 15, at 28; Moglen, *supra* note 26.

259. Bob Zitter, Time Warner, Summit on Intellectual Property and Digital Media Conference, The Cable Center, University of Denver (May 22, 2006). Apple has recently announced it is developing the "iTV"—an application that would bridge the PC and the TV with ease, thus resolving the challenge mentioned in the text. *See* Nick Wingfield & Merissa Marr, *Apple Computer Aims to Take Over Your Living-Room TV*, WALL ST. J., at B1 (Sept. 13, 2006).

260. Fisher generally acknowledges that with the shift to the ACS model, such changes might occur, but does not elaborate as to their nature or their subsequent effects. FISHER, *supra* note 5, at 237.

pressures will come with the adoption of the ACS model, which will present users with many millions of content options online, at a marginal cost of zero. Such a variety provided to consumers with no financial constraints, would possibly create a tendency to engage in constant "flipping;" mercilessly skipping from one form of content to another at the moment they are displeased with what they are receiving. Such behavior would resemble television viewing in a multi-channel medium, which offers thousands of channels to viewers with only limited time and attention span. In this latter example, many viewers indeed respond to this variety by engaging in constant "flipping" switching from one channel to another.

Such enhanced "flipping" behavior online could have several outcomes. At first, it might allow content consumers to become more demanding in their pursuit of quality content, and less willing to settle for mediocre products. Thus, content that rises to the top of the "most watched and listened to" list would be better than the content that is at the top in today's market dynamic. This point could be strengthened by a recent controversial study comparing prime time television programs of today and those shown in previous decades.[261] This study argues that some of today's leading television programming introduces shows with many interweaving story lines, intelligent writing, and thicker and more intense plots. Given the fact that deciding upon the quality of such content products is an extremely subjective task, these factors might objectively indicate that indeed the content available has improved substantially. One possible reason for this improvement could be the intense competition for the consumers' attention in a multi-channel age.

Yet the users' limited attention span, and the "flipping" phenomenon may have an adverse affect as well. This adverse effect concerns forms of content that are of social significance, yet are only fully appreciated after being experienced, watched, or heard in their entirety and perhaps even only after several such "experiences." I will refer to these works as "Masterpieces."[262]

In a cultural environment that allows for constant "flipping" between forms of content that are all available at marginal cost of zero, it could be assumed that users will not provide Masterpieces the second and third chances that are required for their full appreciation. Therefore, the depth and genius of these works will not be recognized and acknowledged. Rather, users will immediately switch to other forms of content that would satisfy their immediate "needs" for entertainment and leisure.

---

261. For this analysis, see STEVEN JOHNSON, EVERYTHING BAD IS GOOD FOR YOU 62-116 (2005).

262. Clearly this definition does not comply with the common definition of such works. In addition, there are, of course, many examples of great works of art which have immediately achieved commercial success and broad public appreciation.

Should this dynamic take place, it could generate the following chain reaction: Users and consumers will not acknowledge the existence and value of such works, and therefore will not review these forms of content in their entirety (which, as mentioned above is a prerequisite for generating compensation to the authors in view of "gaming" concerns), nor return to them at a later time. In addition, they would not mention the existence of such works to their peers and friends (while acting as intermediaries in the various networks and structures mentioned above) who would therefore not learn of these works and refrain from their usage. Thereafter, authors and creators of such Masterpieces will suffer a drop in the compensation they receive in an ACS-governed market. Finally, content producers will choose to under-produce these forms of content (or Masterpieces) in view of the limited compensation they will reap. This, in turn, would harm society in general, which would now be deprived of important cultural resources.[263]

It is interesting to note, as a comparison, that today's model for compensation (in which users directly compensate authors, usually in advance, for access to content) provides for a more supportive environment for Masterpieces and their authors. Here, the authors of such works might gain access to the market by receiving support from a large media conglomerate that attends to funding and distribution after deciding to promote this specific artist and her work (to meet personal, social or even financial objectives).[264] Consumers, who would purchase such works (acting upon the recommendation of the media conglomerates in their capacity as content intermediaries) would be more willing to experience such works in their entirety and perhaps even several times. This is because they have already paid for such content, and would be willing to devote more time to it,[265] to justify (in their eyes, at least) such past expenditures. They will also do so, because alternative forms of content (which must be directly purchased) are scarce. A result of devoting time

263. As mentioned above, this part of the analysis is somewhat shaky, as it is based on a problematic premise—that the authors of Masterpieces are indeed motivated by the level of compensation they are likely to reap. Those disagreeing with this argument might further argue that in this context, authors of Masterpieces rarely take into account the success of their writing, and are lead by other intrinsic or extrinsic incentives (and many of which rarely receive meaningful compensation for such works during their lifetime). I would argue, that at least some authors of Masterpieces are not motivated by internal incentives alone, and therefore this discussion is at least somewhat relevant.

264. *See* Baker, *supra* note 233, at 878. Baker explains that in the publishing context, publishers have been known to invest their profits from lucrative publications in important projects that will probably not make them any money. However, he also explains that these practices are quickly disappearing as this business as well is becoming "bottom line" oriented. On this issue, see also Andre Schiffrin, THE BUSINESS OF BOOKS: HOW INTERNATIONAL CONGLOMERATES TOOK OVER PUBLISHING AND CHANGED THE WAY WE READ 91, 95, 108 (2000).

265. My assumption stated in the text as to the consumers willingness to allocate more time and attention resources to products they have already purchased and paid for relies on the "Sunk Costs Fallacy"—a cognitive phenomenon, according to which individuals want to cut their losses, but continue to engage in actions that allow them to capitalize on costs they already incurred. For more on this phenomenon, see Wikipedia, http://en.wikipedia.org/wiki/Sunk_cost_fallacy.

to this work would be the users' ability to recognize the true value of the Masterpieces and thereafter revisit such content and recommend it to others (who would go ahead and purchase such content). This, in turn, will provide for additional compensation to the Masterpiece right holders.

Yet the challenges the ACS model presents to the compensation for and production of Masterpieces might be countered by several advantages this model has in store; advantages that arise from the model's distribution dynamics. The ACS models can easily integrate and promote content distribution networks that employ small groups congregating online, and allow for the quick and efficient diffusion of ideas and content between people with a common interest and in many cases similar traits and preferences.[266] These "virtual communities" can prove a fertile ground for effectively distributing information about Masterpieces in a dynamic that could overcome the "flipping" threat addressed above. This is because, within such a community, members recommend to each other works they believe will meet their specific taste and liking. Within these circles, it is fair to assume that users would be willing to accept recommendations and act upon them, even if this would mean resisting the "temptation" to flip to another form of content if the recommended work seems at first unsatisfactory. They will do so because users would learn from experience that recommendations given within these systems will prove worthy of their time and attention—even though at first glance they might not appear as such. In addition, within these secluded circles, members will know a great deal about each other—a factor that would contribute to the effectiveness of the recommendation.[267]

The notion of distributing information regarding Masterpieces through social networks is of course far from novel—and is probably one of the main ways in which information regarding these forms of content travels. However, the online social networks include several important improvements and advantages: They allow for the creation of broader and richer communities that are not hindered by geographical distances and bring together people with common interest from very different backgrounds. Furthermore, in a market operating according to the ACS model, such networks not only can recommend the work but provide it directly, either by linking or allowing downloading.

Therefore, to sum up this point, examining the possible effects of the ACS model on the creation of Masterpieces provides us with an addi-

---

266. In many cases, this common interest is the reason for the creation of the online forum or community (for instance, a recommendation on a new jazz album in a jazz-fan forum). In others, the common interests might be incidental (for instance, a neighborhood forum, in which one neighbor who know the others well, recommends a book she believes they will appreciate).

267. *See* Forman, Ghose & Wiesenfeld, *supra* note 255 (discussing the importance of personal information regarding participants to the success of the internal dynamics of virtual communities).

tional reason as to why the implementation of the ACS model should include tools, funds, and applications that would promote social online networks. With such robust networks, the potential damages the pressure and limitation on the users' attention cause Masterpiece production, could be substantially mitigated.

However, with the implementation of the ACS model, policy makers must remain alert and constantly examine whether this model facilitates sufficient exposure, consumption, and thereafter, production of Masterpieces. Should this new marketplace lead to underproduction of Masterpieces, policymakers must consider promoting the production of such works using other, more direct measures. For instance, the administrative agency could be required to set aside limited sums from the overall, levy-financed, fund. These sums could be than used towards the direct promotion of Masterpiece production, by providing prizes for extraordinary works and scholarships for their authors. This dynamic has indeed been suggested (for other reasons) by some of the ACS scholars.[268] It is also commonly applied in Europe with regard to the funds collected via levies to compensate artists for non-authorized personal uses.[269]

Clearly this last suggestion presents several shortcomings. The funds and scholarships mentioned will be distributed by a committee of "experts" that might be biased, engage in elitism and paternalism when deciding what does and does not amount to a Masterpiece worth financing and promoting. The famous controversies surrounding the National Endowment for the Arts would surely be echoed,[270] with all their political implications, as politicians decide what form of content is worthy of public funding. However, in specific instances, this partial solution will indeed be required in order to allow the sponsoring of Masterpieces, which might be lost as a result of the contemplated shifts in compensation policy and trends of content consumption.

CONCLUSION

In this article, I chose to address and contribute to a new line of scholarship that offers an interesting policy solution to the difficult challenges of today's digital content market. The ACS scholars have drawn

---

268. *See* Lunney, *supra* note 15, at 915-16. (arguing for such allocation from the fund to provide to marginalized artists). *But see* Netanel, *supra* note 2, at 58 (arguing that such practices should be limited, as they might lead to "rent seeking"). I agree with Netanel's overall concern, but disagree with his result given my analysis of Masterpiece consumption above.

269. *See* HUGENHOLTZ, GUIBAULT, & GEFFEN, *supra* note 129, at 68-69 (summarizing the states that allocate some of the funds to a "social fund").

270. FISHER, *supra* note 5, at 217. For more information regarding the case law and controversy this fund involved, see Freedom of Expression at NEA, http://www.csulb.edu/~jvancamp/intro.html (last visited Oct. 20, 2006).

out an elaborate and thoughtful blueprint for an extensive and innovative model. I have attempted to continue this line of scholarship by sharpening issues that remain open, pointing out some matters that require additional thought and analysis and examining others that should be somewhat changed. I have emphasized the need to develop mechanisms for content distribution (in addition to those focused on compensation) and mentioned several existing online models that might be fitting for the task.

Yet, at the end of the day, I assume that many readers will still remain unconvinced by the arguments set forth and maintain their position that the implementation of an ACS model (even with the improvements suggested herewith) is politically, technologically, or economically infeasible. To these readers, I say that the journey this article draws out was not traveled in vain. The analysis conducted above has taught us important lessons regarding the business and policy implications of the development of new technological tools in an ever-changing content market. These lessons will prove fruitful when facing future challenges that will be sure to arise in the Internet society.

DISPUTED-CONSENT SEARCHES: AN UNCHARACTERISTIC STEP TOWARD REINFORCING DEFENDANTS' PRIVACY RIGHTS

INTRODUCTION

The Fourth Amendment of the U.S. Constitution protects citizens' rights to privacy by prohibiting unreasonable government searches and seizures.[1] Government officials, however, may overcome this barrier by obtaining a warrant based upon probable cause.[2] Courts assess probable cause through a totality of the circumstances approach that, in order for a finding, requires information sufficient to justify a reasonably prudent individual in believing that an offense has been committed.[3] In general, a person claiming a violation of his or her Fourth Amendment rights must satisfy two requirements. First, that individual must prove that a search or seizure actually occurred within the meaning of the Fourth Amendment.[4] Second, the person must show that the search or seizure violated reasonableness.[5] Failure to establish both requirements destroys any Fourth Amendment claim. However, the warrant requirement embodied in the Fourth Amendment is not unconditional, and several exceptions exist which may allow police officers to engage in a warrantless search. Specifically, exceptions include: searches incident to a lawful

1. U.S. CONST. amend. IV.

2. *Id.*; *see also* Akhil Reed Amar, *The Fourth Amendment, Boston, and the Writs of Assistance,* 30 SUFFOLK U. L. REV. 53, 54-65 (1996) (providing a general discussion of the Fourth Amendment).

3. *See, e.g.*, Brinegar v. United States, 338 U.S. 160, 175-76 (1949) (explaining further that probable cause is a practical, non-technical consideration based on the everyday observations of reasonable people).

4. *See, e.g.*, PHILLIP A. HUBBART, MAKING SENSE OF SEARCH AND SEIZURE LAW: A FOURTH AMENDMENT HANDBOOK 10-12 (2005) (explaining that standing requires a showing that 1) the complaining party's privacy right is at issue, 2) a government agent conducted the search or seizure, and 3) the search or seizure was of the party's person, house, paper, or effects).

5. *Id.* at 167 (explaining that unreasonableness determinations require a balancing of the degree and nature of intrusiveness against the weight of the government interest served by the search or seizure); *see also* Katz v. United States, 389 U.S. 347, 357 (stating generally that "searches conducted outside the judicial process, without prior approval by judge or magistrate, are per se unreasonable under the Fourth Amendment – subject only to a few specifically established and well-delineated exceptions").

arrest,[6] searches made under exigent circumstances,[7] moving vehicle searches,[8] and searches made upon voluntary consent.[9]

This comment addresses the voluntary consent exception to the warrant requirement, as applied in the recent *Georgia v. Randolph*[10] decision. Initially, the Supreme Court qualified the consent exception as applicable against the individual who provided consent to police.[11] However, soon thereafter, the Court verified the legitimacy of consent given by a third party, holding that any person reasonably presumed to possess common authority over the premises could consent to a warrantless search.[12] Specifically, the Court upheld the lawfulness of a search made pursuant to consent given by a *present* occupant against the subsequent objection of the *absent* nonconsenting co-occupant.[13] Although these cases were instructive for interpreting the consent exception when the nonconsenting party was absent, the Court did not address the constitutionality of a warrantless search when two or more present co-occupants disagreed as to whether police could enter. Tellingly, lower courts across the nation facing this issue reached opposite holdings concerning the validity of disputed consent searches based on differing interpretations of the relevant Supreme Court precedents.[14]

Spurred by this unsettling disparity, the Supreme Court recently addressed the issue in *Georgia v. Randolph*.[15] In Part I, this comment will provide an overview of the relevant Supreme Court decisions that evaluate the Fourth Amendment warrant requirement as it relates to the voluntary consent exception. Part II will review the majority, concurring, and

---

6. *E.g.*, Chimel v. California, 395 U.S. 752, 763 (1969) (holding that "it is reasonable for the arresting officer to search the person arrested in order to remove any weapons that the latter might seek to use in order to resist arrest or effect his escape.").

7. *E.g.*, Cupp v. Murphy, 412 U.S. 291, 295-96 (1973) (defining exigent circumstances as those that would endanger the lives of the responding officers, or of others, in the absence of a timely investigation).

8. *E.g.*, Carroll v. United States, 267 U.S. 132, 153-58 (1925) (permitting warrantless searches of moving vehicles if probable cause suggests the vehicle contains contraband, instruments of a crime, or evidence of a crime).

9. *E.g.* Schneckloth v. Bustamonte, 412 U.S. 218, 243 (1973) (permitting warrantless searches without probable cause when express or implied consent is given).

10. 126 S. Ct. 1515 (2006).

11. *Schneckloth*, 412 U.S. at 243-45 (stating, however, that under some circumstances a third party may give lawful consent).

12. *See* United States v. Matlock, 415 U.S. 164, 171 (1974) (holding that a co-occupant possessing common authority over the home may consent to a police search against another occupant of that home, even in spite of that absent, nonconsenting party's refusal); *see also* Illinois v. Rodriguez, 497 U.S. 177, 188-89 (1990) (holding that whether the occupant truly possesses common authority over the premises is immaterial in determining the legality of the consent given; instead, the only consideration courts must address is whether police were objectively reasonable in thinking that individual possessed common authority).

13. *Matlock*, 415 U.S. at 170.

14. *Compare* Primus v. State, 813 N.E.2d 370, 376 (Ind. Ct. App. 2004) (holding that a search of Defendant's premises upon consent of his roommate was valid despite the defendant's presence), *with* State v. Randolph, 604 S.E.2d 835, 837 (Ga. 2004) (holding that a search of Defendant's home upon consent of his wife, but in his presence and with his clear refusal, was invalid).

15. Georgia v. Randolph, 126 S. Ct. 1515, 1528 (2006).

dissenting opinions in *Georgia v. Randolph*.[16] Finally, Part III will analyze the Court's decision by focusing on: 1) the practical implications of the decision, specifically with regard to the inherent conflict between personal privacy and law enforcement, 2) the holding's place within the established jurisprudential trend of limiting privacy through the third-party consent legal standard, 3) the benefits and latent costs of the logic employed in the *Randolph* opinion, and 4) potential solutions available for addressing the problems likely to arise in the decision's aftermath. Ultimately, this comment will conclude that *Randolph* occupies a position of limited authority because its interpretation of the Fourth Amendment applies only to the rare situation in which multiple, present co-occupants disagree on the issue of consent. Further, it will be suggested that the *Randolph* holding encourages police to circumvent the protective nature of the holding. And finally, this comment will suggest that *Randolph* is likely to generate confusion amongst lower courts and law enforcement agents, because it is presently unclear whether the holding signals a reversal of the Court's view of Fourth Amendment searches, or merely a deviation from a viable trend.

I. BACKGROUND

Among the several exceptions to the warrant requirement, investigating officers find the voluntary consent exception particularly attractive as an alternative to obtaining a warrant.[17] This is true for several reasons. First and foremost, police often seek consent from a suspect because it permits entry in situations where probable cause does not exist to substantiate a warrant.[18] Further, even when probable cause does exist, consent searches function as an instrument of convenience that police can use to skirt the burdensome process of obtaining a warrant.[19] This factor is particularly true when police would otherwise have to travel long distances to acquire a warrant.[20] Last, officers often prefer consent searches in lieu of a warrant because the scope of a permissible search may extend to areas it likely would not reach under the auspices of a warrant.[21]

16. *Id.* at 1515-43.

17. WAYNE R. LAFAVE, JEROLD H. ISRAEL & NANCY J. KING, CRIMINAL PROCEDURE 250 (4th ed. 2004) (stating that "[t]he practice of making searches based on consent is by no means a disfavored one.").

18. *See* WAYNE R. LAFAVE, SEARCH AND SEIZURE: A TREATISE ON THE FOURTH AMENDMENT 5 (4th ed. 2004) (explaining that the scope of a consent search is frequently broader than the scope of a search justified by a warrant, because the consenting party rarely conditions or qualifies the scope of consent).

19. *See id.* at 4-5.

20. *See id.* at 5.

21. *See id.* (explaining that the objectively reasonable scope of a consent search is frequently broader than that justified by a search warrant, because the consenting party rarely conditions or qualifies the scope of consent).

The general rule governing the voluntary consent exception to the Fourth Amendment warrant requirement holds that the consenting individual must give permission voluntarily and under non-coercive circumstances.[22] Courts regularly uphold both written and verbal consent, whether given pursuant to a police request, or of the suspect's own volition. And further, the Supreme Court has upheld the lawfulness of consent searches even under circumstances in which the suspect is not aware of the right to refuse,[23] as well as situations where police have already taken the suspect into custody.[24] In addition to the requirement that consent to a police search be given voluntarily, these searches are limited in scope by a standard of "objective reasonableness." [25] Specifically, "objective reasonableness" evaluates what a reasonable person would understand the scope of consent to include after considering the interchange between the officer and the suspect.[26] The Supreme Court has addressed the voluntary consent exception to the Fourth Amendment warrant requirement on several occasions.[27] Those decisions are briefly characterized here.

A. Reasonableness as the Standard

In *Stoner v. California*,[28] the Court considered whether a hotel clerk could lawfully provide police with consent to search a guest's room. Though recognizing that a clerk does possess some limited right of entry for purposes such as cleaning or inspecting the room, the Court decided that a hotel employee does not retain the authority to consent to a search against the guest.[29] The Court reasoned that police have no rational foundation to conclude that a hotel clerk possesses such authority.[30] Three years later, in *Katz v. United States*,[31] the Court elaborated on the

22. *Schneckloth*, 412 U.S. at 248 (holding that the prosecution has the burden of showing by a preponderance of the evidence that consent was "in fact voluntarily given and not the result of duress or coercion, express or implied.").

23. *Id.* at 248-49 (stating, however, that a person's knowledge of the right to refuse should be taken into account in the totality of the circumstances determination of whether the consent was voluntary).

24. *See, e.g.*, United States v. Watson, 423 U.S. 411, 424 (1976), *but see* Florida v. Royer, 460 U.S. 491, 507-08 (1983) (holding that, on the contrary, when consent given by a suspect *unlawfully* taken into custody is acted on, it should be considered involuntary and therefore in violation of the Fourth Amendment).

25. *See* Florida v. Jimeno, 500 U.S. 248, 249 (1991).

26. *Id.* (referring to "area of consent" as meaning the actual physical parts of the premises that the consenter has given the officer permission to search).

27. *See generally* Nancy J. Kloster, Note, *An Analysis of the Gradual Erosion of the Fourth Amendment Regarding Voluntary Third Party Consent Searches: The Defendant's Perspective*, 72 N.D. L. REV. 99, 104-15. (1996) (providing a detailed analysis of the Court's trend in deciding co-occupant consent cases).

28. 376 U.S. 483 (1964).

29. *Stoner*, 376 U.S. at 489.

30. *Id.* at 488; *see also* Chapman v. United States, 365 U.S. 610, 616-18 (1961) (holding that a landlord may not consent to a search of the home leased to a tenant, because it is not reasonable for police to assume that a landlord possesses that authority).

31. 389 U.S. 347 (1967).

meaning of "reasonableness" by holding that Fourth Amendment's protection against warrantless searches was not confined to the home.[32] Instead, this protection extended to all areas where the suspect enjoys a reasonable expectation of privacy.[33] Of particular influence in subsequent cases was Justice Harlan's concurring opinion which established a two-prong inquiry focusing on whether: 1) the suspect had a subjective expectation of privacy, and 2) society would view that expectation as reasonable.[34]

B. Assumption of the Risk and Third Party Consent

In *Frazier v. Cupp*,[35] the Court upheld the legality of consent given by a third-party based on an assumption of the risk doctrine. Specifically, the Court held that whenever two or more people jointly use an item or place, each individual implicitly assumes the risk that the co-occupant (or co-possessor) might share that item or place with others.[36] Two years later, in *Coolidge v. New Hampshire*,[37] the Court considered whether the husband-wife relationship supported an inference that either spouse could consent to a search against the other.[38] Acting on the standard established in *Katz*, the Court determined that this relationship did support such an inference because, unlike the situation in *Stoner*, it is reasonable for police to assume that one spouse has the authority to consent to a search of the home in the other spouse's absence.[39] Shortly thereafter, the Supreme Court decided the seminal case of *United States v. Matlock*,[40] and held that any time two or more individuals possess "common authority" [41] over a premises, the consent of any individual alone trumps the subsequent objection of another, absent co-occupant.[42] In *Minnesota v. Olson*[43] however, the Court upheld the privacy rights of an overnight social guest arrested in the course of a warrantless search, and held that houseguests have privacy rights parallel to those of occu-

32. *Katz*, 389 U.S. at 359 (finding, under the factual circumstances, that this protection extended to the content of phone conversations made in a public phone booth).

33. *Id.* (stating specifically that "[t]hese considerations do not vanish when the search in question is transferred from the setting of a home, an office, or a hotel room to that of a telephone booth. Wherever a man may be, he is entitled to know that he will remain free from unreasonable searches and seizures.").

34. *Id.* at 361 (Harlan, J., concurring).

35. 394 U.S. 731 (1969).

36. *Frazier*, 394 U.S. at 740 (holding that when two cousins shared use of a bag, one cousin had authority to consent to a police search because through their joint use, both parties assumed the risk that the other would share the contents of the bag).

37. 403 U.S. 443 (1971).

38. *Coolidge*, 403 U.S. at 489-90.

39. *Id.* at 474, 487-90.

40. 415 U.S. 164 (1974).

41. BLACK'S LAW DICTIONARY 292 (8th ed. 1999) (defining common-authority as: "the principle that a person may consent to a police officer's search of another person's property if both persons use, control, or have access to the property.").

42. *Matlock*, 415 U.S. at 171.

43. 495 U.S. 91 (1990).

pants because "it is unlikely that [the host would] admit someone who want[ed] to see or meet with the guest over the objection of the guest."[44]

C. Further Limiting Privacy

In *Schneckloth v. Bustamonte*,[45] the Court faced the issue of whether, in order for the search to comply with the "voluntary" requirement,[46] police were required to inform the consenting occupant of his or her right to refuse the search. In light of the expansive view the Court had taken with regard to warrantless searches, it unsurprisingly held that a suspect did not have to be aware of the right to refuse the warrantless search in order for consent to remain valid.[47] Finally, in *Illinois v. Rodriguez*[48] the Court took its greatest leap yet in restricting defendants' privacy rights. In *Rodriguez,* the Court considered a third-party consent search when the individual who provided consent to police did not actually reside at the home, and had no common authority over it.[49] Nevertheless, the Court held that whether the third-party actually possesses common authority over the premises is irrelevant to the determination of the search's validity.[50] Instead, courts must only address whether it was objectively reasonable for police to have believed that the third-party had common authority.[51]

Therefore, prior to *Georgia v. Randolph*, the Court established a clear trend of gradually eroding defendants' Fourth Amendment privacy rights, while correspondingly broadening the scope of lawful police searches. By late 2005, an opportunity arose for the Court to address the consent exception as it applied to multiple co-occupants possessing common authority over the premises, but who disagreed over consent to a police search.[52] In light of precedent, as well as the Court's trend of limiting privacy, little hope must have remained for pro-privacy advocates seeking reinstatement of the protective measures once apparent in the Fourth Amendment.

---

44. *Olson*, 495 U.S. at 99.

45. 412 U.S. 218 (1973).

46. *See, e.g., Schneckloth*, 412 U.S. at 227 (referring to the requirement that consent be given voluntarily, and under non-coercive circumstances).

47. *Id.* (holding that, as long as the consenter isn't coerced by police, the prosecution need not show that the consenter had knowledge of his or her right to refuse in order to establish that consent was voluntary).

48. 497 U.S. 177 (1990).

49. *See Rodriguez*, 497 U.S. at 179.

50. *Id.* at 188-89.

51. *Id.* (holding that the search was valid despite the consenting third-party not having common authority, because it was objectively reasonable to assume that party had common authority when she was holding a baby and referring to the apartment as "ours").

52. *Randolph*, 126 S. Ct. at 1520 (stating "[n]one of our co-occupant consent-to-search cases . . . has presented the further fact of a second occupant physically present and refusing permission to search").

II. *GEORGIA V. RANDOLPH*

A. Facts

In late May of 2001, Scott and Janet Randolph separated when Janet, along with their son, left the family's Georgia residence and moved to her parents' home in Canada.[53] Janet and her son did not stay long though, and they returned to Scott at the family's home in Georgia within two months.[54] On the morning of July sixth, police responded to a domestic dispute at the Randolph home.[55] When the police arrived, Janet informed the officers that Scott had removed their son from the home, and that he would not tell her where the child was.[56] Furthermore, Janet informed police that Scott was a cocaine addict, and that there were "items of drug evidence" in the home.[57] Scott proceeded to inform the officers that he had taken the child to a neighbor's house out of fear that Janet would leave the country with him again, and further, that it was Janet and not he who abused drugs and alcohol.[58] After locating the Randolph's child at the neighbor's house, the responding officers confronted Scott and Janet, both of whom were present outside the house, and asked for consent to search the home.[59] Despite Scott's unwavering refusal to consent, Janet gave the officers her permission to enter the house.[60] Acting on Janet's consent, the officers entered and followed her into an upstairs bedroom where they found a straw caked with cocaine residue.[61]

B. Procedural History

Scott Randolph was subsequently indicted for possession of cocaine on the basis of the evidence obtained by the search.[62] Upon a motion to suppress the evidence, the trial court denied Scott's motion. Specifically, the trial court held that the search was valid because Janet had common authority over the home, and therefore had the power to consent to the search.[63] After losing at trial, Scott appealed the ruling to the Georgia Court of Appeals. In reviewing the decision, that court reversed the trial court's holding, and was later upheld by the state supreme court on the basis that "consent to conduct a warrantless search of a residence given by one occupant is not valid in the face of the refusal of another occupant

53. *Id.* at 1519.
54. *Id.*
55. *Id.*
56. *Id.*
57. *Id.*
58. *Id.*
59. *Id.*
60. *Id.*
61. *Id.*
62. *Id.*
63. *Id.* (referring to common authority as defined in *Matlock*).

who is physically present at the scene."[64] The Georgia Supreme Court acknowledged *United States v. Matlock*[65] in its decision, but distinguished it because Scott Randolph was physically present to refuse consent.[66] The U.S. Supreme Court granted certiorari to evaluate the decision. It affirmed, holding that "a warrantless search of a shared dwelling for evidence over the express refusal of consent by a physically present resident cannot be justified as reasonable as to him on the basis of consent given to the police by another resident."[67]

C. The U.S. Supreme Court's Decision

1. Majority Opinion

The majority opinion in *Georgia v. Randolph* holds "that a physically present inhabitant's express refusal of consent to a police search is dispositive as to him, regardless of the consent of a fellow occupant."[68] The majority rests its decision on three points: 1) the social expectations analysis originally outlined by Justice Harlan in his concurring opinion in *Katz v. United States,*[69] 2) United States Supreme Court precedents, and 3) the resolution of potentially troublesome implications of their holding.

a. "Social Expectation" Analysis

The majority's primary theory evaluates reasonableness as a function of social expectations between co-occupants.[70] To that end, the opinion presents a hypothetical situation in which someone arrives at a shared residence and encounters two co-occupants: one inviting entrance, and another refusing.[71] In such a situation, the majority assumes that, due to widely understood social expectations, "no sensible [visitor] would go inside."[72] The majority then analogizes the hypothetical guest to a police officer seeking permission to search, and concludes that because it would be unreasonable for the guest to enter facing such a disagreement, it would likewise be unreasonable for the officer to enter.[73] The majority's syllogistic logic leads it to conclude that "[d]isputed permission is thus no match for this [privacy] value of the Fourth Amend-

64. *Id.* (quoting Randolph v. State, 590 S.E.2d 834 (Ga. Ct. App. 2003)).
65. 415 U.S. 164 (1974).
66. Randolph v. State, 604 S.E.2d 835, 836-37 (Ga. 2004).
67. *Randolph*, 126 S. Ct. at 1526.
68. *Id.* at 1528.
69. 389 U.S. 347, 360-61 (1967) (Harlan, J., concurring).
70. *Randolph*, 125 S. Ct. at 1521 (citing *Matlock* for the proposition that "the reasonableness of such a search is in significant part a function of commonly held understanding about the authority that co-inhabitants may exercise in ways that affect each other's interests.").
71. *Id.* at 1522-23.
72. *Id.* at 1523.
73. *Id.*

ment, and the State's other countervailing claims do not add up to outweigh it."[74]

b. Precedent

Additionally, the majority considers a prior case that "took a step toward the issue" of warrantless searches.[75] In *Minnesota v. Olson*,[76] the Court assessed the privacy rights of an overnight social guest arrested in the course of a warrantless search, and held that houseguests have privacy rights parallel to those of occupants because "it is unlikely that [the host would] admit someone who want[ed] to see or meet with the guest over the objection of the guest."[77] Based on that decision, the majority goes on to reason that "it presumably should follow that an inhabitant of shared premises may claim at least as much, and it turns out that the co-inhabitant naturally has an even stronger claim."[78] Accordingly, the U.S. Supreme Court reinforced the Georgia Supreme Court's holding that *Randolph* is distinguishable from *Matlock*, and therefore not contradictory, due to the physical presence of the nonconsenting occupant.[79]

c. "Loose Ends"[80]

The majority concludes by addressing two potentially troublesome results, and then attempting to dispel them as inconsequential concerns that are not weighty enough to overpower its ultimate holding.[81] First, it tackles the apparent double standard created by comparing this holding with the Court's long-standing endorsement of citizens bringing criminal activity to light.[82] Ultimately, the majority reconciles the two by stating that "society can often have the benefit of [citizens bringing to light criminal activity of others] without relying on a theory of consent that ignores an inhabitant's refusal to allow a warrantless search."[83] Specifically, the majority reasons that the consenting co-occupant can either deliver the incriminating evidence to police,[84] or inform police of the relevant information to assist in obtaining a warrant.[85] Second, the majority addresses the dissent's primary objection that the holding will

74. *Id.* at 1524.
75. *Id.* at 1522 (referring to *Olson*, 495 U.S. 91 (1990)).
76. 495 U.S. 91 (1990).
77. *Randolph*, 126 S. Ct. at 1522 (quoting *Olson*, 495 U.S. at 99).
78. *Id.*
79. *See id.* at 1527.
80. *Id.* at 1524.
81. *Id.*
82. *Id.* at 1527; *see, e.g.*, *Coolidge*, 403 U.S. 443, 488 (1971) (stating that "it is no part of the policy underlying the Fourth . . . Amendment[ ] to discourage citizens from aiding to the utmost of their ability in the apprehension of criminals.").
83. *Randolph*, 126 S. Ct. at 1524.
84. *Id.* (citing *Coolidge*, 403 U.S. at 487-89 (referring to a situation where the defendant's wife turned incriminating guns over to the police)).
85. *Id.*

shield spousal abusers.[86] While recognizing the prevalent domestic violence problem,[87] the majority characterizes this criticism as a "red herring"[88] and argues that "this case has no bearing on the capacity of the police to protect domestic victims."[89] In support of this proposition, the majority asserts that police may always enter, under the exigent circumstances exception, when there is reasonable evidence that a domestic threat exists.[90]

2. Concurring Opinions

a. Justice Stevens

Justice Stevens writes a very brief concurring opinion which largely strays from addressing the issue of disputed consent searches. The concurrence, which appears to be directed at (and later responded to by) Justice Scalia, serves almost exclusively to criticize "originalism" as a tool for constitutional interpretation because it fails to account for changes in contemporary society.[91] Specifically, Stevens explains that when the Constitution was adopted in 1791 only a man's consent would be of import as a female could not at that time make legal decisions.[92] Accordingly, under an originalist's viewpoint, police officers seeking consent to search would only be required to receive permission from the husband. However, taking special care to highlight the discrepancy, Stevens points out that this concept is no longer valid because, "as a matter of constitutional law . . . the male and the female are equal partners."[93]

b. Justice Breyer

Justice Breyer also writes a concurring opinion, his purpose being to defend against the dissent's assertion that this holding will shield spousal abusers. Specifically, Breyer stresses the importance of reasonableness when determining the legality of Fourth Amendment searches, and further, that evaluation of reasonableness must turn on the totality of the circumstances.[94] Accordingly, Breyer notes that under circumstances in which police officers engage in a consent-based search with the sole objective being to recover evidence, and one occupant makes a clear objection, the search would be unreasonable absent any other Fourth Amend-

86. *Id.* at 1525.

87. *Id.* (citing numerous government and private statistical reports addressing the widespread domestic violence problem plaguing U.S. society).

88. *Id.* at 1526.

89. *Id.* at 1525.

90. *See id.* at 1526 (citing United States v. Donlin, 982 F.2d 31, 32 (1st Cir. 1992); United States v. Hendrix, 595 F.2d 883, 885-86 (D.C. Cir. 1979) (per curiam); People v. Sanders, 904 P.2d 1311, 1313-15 (Colo. 1995)).

91. *Id.* at 1528-29 (Stevens, J., concurring) (referring to all societal changes, including specifically the modern day equality of sexes).

92. *Id.*

93. *Id.* at 1529 (citing Reed v. Reed, 404 U.S. 71 (1971)).

94. *Id.* (Breyer, J., concurring).

ment exceptions.[95] However, Breyer goes on to remark that "were the circumstances to change significantly, so should the result."[96] In making this statement, Breyer implies that any sign of domestic violence or other dangerous conditions would alter the totality of the circumstances, and potentially make entrance reasonable under the exigent circumstances exception.

3. Dissenting Opinions

a. Chief Justice Roberts

In the dissenting opinion, Chief Justice Roberts, who is joined by Justice Scalia, begins by pointing to several deficiencies in the majority holding. Ultimately, they conclude by identifying what, in their opinion, is "the correct approach."[97]

First, the dissenters argue that the holding establishes a rule that will apply to various police searches in such a random fashion that it misses the true interest protected by the Fourth Amendment: a reasonable right to privacy.[98] In making this point, the dissent draws attention to the factual similarities present between *Randolph*, and those in the *Matlock*[99] and *Illinois v. Rodriguez*.[100] This comparison suggests that because *Randolph* implicitly contradicts the precedents, the holding is illogical.[101] In other words, the dissenters argue, a rule that endorses protection for defendants lucky enough to be standing at the threshold,[102] but not those sleeping in an adjacent room or detained in a nearby squad car[103] is arbitrary and unfair.[104] Specifically, the dissent challenges the decision to draw such a fine line between the nonconsenting occupant who is present at the threshold to refuse consent, and the nonconsenting occupant who happens to be away when the police arrive, but later contests the validity of the search.[105]

The second criticism, which the dissenters argue as being rooted in the randomness deficiency, suggests that the holding will shield spousal abusers by prohibiting police intervention when the perpetrator refuses to

95. *Id.* at 1530.
96. *Id.*
97. *Id.* at 1531 (Roberts, C.J., dissenting).
98. *See id.* at 1531, 1536, 1539.
99. 415 U.S. 164.
100. 497 U.S. 177 (1990).
101. *Randolph*, 126 S. Ct. at 1536 (Roberts, C.J., dissenting) (stating that "when the development of *Fourth Amendment* jurisprudence leads to such arbitrary lines, we take it as a signal that the rules need to be rethought").
102. *Id.*
103. *Id.* at 1534, 1536 (comparing the protection provided to the defendant in *Randolph* with the lack of protection afforded to the defendant in *Matlock*, who was detained in the police squad car; or the defendant in *Rodriguez*, who was napping in the next room).
104. *Id.* at 1536.
105. *See id.*

consent.[106] Although they acknowledge the majority's response,[107] Roberts and Scalia rebut the majority's presumption that domestic violence automatically gives rise to the exigent circumstances necessary to authorize warrantless entrance.[108] And even in circumstances where it does apply, justifying the rule on such a contingency is, according to the dissent, unreasonably "strange."[109]

Third, the dissenters argue that the social expectation analogy is plagued with unfounded assumptions and faulty logic for presupposing that any reasonable houseguest would automatically leave upon disagreement between roommates.[110] The dissenters attack this assumption with their own hypotheticals which suggest less clear "social expectations." For example, what if the guest was a family member; what if the guest had traveled long distances; or what if the guest was accepted by several, and only rejected by one roommate?[111] And further, the dissent questions, if the law recognizes no superiority or inferiority between co-occupants as the majority so clearly pronounces,[112] then why should the law implicitly assume that the occupant refusing consent prevails?[113]

Finally, the dissent asserts that the majority opinion incorrectly applies the social expectation analysis as originally articulated in *Katz*.[114] Specifically, the dissent argues that the majority is mistaken because it uses the analysis to review the issue of consent, rather than correctly

---

106. *Id.* at 1531 ("[T]he cost of affording such random protection is great, as demonstrated by the recurring cases in which abused spouses seek to authorize police entry into a home they share with a nonconsenting abuser."); *see also id.* at 1537 ("Perhaps the most serious consequence of the majority's rule is its operation in domestic abuse situations . . . .").

107. *Id.* at 1538 (referring to the majority response that the dissent's spousal abuse argument is a "red herring," and the ruling will not shield spousal abusers because of the continued prevalence of the exigent circumstances exception).

108. *Id.* (citing United States v. Davis, 290 F.3d 1239, 1240-41 (10th Cir. 2002) (finding no exigent circumstances justifying entry when police responded to a report of domestic abuse, officers heard no noise upon arrival, defendant told officers that his wife was out of town, and wife then appeared at the door seemingly unharmed but resisted husband's efforts to close the door)); *see also id.* at 1537 (inferring, as well, that in circumstances where domestic abuse is present but no exigent circumstances exist to allow police entry, the nonconsenter will "inflict retribution" on the consenter as soon as police leave).

109. *Id.*

110. *Id.* at 1532 (stating that "such shifting expectations are not a promising foundation on which to ground a constitutional rule, particularly because the majority has no support for its basic assumption – that an invited guest encountering two disagreeing co-occupants would flee – beyond a hunch about how people would typically act in an atypical situation").

111. *Id.*

112. *Id.* at 1523 (Souter, J.) (stating that "there is no societal understanding of superior and inferior" between co-occupants).

113. *Id.* at 1532 (Roberts, C.J., dissenting) (responding to the majority's presumption that a co-occupant has no authority to demand his guest be admitted by stating that "it seems equally accurate to say . . . that the objector has no 'authority' to insist on getting *his* way over his co-occupant's wish that her guest be admitted."); *see also id.* at 1541 (Scalia, J., dissenting) (stating that "men and women are no more 'equal' in the majority's regime, where both sexes can veto each other's consent, than on the dissent's view, where both sexes cannot").

114. *Id.* at 1532 (Roberts, C.J., dissenting) (citing *Katz*, 389 U.S. at 361 (Harlan, J., concurring) (referring to the two-part test which asked whether: 1) the suspect had a subjective expectation of privacy, and 2) society would view that expectation as reasonable)).

applying it to the overall inquiry of reasonableness in the expectation of privacy.[115] Moreover, as Chief Justice Roberts points out, "the social expectations concept has not been applied to all questions arising under the Fourth Amendment, least of all issues of consent."[116] Therefore, according to the dissenters, the majority utilizes the social expectation analysis inappropriately, because it confuses the contextual meaning of reasonableness. As a result, the dissent claims, social expectations should not be used to evaluate the Fourth Amendment's protection of privacy as it relates to the consent exception.[117]

The dissent ultimately offers a "correct approach"[118] to the issue presented in *Randolph*, based on the same assumption of the risk doctrine prevalent throughout the precedents.[119] Its rationale is very simple: "[i]f an individual shares information, papers, *or places* with another, he assumes the risk that the other person will in turn share access to that information or those papers *or places* with the government."[120] Accordingly, it is reasonable for police to search the premises upon the consent of any single occupant, because all other occupants impliedly assume that risk by living together.[121] The dissent argues that its approach is preferable because it flows naturally from the precedents,[122] and is logically grounded in the concept of privacy.[123]

b. Justice Scalia

In addition to joining Chief Justice Roberts, Justice Scalia offers an independent response to Justice Stevens's criticism of "originalism."[124] First, he challenges Stevens's assertion that "originalism" would have justified police in searching the home over the objection of a married woman when the Constitution was adopted in 1791.[125] He further defends "originalism" by arguing that this mode of constitutional interpretation can operate flexibly, despite Stevens's doubts. Specifically, Scalia states that "there is nothing new or surprising in the proposition" that the Constitution may remain unchanged while bodies of law to which it refers do change.[126] Scalia concludes by criticizing Stevens's "celebration"[127] of the majority's professed endorsement of women's rights, ar-

115. *Randolph*, 126 S. Ct. at 1537.
116. *Id.* (suggesting that the social expectations concept should not be used to examine the issue of whether consent was reasonable).
117. *Id.*
118. *Id.* at 1531.
119. *See, e.g.*, *Frazier*, 394 U.S. 731, 740 (1969).
120. *Randolph*, 126 S. Ct. at 1531 (Roberts, C.J., dissenting).
121. *Id.* at 1534 (citing *Matlock*, 415 U.S. at 171 n.7).
122. *Randolph*, 126 S. Ct. at 1536.
123. *Id.*
124. *Id.* at 1539-41 (Scalia, J., dissenting).
125. *Id.* at 1540.
126. *Id.* (arguing that "originalism" does not prohibit the Constitution from taking account of changes in other bodies of law).
127. *Id.* at 1541.

guing that in light of domestic violence patterns the decision will actually defeat the basis of their pronouncement, and instead lay the foundation for even greater violence toward women.[128]

c. Justice Thomas

Justice Thomas pursues a simple argument unaddressed in the other opinions. Specifically, he asserts that *Coolidge v. New Hampshire*[129] "squarely controls this case."[130] In *Coolidge*, officers questioned the wife of an absent defendant pursuant to a homicide investigation. Subsequently, and of her own accord, the wife invited the officers into the home and permitted them to conduct a search. By their search, the police obtained guns and clothing belonging to her husband. The evidence later served to convict him of murder.[131] Thomas relies upon the holding in *Coolidge* that "when a citizen leads police officers into a home shared with her spouse to show them evidence . . . that citizen is not acting as an agent of the police, and thus no Fourth Amendment search has occurred."[132] In doing so, Thomas argues that the facts in *Randolph* are indistinguishable from *Coolidge*.[133] Accordingly, Janet Randolph was not acting as an agent of the officers. For that reason, Justice Thomas explains, no Fourth Amendment search ever occurred. Therefore, the trial court's initial denial of the motion to suppress was appropriate.[134]

III. ANALYSIS

The Supreme Court's decision in *Georgia v. Randolph*[135] addressed whether police could lawfully search the home when one occupant disputes another occupant's consent.[136] Under existing law, lower courts had reached incongruous results in answering this question,[137] and the justice system as a whole needed an authoritative ruling from the Supreme Court. In *Randolph*, the Court answered by formulating a clear rule[138] which will help lower courts shape future decisions. Jurisdictional differences in holdings deciding this issue, albeit a limited one, will disappear and the justice system should benefit from a consistent

128. *Id.*
129. 403 U.S. 443 (1971).
130. *Randolph*, 126 S. Ct. at 1541 (Thomas, J., dissenting).
131. *Coolidge*, 403 U.S. at 446.
132. *Randolph*, 126 S. Ct. at 1541-42 (Thomas, J., dissenting) (citing *Coolidge*, 403 U.S. at 488-498).
133. *Randolph*, 126 S. Ct., at 1542.
134. *Id.* at 1542-43.
135. 126 S. Ct. 1515 (2006).
136. *Randolph*, 126 S. Ct. at 1518-19.
137. *Compare* Primus v. State, 813 N.E.2d 370, 376 (Ind. Ct. App. 2004) (holding that a search of Defendant's premises upon consent of his roommate was valid despite the Defendant's presence), *with* Randolph v. State, 604 S.E.2d 835, 837 (Ga. 2004) (holding that a search of Defendant's home upon consent of his wife, but his presence and clear refusal, was invalid).
138. *Randolph*, 126 S. Ct. at 1526 ("We therefore hold that a warrantless search of a shared dwelling for evidence over the express refusal of consent by a physically present resident cannot be justified as reasonable as to him on the basis of consent given to the police by another resident.").

and predictable approach to the validity of disputed-consent searches. To that extent, the holding in *Randolph* is clearly beneficial.

However, the *Randolph* holding simultaneously forecasts a novel set of problems that lower courts will have to address as a result. Generally, this analysis will to point out *Randolph's* potential deficiencies, balancing them with the decision's clear benefits for a comprehensive assessment. Attention will focus on unanticipated practical effects likely to arise in the aftermath of the holding, scrutinizing the tension between citizens' privacy rights, and the police's ability to enforce the law. Additionally, the Court's well-established trend of limiting Fourth Amendment privacy rights will be evaluated, paying heed to issues that may arise as a result of *Randolph's* departure from this trend. And finally, potential solutions available to both courts and police will be suggested to assist in dealing with these new issues.

A. Privacy Rights Versus Law Enforcement: Practical Implications of Georgia v. Randolph

1. Practical Implications for Citizens' Privacy Rights

At first glance, the *Randolph* decision appears to expand privacy rights, for the first time in several decades of Supreme Court Fourth Amendment decisions,[139] by holding in favor of the defendant on a voluntary consent issue.[140] The Court announces a straightforward holding that attempts to limit the scope of permissible police searches in favor of protecting the refusing defendant's privacy.[141] However, in spite of this seemingly simple directive, there is a strong possibility that *Randolph* will actually act to *limit* the right of privacy even further. This restriction of rights is most likely to manifest itself through changing police behavior in response to situations of disputed consent.

Specifically, *Randolph* creates an incentive for police, upon encountering a situation in which co-occupants might disagree on the issue of consent, to immediately detain and remove the occupant most likely to refuse a search. In taking this action, police can then pursue consent from the remaining occupant without regard to whether the detained occupant would agree.[142] Accordingly, the holding encourages police to take action that, from a practical standpoint, completely undermines the rule's original purpose of preserving a nonconsenting co-occupant's privacy rights in the face of another, consenting occupant.[143] This induce-

139. *See supra* Part I.

140. *Randolph*, 126 S. Ct. at 1528 ("This case invites the straightforward application of the rule that a physically present inhabitant's express refusal of consent to a police search is dispositive as to him, regardless of the consent of a fellow occupant.").

141. *Id.*

142. *See id.* at 1527 ("[T]he potential objector, nearby but not invited to take part in the threshold colloquy, loses out.").

143. *See id.* at 1528.

ment is further supported because the holding does not purport to overrule *United States v. Matlock*[144] or *Illinois v. Rodriguez*,[145] but merely distinguishes them. In those two cases, the Court held that neither an occupant sleeping in the next room,[146] nor one detained in a squad car,[147] could later object to a search. By refusing to overrule those two decisions, the Court essentially endorsed their holdings as good law, thereby informing lower courts and police that searches made under similar circumstances are still legally viable.[148] As a consequence, *Randolph* will likely encourage police to circumvent its holding by removing potential nonconsenting occupants, thereby nullifying the very Fourth Amendment protections it seeks to espouse.

A second implication of this decision concerns its limited effect on Fourth Amendment jurisprudence. As the majority opinion readily admits,[149] *Randolph* confines its authority to such a unique set of circumstances that it will very rarely apply.[150] It is interesting that the majority so readily admits the narrowness of its holding, because earlier in the opinion it justifies acceptance of this issue in the first place on the judicial system's need for a decision on this issue.[151] Nevertheless, this decision will be of only limited effect in protecting privacy rights because, coupled with police incentive to detain hostile suspects, it will be the rare situation in which lower courts and legal counsel can point to *Randolph* as authoritative, controlling authority. Instead of "drawing a fine line,"[152] as it did, the Court might have been better served by opting to rely on the broad and longstanding precedents through application of the assumption of the risk doctrine,[153] as it concerns co-occupancy. Though reliance on this doctrine would have forced the Court to come to the opposite conclusion,[154] doing so would have enabled the Court to avoid the

144. 415 U.S. 164 (1974).

145. 497 U.S. 177 (1990).

146. *Rodriguez*, 497 U.S. at 180.

147. *Matlock*, 415 U.S. at 179 (Douglas, J., dissenting).

148. *Randolph*, 126 S. Ct. at 1527.

149. *See id.* ("[I]f [*Rodriguez* and *Matlock*] are not to be undercut by today's holding, we have to admit that we are drawing a fine line; if a potential defendant with self-interest in objecting is in fact at the door and objects, the co-tenant's permission does not suffice for a reasonable search, whereas the potential objector, nearby but not invited to take part in the threshold colloquy, loses out.").

150. *See id.*

151. *Compare Randolph*, 126 S. Ct. at 1527 ("*[W]e have to admit that we are drawing a fine line*; if a potential defendant with self-interest in objecting is in fact at the door and objects, the co-tenant's permission does not suffice for a reasonable search, whereas the potential objector, nearby but not invited to take part in the threshold colloquy, loses out.") (emphasis added), *with Randolph*, 126 S. Ct. at 1520 ("We granted certiorari to resolve a split of authority on whether one occupant may give law enforcement effective consent to search shared premises, as against a co-tenant who is present and states a refusal to permit the search.").

152. *Randolph*, 126 S. Ct. at 1527.

153. *See, e.g.*, *Frazier*, 394 U.S. 731, 740 (1969).

154. *Randolph*, 126 S. Ct. at 1531 (Roberts, C.J., dissenting) (suggesting that if the Court had relied on the assumption of the risk doctrine, it could only have come to the conclusion that the search against Mr. Randolph did not violate the Fourth Amendment).

implicit contradiction with the relevant precedents,[155] while simultaneously reinforcing the vitality of a consistent line of decisions dealing with Fourth Amendment consent searches.[156]

2. Practical Implications for Law Enforcement

Because all Fourth Amendment decisions implicitly struggle with the delicate balance between citizens' privacy rights and law enforcement capabilities, *Randolph* equally carries with it practical implications for police officers in the course of their work. Although it is difficult to predict the full extent of changing police behavior following this decision, both the majority and dissenting opinions reflect on one recurring situation this holding is almost sure to affect on a regular basis: domestic violence. Because domestic violence already plagues U.S. society,[157] and because this rule undisputedly affects police response to domestic violence situations,[158] the opposing opinions devote considerable attention to the rule's likely effect on the issue.[159] Not surprisingly, they reach wholly inconsistent conclusions.

While the majority presents a solid argument that the dissent's criticism is a "red herring,"[160] its opinion fails to look beyond the legal theory in order to appreciate the reality of domestic violence. Moreover, evidence of domestic violence may permit police entry under the exigent circumstances exception[161] in some disputed-consent circumstances; however, often times no such evidence will exist to, literally, open the door to a search. It is easy to foresee a situation in which police respond to a domestic violence call only to find that no actual violence has yet been inflicted. Police are powerless to act at that point. However, as recognized by the dissent, consider the retribution that will be carried out once the police leave.[162] Similarly, it is likely that police will rarely get an opportunity to view evidence of violence in the first place, as the bat-

155. *E.g.*, *Matlock*, 415 U.S. 164; *Rodriguez*, 497 U.S. 177, 188-89 (1990) (referring to contradiction created when comparing the decision in *Randolph* which disallows searches made when the suspect is standing at the door, to the decisions in *Matlock* and *Rodriguez* which allow searches made when the suspect is either detained in the nearby squad car, or sleeping in an adjacent room).

156. *See supra* Part I.

157. *See* Kapila Juthani, Note, *Police Treatment of Domestic Violence and Sexual Abuse: Affirmative Duty to Protect vs. Fourth Amendment Privacy*, 59 N.Y.U. ANN. SURV. AM. L. 51, 56 (2003). *See generally Randolph*, 126 S. Ct. at 1525 (presenting a substantial list of government and private reports analyzing the extent and nature of the domestic violence problem in the U.S.).

158. *See, e.g.*, *Randolph*, 126 S. Ct. 1515; *see also* Scheiber v. City of Philadelphia, 156 F. Supp. 2d 451, 458-59 (E.D. Pa. 2001) (referring to two, of many, relevant Fourth Amendment consent cases where the original police response arose from domestic violence).

159. *See Randolph*, 126 S. Ct. at 1525-26 (explaining the majority's perspective of the holding's affect on domestic violence); *see also id.* at 1537-38 (Roberts, C.J., dissenting).

160. *Randolph*, 126 S. Ct. at 1526 (referring to its failure to account for the exigent circumstances exception).

161. *Id.* at 1525 n.6 (listing the exigent circumstances that render consent irrelevant).

162. *See id.* at 1537 (Roberts, C.J., dissenting) (arguing that violence will be exacerbated when domestic disputes are reported, but police are unable to take action due to the lack of evidence which would permit intervention).

tered and frightened woman will remain out of view for fear of severe future consequences if she comes out and pleads for protection.

On a broader level, *Randolph* will also have the practical effect of contributing to police officers' general confusion concerning a suspect's privacy rights. Responding officers will constantly confront unclear situations where there is not an obvious consenting or nonconsenting occupant standing at the doorway. Does the refusal still trump if the refusing occupant is shouting his rebuff up the stairs from the basement? What if either occupant is under the influence of alcohol or drugs? What if the nonconsenting occupant is on the telephone with the consenting occupant, stating his refusal through the phone? Prior to *Randolph*, these potential contingencies did not present an issue. Police officers, operating under the established assumption of the risk principle,[163] needed only to concern themselves with obtaining consent from the co-occupant standing in the doorway. Thus, although these possibilities were irrelevant prior to *Randolph*, drawing such a fine line makes these questions suddenly important.

B. A Departure from the Court's Trend of Limiting Privacy Rights

Beginning with the decision in *Frazier v. Cupp*,[164] and extending over thirty years up to *Randolph*, the Supreme Court adhered to a generally recognized trend of expanding the validity of warrantless, consent-based police searches, and correspondingly limiting defendants' privacy rights under the Fourth Amendment.[165] Consequently, *Randolph*, which reinforces the individual privacy interest at the expense of valid police searches,[166] comes as an unexpected departure from this trend.[167] While the consequences of departing from an identifiable trend in constitutional interpretation may not necessarily be severe, the resulting judicial uncertainty creates problematic issues.

Prior to *Randolph*, lower courts across the nation benefited from a clear understanding of the Supreme Court's consistent trend of upholding the validity of police searches. However, post-*Randolph*, courts are suddenly placed into a state of confusion about the direction of Fourth Amendment search and seizure jurisprudence. Specifically, state courts will struggle to decide whether this decision forecasts an impending reversal in Fourth Amendment decision-making, or whether it's merely a detour in the continued path of generally limiting privacy. This creates a

163. *Supra* Part I.B.

164. 394 U.S. 731 (1969).

165. *See generally* Kloster, *supra* note 26, at 104 (providing a detailed analysis of the Court's trend in deciding Fourth Amendment search cases).

166. *Randolph*, 126 S. Ct. at 1526.

167. *See supra* Part I (referring to the trend of limiting privacy rights and expanding the boundaries of valid warrantless police searches).

negative impact on the judicial system as a whole because it signals inconsistency and unpredictability.

C. Unreasonable Reasoning

The logic employed by the majority opinion can be questioned with regard to two specific facets of the decision. First, a recurrently problematic result of the holding concerns its apparent neglect of precedent. Moreover, the majority attempts to deal with this problem by "drawing a fine line"[168] and dismissing the contradiction with precedent cases as a "loose end"[169] to be tied up.[170] However, the reality of the holding, especially concerning *Matlock* and *Rodriguez*, is that the decisions considered collectively bear little logical relationship to one another. Worse, the logic employed by the Court in each decision serves to effectively contradict the logic used in the others.[171] Consideration of the holdings together begs the question of how justice may be reconciled when a nonconsenting occupant lucky enough to be standing at the threshold prevails, but a nonconsenting occupant napping in the next room, or detained in a nearby squad car, loses out. Factually, the differences are so minute that the majority struggles to distinguish them.[172] Consequently, following *Randolph* a stalemate of Supreme Court decisions pits limited privacy rights against expansive police discretion to search, and the determinative factor as to which rule will apply falls solely upon the arbitrary circumstances surrounding the suspect's location with respect to the doorway.

Additionally, the "social expectation" reasoning utilized by the majority as a tool to assess reasonableness is questionable for two reasons. Initially, the model is faulty because it misses the true inquiry of reasonableness as defined by precedent, which is the individual's "legitimate expectation of privacy."[173] The majority mistakenly equates this legitimate expectation with the disputed-consent social expectation hypothetical,[174] assuming without explanation that it leads to the correct conclusion that entry violates reasonableness. However, as the dissent points out, social expectations comprise a very different set of assumptions than those encompassing the legitimate expectation of privacy.[175] Accordingly, one might assume that his roommate won't turn over evidence of

---

168. *Randolph*, 126 S. Ct. at 1527.

169. *Id.*

170. *See supra* Part II.C.1.c.

171. *Randolph*, 126 S. Ct. at 1527 (referring to the *Randolph* holding as compared to *Matlock* and *Rodriguez*).

172. *Id.* (admitting that they are drawing a fine line, and arguing that the "formalism is justified").

173. Rakas v. Illinois, 439 U.S. 128, 143 n.12 (1978).

174. *See Randolph*, 126 S. Ct. at 1522-23 (referring to the hypothetical situation laid out in the majority opinion, where a social guest arrives at a residence and is greeted by two co-occupants in disagreement over whether the guest may enter).

175. *Id.* at 1533.

his drug problem to police, based on social expectations between them.[176] However, because he shares a living space with that roommate, any legitimate expectation of privacy with regard to their shared dwelling has "already been frustrated,"[177] and regardless of friendship-based expectations between them, there can be no legitimate expectation of protection from police searches if consent is given by one roommate. Thus, the majority's reliance on the social expectation hypothetical could be considered inappropriate because it equates the social expectations of co-occupants with the legitimate expectations of privacy. The concepts are different, and accordingly yield different results.[178]

The second problem that arises with regard to the hypothetical is the majority's unwarranted assumption that the nonconsenting occupant's decision prevails in the event of a disputed-consent situation.[179] Moreover, the majority opinion clearly holds that whenever two co-occupants disagree on the issue of consent to a police search, the nonconsenting occupant prevails. This assumption generates query as to why, when the occupants share an equal status of authority,[180] the party refusing consent must be considered the winner.[181] Thus, the social expectation analysis could be considered imperfect even within its own boundaries, because it stresses the fact that co-occupants share equal authority, but then assumes without justification that the nonconsenting occupant prevails in situations of dispute.

D. Solutions

Despite the unanticipated problems likely to arise following the Supreme Court's holding in *Randolph*, some possible solutions exist that may assist to alleviate any negative affects. First and foremost, judges deciding Fourth Amendment disputed-consent search cases should engage in an intensive fact-based analysis of the situation in front of them, and require the particular situation to precisely match *Randolph*'s distinct factual background before applying it as precedent. Although *Randolph* seems to inherently conflict with prior precedents such as *Matlock* and *Rodriguez*, if courts can devote special attention to the factual background so as to apply the respective precedents only to identical factual

176. *Id.*

177. United States v. Jacobsen, 466 U.S. 109, 117 (1984) (referring to the meaning of a legitimate expectation of privacy in the context of co-occupancy).

178. *Id.*

179. *Randolph*, 126 S. Ct. at 1523 (stating that "there is no societal understanding of superior and inferior" between co-occupants).

180. *Id.* at 1523 ("[T]here is no common understanding that one co-tenant generally has a right or authority to prevail over the express wishes of another, whether the issue is the color of the curtains or invitations to outsiders.").

181. *Id.* at 1532 (Roberts, C.J., Scalia, J., dissenting) ("It seems equally accurate to say – based on the majority's conclusion that one does not have a right to prevail over the express wishes of his co-occupant – that the objector has no 'authority' to insist on getting *his* way over his co-occupant's wishes that her guest be admitted." (emphasis added)).

situations, no overbearing contradictions will arise. Doing this will alleviate the confusion created by the Supreme Court's seemingly conflicting decisions by limiting the scope of future holdings to situations which match the appropriate precedents.

Second, courts should embrace a Preliminary Effort Test,[182] which requires officers pursuing a co-occupant consent search to, at least, make an initial attempt to obtain consent from all accessible inhabitants. Specifically, this test would initially require responding officers engaged with any occupant of the residence to inquire as to whether any other co-occupants are present on the premises. If the officers are led to believe that other occupants are at home, this test would additionally require the police to obtain consent from all additional occupants before entering to search. Adoption of this test would be beneficial for several reasons. First, the Preliminary Effort Test is commensurate with the reasonableness inquiry explicit in the Fourth Amendment,[183] because it requires police to take the simple step of initially asking whether other occupants are present, and then obtaining consent from any that are there. Similarly, this test reinforces the enduring vitality of the warrant requirement in the Fourth Amendment[184] because police would be required to pursue the alternate course of obtaining a warrant in the event that a roused co-occupant decides to refuse the search. Last, application of this test would benefit the judicial system by preventing officers from circumventing the purpose of *Randolph's* holding by immediately detaining potential non-consenting occupants.[185] Incongruous situations similar to *Matlock* or *Rodriguez* would not arise to frustrate the holding, because police would be required to make an initial effort to seek out other inhabitants, and receive consent from them before engaging in a search. Although this test creates an added burden on police, it would greatly reduce the possibility of engaging in an illegal search.

Finally, in the event that police are unable to obtain a warrant, they should be encouraged to pursue the traditional alternatives to a warrantless search more readily. For one, officers responding to situations of disputed consent should simply request that the consenting occupant retrieve and hand over any incriminating evidence that the officers hope to discover in the event of a search. This alternative is likely to be viable in many disputed consent circumstances where the two occupants are at odds with one another,[186] as the consenting individual will be willing to

182. *See* Elizabeth A. Wright, Note, *Third Party Consent Searches and the Fourth Amendment: Refusal, Consent, and Reasonableness*, 62 WASH. & LEE L. REV. 1841, 1874-77 (2005) (proposing the "Reasonable Attempt Test"); Gregory S. Fisher, *Search and Seizure, Third Party Consent: Rethinking Police Conduct and the Fourth Amendment*, 66 WASH. L. REV. 189, 202-08 (1991) (proposing "Discretionary Restraint" as a framework for addressing third-party consent).

183. U.S. CONST. amend. IV.

184. *Id.*

185. *See supra* Part III.A.2

186. *See, e.g.*, *Randolph*, 126 S. Ct. at 1519.

take whatever steps are necessary to implicate their co-occupant. Additionally, when the consenting occupant is unable or unwilling to deliver the incriminating evidence to the responding officers, they can at least aid the officers in obtaining a legitimate search warrant by establishing probable cause. In choosing to pursue either of these warrantless search alternatives, police will benefit by preserving the ability to obtain incriminating evidence against the suspect, while at the same time eliminating the possibility of engaging in a warrantless search.

CONCLUSION

In the immediate wake of *Georgia v. Randolph*, the decision's long-term impact on Fourth Amendment search and seizure jurisprudence is unclear. While the holding may signal a pending reversal in the Court's prior predilection to limiting citizens' privacy rights, it may also represent only a slight divergence in a continually viable trend. Perhaps the Court will have the opportunity to resolve this query in future decisions, potentially even those based on *Randolph's* shortcomings. Regardless of its eventual impact, *Randolph* currently holds a position of limited application which is likely to create confusion and foster practical dilemmas. As *Randolph* begins to assert its presence in the world of Fourth Amendment privacy rights, one can only hope that, as the majority opinion suggests, these problematic issues prove to be mere "loose ends"[187] capable of being tied up.

Andrew Fiske[*]

187. *Id.* at 1527.

*. J.D. Candidate, 2008, University of Denver Sturm College of Law. The author would like to thank Professor Alan Chen for his valuable input.

BROWN V. SANDERS: INVALID FACTORS AND APPELLATE REVIEW IN CAPITAL SENTENCING

INTRODUCTION

Death penalty jurisprudence in America is dynamic. Since *Furman v. Georgia*[1] in 1972, the states and the United States Supreme Court have elaborated constitutional and practicable systems of capital punishment and sentencing.[2] The Court has worked to ensure that the death penalty cannot be imposed arbitrarily, and to allow sentencers to review mitigating factors that can support lesser sentences.[3] Meanwhile, state appellate courts have examined death penalty statutes to ensure they meet revised sentencing guidelines.[4] The Supreme Court has carved out a jurisprudential approach to sentences rendered using invalid sentencing factors after new statutory factors were found too vague to ensure the constitutional rights of offenders.[5] The Court has maintained guidelines for valid factors[6] and has addressed cases in which sentences were imposed after a jury considered factors later determined invalid.[7]

This comment addresses a recent capital decision by the United States Supreme Court. *Brown v. Sanders*[8] is the latest in a series of cases addressing death sentences issued after the consideration of invalid sentencing factors. In deciding *Brown*, the Supreme Court eliminated the distinction between "weighing" and "non-weighing" jurisdictions[9] which had been in place for fifteen years.[10] While this change in jurisprudence will simplify the examination of sentences derived from invalid sentencing factors, the majority opinion in *Brown* failed to clarify the role of appellate review under the new system.

Part I provides an overview of the constitutional requirements for death sentences, and the approaches taken with respect to invalid sentencing factors. Part II discusses the decision in *Brown*. Part III analyzes the decision, first in a discussion of its elimination of weighing and

---

1. 408 U.S. 238 (1972) (per curiam).
2. Srikanth Srinivasan, Note, *Capital Sentencing Doctrine and the Weighing – Nonweighing Distinction*, 47 STAN. L. REV. 1347, 1347-48 (1995).
3. Stephen Hornbuckle, Note, *Capital Sentencing Procedure: A Lethal Oddity in the Supreme Court's Case Law*, 73 TEX. L. REV. 441, 444-46; Srinivasan, *supra* note 2, at 1352-53.
4. *See, e.g.*, Brown v. Sanders, 126 S.Ct. 884, 888-89 (2006).
5. Tuilaepa v. California, 512 U.S. 967, 973-75 (1994).
6. *Id.* at 972-73.
7. *See, e.g.*, *Brown*, 126 S. Ct. at 888.
8. 126 S. Ct. 884 (2006).
9. *Id.* at 891-92.
10. *See generally* Clemons v. Mississippi, 494 U.S. 738 (1990) (holding that in death penalty cases it is constitutionally permissible for courts to weigh or reweigh aggravating or mitigating circumstances).

non-weighing jurisdictions, and then in regard to the risks to future cases regarding the requirement for appellate review which is strikingly absent from the opinion. In eliminating the distinction, the Court attempted to clarify the sentencing process for all jurisdictions, but neglected to discuss the crucial role of appellate review under the new system.

I. BACKGROUND

Three requirements apply to all death sentences: guided discretion, individualized sentencing, and appellate review.[11] When a sentencing factor used to meet either of the first two requirements is found invalid, appellate courts have previously determined whether the state is a weighing or a non-weighing jurisdiction to decide whether the sentence may stand.[12] Appellate review is always a requirement, but any sentence rendered after consideration of invalid sentencing factors may be "reweighed" or may go through harmless error analysis during the appellate process.[13]

A. Guided Discretion

Furman v. Georgia[14] established that a sentence of death is unconstitutional if a sentencing body had complete discretion in imposing it.[15] The Supreme Court explained that "where discretion is afforded a sentencing body on a matter so grave as the determination of whether a human life should be taken or spared, that discretion must be suitably directed and limited so as to minimize the risk of wholly arbitrary and capricious action."[16] Statutes developed in response to *Furman* were upheld if they limited the group of offenders eligible for death, thus guiding the discretion of the sentencing bodies.[17] Most jurisdictions now meet this requirement by defining eligibility factors for the death penalty.[18] If the nature of a crime satisfies the factors required by the state, the sentencer has the opportunity to impose the death sentence.[19]

---

11. *See, e.g.*, Gregg v. Georgia, 428 U.S. 153, 189, 193 (1976) (guided discretion); Lockett v. Ohio, 438 U.S. 586, 605 (1978) (individualized discretion); Zant v. Stephens, 462 U.S. 862, 890 (1983) (appellate review).

12. *E.g.*, *Brown v. Sanders*, 126 S. Ct. 884, 889 (2006).

13. *See* Clemons v. Mississippi, 494 U.S. 738, 750-54 (1990).

14. 408 U.S. 238 (1972) (per curiam).

15. Hornbuckle, *supra* note 3, at 441-42; Srinivasan, *supra* note 2, at 1349.

16. *Gregg*, 428 U.S. at 189.

17. *Zant*, 462 U.S. at 874, 879; Srinivasan, *supra* note 2, at 1349-50; Marcia A. Widder, *Hanging Life in the Balance: The Supreme Court and the Metaphor of Weighing in the Penalty Phase of the Capital Trial*, 68 TUL. L. REV. 1341, 1347-48 (1994).

18. *See, e.g.*, *Brown*, 126 S. Ct. at 889; *see also* Srinivasan, *supra* note 2, at 1351 (suggesting that all guided discretion statutes require a showing of certain aggravated factors before imposing a death sentence); Tuilaepa v. California, 512 U.S. 967, 971-72 (1994) (explaining that a defendant cannot be sentenced to death without a finding of at least one "aggravating circumstance," which serves to limit the number of defendants eligible for the death penalty).

19. *Brown*, 126 S. Ct. at 889; *see also* Hornbuckle, *supra* note 3, at 446 (explaining that a factfinder can sentence a defendant to death only where at least one aggravating circumstance has been proven).

B. Individualized Sentencing

The second requirement for a constitutional death sentence is that the jury analyzes any mitigating factors in the circumstances of the crime or the character of the defendant.[20] This provides an individualized sentence for every offender eligible for death.[21] Because of the severity and finality of the death penalty, the Constitution requires that even if an offender is found eligible through guided discretion, his character and circumstances must be weighed against the aggravating factors found by the sentencer.[22] As opposed to their limited discretion in determining eligibility for death, the individualized sentencing requirement ensures that juries may consider any mitigating evidence that comes to light.[23] Though sentencers may impose the death penalty on eligible offenders, they may always consider mitigating factors, and they are never required to issue a sentence of death.[24]

C. Appellate Review

After *Furman*, appellate review is a safeguard that has ensured the constitutionality of death sentences.[25] *Gregg v. Georgia*[26] emphasized that in a system of guided discretion "the further safeguard of meaningful appellate review is available to ensure that death sentences are not imposed capriciously or in a freakish manner."[27] The Court in *Zant v. Stephens*[28] noted that Georgia's sentencing procedure could be approved in part because every death sentence was reviewed by the state supreme court "to determine whether the sentence was arbitrary or disproportionate."[29] In the context of invalid sentencing factors, appellate review becomes even more important to prevent an unconstitutional sentence.

Barclay v. Florida[30] presents Florida's approach to appellate review of death sentences. As in many other jurisdictions, there is an automatic appellate review of any death penalty case by the state supreme court.[31] If a jury used an invalid sentencing factor to determine eligibility for a

20. Barclay v. Florida, 463 U.S. 939, 950 (1983); Eddings v. Oklahoma, 455 U.S. 104, 112 (1982).

21. *See Barclay*, 463 U.S. at 950; *Zant*, 462 U.S. at 879.

22. *Barclay*, 463 U.S. at 950. Though not declared a constitutional requirement in the cases immediately following *Furman v. Georgia*, the court recognized prior to its decision in *Eddings v. Oklahoma*, that any jury's sentencing procedure involved a balance of the aggravating and mitigating factors of the case. Widder, *supra* note 17, at 1358.

23. Srinivasan, *supra* note 2, at 1353.

24. *See, e.g.*, Woodson v. North Carolina, 428 U.S. 280, 304-05 (1976) (striking down a mandatory death penalty statute).

25. *See Gregg*, 428 U.S. at 195.

26. 428 U.S. 153 (1976).

27. *Id.*

28. 462 U.S. 862 (1983).

29. *Zant*, 462 U.S. at 876.

30. 463 U.S. 939 (1983) (upholding Florida's death penalty statute where the state supreme court reviewed each death sentence).

31. *Barclay*, 463 U.S. at 953.

death sentence, the United States Supreme Court requires either a reweighing of all the factors or a harmless-error review if no mitigating factors are present.[32] However, within this review it is accepted that the sentencing process involves subjective decisions.[33] The subjectivity of a sentencer's decisions is appropriate after the class of defendants eligible for the death penalty has been narrowed. After a defendant is found to be eligible, the sentencer uses discretion to impose a sentence appropriate to the offense and any mitigating factors.[34]

When an invalid factor is used in sentencing, appellate review or harmless-error analysis is required by the Eighth Amendment.[35] *Barclay* requires that when an invalid factor has been used in a death penalty decision, there be a reweighing of the factors leading to the sentence by either a jury or an appellate court.[36] When an appellate court affirms a death sentence, there must be no "automatic assumption that [an invalid] factor has not infected the weighing process."[37] *Barclay* made it clear that appellate review is a constitutional requirement for any death sentence involving an invalid sentencing factor.[38]

The Supreme Court has consistently recognized this precedent in prior and later cases. In *Zant*, the Court described "mandatory appellate review" of a sentence imposed using an invalid sentencing factor as an "important procedural safeguard" that was necessary to "avoid arbitrariness and assure proportionality" in sentencing.[39] When there is a risk of guided discretion going astray, appellate review is the safety measure that ensures the constitutionality of a death sentence. *Clemons v. Mississippi*[40] also discussed the importance of "meaningful appellate review" in cases involving invalid sentencing factors.[41] Though *Clemons* is most often cited as an example of a weighing state, the decision rested on the Mississippi Supreme Court's failure to analyze whether the use of an invalid sentencing factor was a harmless error.[42]

Stringer v. Black[43] reaffirmed the importance of appellate review in cases with sentencing errors.[44] The case framed the harmless-error re-

32. *See id.* at 954-58. Additionally, if a judge imposes a death sentence over the jury's recommendation, the state supreme court applies a clear and convincing standard to all the facts in favor of death to determine if the sentence should stand. *Id.* at 955-56.
33. *Id.* at 950. *See* Widder, *supra* note 17, at 1373.
34. *See Brown*, 126 S. Ct. at 889; Widder, *supra* note 17, at 1374.
35. *See Brown*, 126 S.Ct. at 901 (Breyer, J., dissenting).
36. *See Clemons*, 494 U.S. at 749, 751.
37. Stringer v. Black, 503 U.S. 222, 231 (1992).
38. *See Barclay*, 463 U.S. at 958.
39. *Zant*, 462 U.S. at 890.
40. 494 U.S. 738 (1990).
41. *Clemons*, 494 U.S. at 749.
42. *Id.* at 753-54; *see also* Hornbuckle, *supra* note 3, at 453.
43. 503 U.S. 222 (1992).
44. *Stringer*, 503 U.S. at 237. *But see* Srinivasan, *supra* note 2, at 1367 (arguing that *Stringer* added harmless-error analysis as a new element to individualized sentencing).

view requirement in respect to the distinction between weighing and non-weighing states, which it defined.[45] *Stringer* indicated that where an appellate court has determined that there would have been no difference in a sentence without the analysis of an invalid factor, the sentence is constitutional.[46] But the reviewing court may not assume such, and harmless-error review or appellate reweighing of sentencing factors is necessary to ensure that an offender has been sentenced individually.[47] Nonetheless, whether the state weighs or does not, an appellate review of an invalid sentencing factor is a constitutional requirement for a death sentence to stand.

D. Non-Weighing States

In *Zant v. Stephens*,[48] the Supreme Court addressed a sentence that had been issued according to Georgia's statutes concerning guided discretion and individualized sentencing. On appeal, one of the factors making the defendant eligible for the death penalty was found to be unconstitutionally vague.[49] The Court had to determine whether the use of the invalid factor in determining eligibility required the sentence to be vacated.[50]

In *Zant*, the aggravating (eligibility) factors were used to narrow the class of offenders eligible for death, but the jury was not required to specifically analyze those factors in imposing a sentence.[51] *Zant* is now considered to be an analysis of a non-weighing state because of the jury's ability to consider non-statutory factors in sentencing.[52] Georgia's statute provided that at least one statutory eligibility factor must be found by a jury for a defendant to become eligible for the death penalty.[53] However, once a jury found the existence of one of the eligibility factors beyond a reasonable doubt, it could examine any other evidence from the trial proceeding and any mitigating circumstances to determine the final sentence.[54]

In *Zant*, though one factor used in determining eligibility for the death penalty was found to be invalid, the Court determined that the defendant was still eligible for his sentence, based on the valid eligibility

45. *Stringer*, 503 U.S. at 229. Courts have used the term "weighing" in regard to sentencing for some time. *See Zant*, 462 U.S. at 880. Justice Breyer argues in his dissent to *Brown v. Sanders* that *Stringer v. Black* was the first case to codify the distinction between weighing and non-weighing in jurisdictional approaches to eligibility factors as aggravating factors. *Brown*, 126 S. Ct. at 902 (Breyer, J., dissenting).

46. *Stringer*, 503 U.S. at 232.

47. *Id.*; Widder, *supra* note 17, at 1344.

48. 462 U.S. 862 (1983).

49. *Zant*, 462 U.S. at 867.

50. *Id.* at 864.

51. *Id.* at 879-81; Hornbuckle, *supra* note 3, at 447-48.

52. Hornbuckle, *supra* note 3, at 447-48.

53. *Zant*, 462 U.S. at 871-72.

54. *Id.*

factors that were found by the jury.[55] The purpose of the statutory factors in Georgia was primarily to guide the discretion of the jury in finding defendants eligible for the death penalty.[56] After the jury placed a defendant beyond that barrier, it had the liberty to base its sentence on all the evidence before it.[57] The Supreme Court held that the existence of at least one valid eligibility factor was sufficient to make the defendant eligible for the death penalty under Georgia law.[58] Although the categorizing of some evidence as an "aggravating circumstance" (as eligibility factors are termed by Georgia statute) "might have caused the jury to give somewhat greater weight to respondent's prior criminal record than it otherwise would have given," the Court did not find this to be a Constitutional error, as the jury properly had all available evidence before it in determining the sentence.[59] Because the invalid factor was not specifically a part of the sentencing process (i.e., because the invalid factor was not given any specific "weight"), the Court upheld the sentence.[60]

The Court limited its holding to states with statutory schemes similar to Georgia's.[61] The opinion distinguished the circumstances in *Zant* from a possible case in which a sentencer would be "specifically instructed to weigh statutory aggravating and mitigating circumstances in exercising its discretion whether to impose the death penalty."[62] It was just such a case which elicited the next refinement in death penalty jurisprudence.

E. Weighing States

Seven years after it decided *Zant v. Stephens*, the Supreme Court addressed death sentences imposed with an invalid factor in a jurisdiction using the same set of factors for determining eligibility (guided discretion) and for imposing sentences (individualized sentencing).[63] *Clemons v. Mississippi* distinguished "weighing" jurisdictions from those following *Zant's* model.[64] Mississippi's statute, unlike Georgia's, used a set of statutory aggravating circumstances both to determine eligibility for the death penalty and to determine whether the death penalty was warranted.[65] Rather than using any evidence before it to determine the sentence, Mississippi juries were required to "weigh" specific statutory factors against any mitigating circumstances, also outlined in statute.[66] The

55. *Id.* at 884, 890.
56. *Id.* at 875.
57. *Id.* at 872.
58. *Id.*
59. *Id.* at 888-89.
60. *Id.*; *see also Clemons*, 494 U.S. at 744-45; Hornbuckle, *supra* note 3, at 450-51.
61. *Zant*, 462 U.S. at 890.
62. *Id.*
63. *See* Widder, *supra* note 17, at 1352-53.
64. *Clemons*, 494 U.S. at 744-45; Hornbuckle, *supra* note 3, at 448-49.
65. *Clemons*, 494 U.S. at 745.
66. *Id.*

jury's task was to determine if there were "insufficient mitigating circumstances . . . to outweigh the aggravating circumstances."[67]

The Court found a risk of "skewing" sentences in these jurisdictions if a jury was instructed to weigh an invalid factor during the sentencing process.[68] It held that such an error did not necessarily invalidate the death sentence, but that "meaningful appellate review" or a "reweighing" of aggravating and mitigating circumstances was required to preserve a sentence issued after weighing of an invalid factor.[69] If a new jury or an appellate court determined that the error was harmless, the sentence could be upheld.[70] Because it was not clear in *Clemons* that the appropriate reweighing or review had been performed by the appellate courts, the death sentence was vacated.[71]

The distinction between weighing and non-weighing jurisdictions was crystallized in *Stringer v. Black*. In reviewing another Mississippi case, the court clarified the procedures used under Georgia (non-weighing) and Mississippi (weighing) law:[72]

> In a nonweighing State, so long as the sentencing body finds at least one valid aggravating factor, the fact that it also finds an invalid aggravating factor does not infect the formal process of deciding whether death is an appropriate penalty. Assuming a determination by the state appellate court that the invalid factor would not have made a difference to the jury's determination, there is no constitutional violation resulting from the introduction of the invalid factor in an earlier stage of the proceedings. But when the sentencing body is told to weigh an invalid factor in its decision, a reviewing court may not assume it would have made no difference if the thumb had been removed from death's side of the scale. When the weighing process itself has been skewed, only constitutional harmless-error analysis or reweighing at the trial or appellate level suffices to guarantee that the defendant received an individualized sentence.[73]

In *Stringer*, one of the aggravating factors used in determining a sentence was found to be vague or imprecise.[74] In a weighing state such as Mississippi, the sentence could not stand after the use of such a factor, unless the aggravating and mitigating circumstances were reweighed.[75] Because the aggravating factors were "weighed" in the sentencing process, as opposed to simply determining eligibility, a sentence in a weigh-

67. *Id.* at 745 n.2.
68. *Brown*, 126 S. Ct. at 890; *see also Stringer*, 503 U.S. at 232.
69. *Clemons*, 494 U.S. at 748-50.
70. *Id.* at 748-49; Hornbuckle, *supra* note 3, at 453-54.
71. *Clemons*, 494 U.S. at 753-54.
72. *Stringer*, 503 U.S. at 231.
73. *Id.* at 232.
74. *Id.* at 237.
75. *Id.*

ing state resulting from the use of an invalid factor could not stand without review.[76]

II. INSTANT CASE—*BROWN V. SANDERS*

A. Facts

The respondent, Ronald Sanders, and a companion broke into the home of the two victims.[77] They bound the victims and beat them on their heads with a blunt object.[78] One victim subsequently died.[79] Sanders was convicted of first-degree murder, attempted murder, robbery, burglary, and attempted robbery.[80] At the eligibility phase of sentencing, the jury found four special circumstances under California Penal Code 190.2[81], any of which would have made the defendant eligible for the death penalty.[82] At sentencing, after considering sentencing factors including the special circumstances from the eligibility phase and "the circumstances of the crime," the jury sentenced the respondent to death.[83]

B. Procedural History

The respondent appealed to the California Supreme Court.[84] The supreme court affirmed the death penalty, though it held that two of the special circumstances found by the jury in the eligibility phase were invalid, under the weighing standard from *Zant v. Stephens*.[85] After exhaustion of state remedies,[86] the defendant filed a motion for a writ of habeas corpus in the United States District Court of the Eastern District of California.[87] The district court denied relief.[88] The Court of Appeals for the Ninth Circuit reversed the sentence, on the grounds that the rule from *Zant* applied by the state court was not applicable to California as a weighing state.[89] The United States Supreme Court granted certiorari to determine whether California is a weighing or non-weighing state, and whether the consideration of invalid sentencing factors by the jury required the sentence to be vacated.[90]

---

76. *Id.*
77. Brown v. Sanders, 126 S. Ct. 884, 888 (2006).
78. *Id.*
79. *Id.*
80. *Id.*
81. CAL. PEN. CODE § 190.2 (West 2006).
82. *Brown*, 126 S.Ct. at 888.
83. *Id.*
84. *Id.*
85. *Id.*
86. *See* 28 U.S.C.A. § 2254(b)(A) (West 2006).
87. *Brown*, 126 S.Ct. at 888; *see also* 28 U.S.C.A. § 2254(d)(1).
88. *Brown*, 126 S.Ct. at 889.
89. *Brown*, 126 S.Ct. at 889.
90. *Id.*

C. The Majority Decision

Justice Scalia issued the opinion in *Brown v. Sanders*, joined by Chief Justice Roberts and Justices O'Connor, Kennedy and Thomas.[91] He began by distinguishing the two methods currently employed by the states to meet the narrowing requirement for death sentences required by *Furman v. Georgia*.[92] Scalia explained the procedures used in sentencing by weighing and non-weighing states, specifying, however, that all jurisdictions are required to meet the requirements for guided discretion and individualized sentencing by allowing a sentencer to weigh the factors meriting a death sentence with mitigating circumstances.[93] Scalia also devoted extensive dicta at this point to arguing against Justice Breyer's understanding that all jurisdictions require harmless error review of invalid sentencing factors, as well as highlighting the distinction between the types of jurisdictions as it was discussed in *Stringer*.[94] Scalia argued that *Zant* did not present a requirement for harmless error review, and that there is no such requirement in non-weighing states created in *Clemons*.[95]

However, because both types of jurisdictions face similar problems with invalid factors, the majority declared a new rule, eliminating the distinction between weighing and non-weighing jurisdictions.[96] Invalid sentencing factors will not upset a sentence of death unless analysis of that factor would provide a jury with facts and circumstances to which it would not otherwise have access.[97] The trigger of a requirement for harmless error review would be the presentation of new facts to a sentencer that it would not have seen without analysis of an invalid sentencing factor.[98]

Scalia argued that part of the reasoning for the elimination of the distinction is that most jurisdictions allow evidence to be presented to the sentencer through an eligibility factor or a sentencing factor, but that not all states fit neatly into the weighing/non-weighing categorization.[99] He noted that even in states that were placed by the court into one of the two categories, the particular scheme may have had elements of both.[100]

The opinion noted that under the former classification, California would have been a non-weighing state, validating the death sentence in

91. *Id.* at 884.
92. *Id.* at 889-91.
93. *Brown*, 126 S. Ct. at 889-91.
94. *Id.* at 891 n.3.
95. *Id.*
96. *Id.* at 891-92.
97. *Id.* at 892.
98. *Id.*
99. *Id.* at 891-92.
100. *Id.* at 892 n.5 (discussing *Stringer*'s use of an invalidated aggravating circumstance that was not an eligibility factor).

Brown under both the old and new systems.[101] Because the sentencing jury was able to consider the facts and circumstances related to the invalid factors under the heading of another valid factor, the majority held that the sentencing process was not skewed, and that the sentence was constitutional.[102]

D. Justice Stevens' Dissent

Justice Stevens' dissent, in which he is joined by Justice Souter, argued that the majority failed to address the question on which certiorari was granted.[103] The issue presented to the Court was whether California is a weighing state.[104] Though the majority did provide an answer to this question, Stevens maintained that the Court's choice to change settled sentencing law will complicate future decisions, and does not address the concerns presented by the California court in that context.[105]

E. Justice Breyer's Dissent

Justice Breyer, joined by Justice Ginsburg, reiterated that the question on certiorari was whether California is a weighing state.[106] However, Breyer argued that the more important issue for sentence review should turn on the nature of the sentencing error at trial rather than on the category of the issuing jurisdiction.[107] Appellate review of all death sentences rendered using invalid sentencing factors should be concerned with whether an error in sentencing was harmless beyond a reasonable doubt.[108]

Breyer discussed his opinion in regard to the two stages of sentencing that he found relevant in all jurisdictions, though he mentioned briefly that some states combine the stages into one proceeding.[109] First, sentencers determine eligibility for the death penalty, and only after this process do they weigh mitigating circumstances against aggravating circumstances (often in the form of eligibility factors).[110] Breyer argued that both types of jurisdictions face the same risks when juries consider invalid sentencing factors: giving undue weight to an issue or piece of evidence that should not have been under consideration.[111] Because information is presented under the rubric of an aggravating factor, it is

101. *Id.* at 893.
102. *Id.* at 894.
103. *Id.* at 896.
104. *Id.*
105. *Id.*
106. *Id.*
107. *Id.* at 898.
108. *Id.* at 896, 898.
109. *Id.* at 896-97.
110. *Id.*
111. *Id.* at 898-99.

unduly weighted for sentencers whether or not they may consider additional factors.[112]

Breyer read the decisions in *Zant* and *Clemons*, though they created the weighing/non-weighing distinction, as turning on the harmlessness of the errors made at trial: "Despite the Court's occasional suggestion to the contrary, the weighing/nonweighing distinction has little to do with the need to determine whether the error was harmless . . . reviewing courts should decide if that error was harmful, regardless of the form a State's death penalty law takes."[113] Breyer discussed the prejudice that can result from a sentencer's consideration of an invalid sentencing factor, reiterating that much of the distinction between the types of jurisdictions does not accurately reflect the statutory constructions of a number of states.[114] Breyer presented California as an example of the failure of the categories, as a state which presents one set of factors for guided discretion, and which adds additional factors (which define but do not limit the circumstances considered) for the individualized sentencing procedure.[115]

He analyzed *Stringer v. Black* as the first case to frame the appellate review process as a weighing/non-weighing issue, and to equate invalid factors in non-weighing states with harmless error automatically.[116] For Breyer, this is not an accurate depiction of death penalty jurisprudence. He also found that Scalia denied the importance of the emphasis placed on sentencing evidence inconsistently with *Clemons*, and "diminishe[d] the need to conduct any harmless-error review at all."[117] The majority decision to treat any error that does not present new evidence as harmless will limit the actual and necessary harmless-error review in many death penalty cases to come.[118]

III. ANALYSIS

This comment addresses *Brown v. Sanders*[119] in the light of capital sentencing precedent, and questions its impact on future cases. It commends the majority for the elimination of the distinction between weighing and non-weighing jurisdictions, as this will clarify the issues surrounding invalid sentencing factors. However, it questions the opinion's failure to set a universal standard for appellate review under the new scheme.

112. *See id.*
113. *Id.* at 898, 900-02.
114. *Id.* at 898-900.
115. *Id.* at 900.
116. *Brown*, 126 S. Ct. at 902.
117. *Id.* at 903.
118. *See id.* at 903-04.
119. 126 S. Ct. 884 (2006).

A. Attempting to Clarify Invalid Factor Issues – Eliminating the Weight

The majority in *Brown* eliminated the developing distinction between weighing and non-weighing states[120] that has been in the background of invalid factor cases since *Zant v. Stephens*.[121] The *Stringer v. Black* decision considered that distinction to be of "critical importance,"[122] but the lengthy analysis of the distinction in that opinion has not been easy to understand. The majority and Breyer's dissent in *Brown* followed different interpretations of the rule presented in *Stringer*,[123] and lower courts could easily share the confusion.

The decision in *Brown* recognized that categorizing states as weighing or non-weighing[124] was not accurate and did not solve the problem inherent in sentences imposed after consideration of invalid factors.[125] As one scholar noted, "[t]hat nonweighing states retain the weighing metaphor to describe the sentencing process suggests that the distinction the Court has created an illusion. Moreover, the distinction has engendered confusion and led to incoherent decisions."[126] The weighing/non-weighing distinction raised concerns that the same analysis took place in both types of jurisdictions, that the terms used in distinguishing the jurisdictions provided undue weight to sentencing factors, and that the states could not be broken into the two categories previously recognized by the Supreme Court.

1. Similar Analysis Across Jurisdictions

The Scalia opinion argued that the distinction between weighing and non-weighing jurisdictions can be eliminated because the same essential process of balancing aggravating and mitigating factors is used by all juries.[127] This view was the foundation for the opinion's emphasis on what evidence the jury has access to during the sentencing process.[128] In the instant case:

> [T]he jury's consideration of the invalid eligibility factors in the weighing process did not produce constitutional error because all the facts and circumstances admissible to establish [the invalid factors] were also properly adduced as aggravating facts bearing upon the "circumstances of the crime" sentencing factor. They were properly

120. *Brown*, 126 S. Ct. at 892.
121. Zant v. Stephens, 462 U.S. 862, 888 n.24 (1983); *see also* Stringer v. Black, 503 U.S. 222, 229 (1992).
122. *Stringer*, 503 U.S. at 231-32.
123. *Brown*, 126 S. Ct. at 891 n.3.
124. *See Stringer*, 503 U.S. at 231-32.
125. *Brown*, 126 S. Ct. at 891-92.
126. Widder, *supra* note 18, at 1365 (noting an Illinois decision which mischaracterized the state as a weighing jurisdiction, relying on the analysis provided in *Stringer*).
127. *Brown*, 126 S. Ct. at 892.
128. *See id.* at 892.

> considered whether or not they bore upon the invalidated eligibility factors.[129]

This is the justification the Court relied upon for the *Brown* decision,[130] but there were other points in favor of eliminating the distinction as well.

2. Weighing Terminology Skewed Analysis of Sentencing Factors

Some scholarship on the weighing/non-weighing distinction focused on the terms used by the courts as much as on their analysis. When a jury is given a set of facts and circumstances, increased significance is given those which are recognized by statute or considered to be automatically "aggravating," no matter whether the jurisdiction officially "weighs" those factors or not.[131] One analysis of *Clemons v. Mississippi* highlighted the importance of these labels to a sentencing jury, and the inconsistent treatment of it by the Court:

> The underlying rationale of the *Clemons* opinion must be that the jury goes about its decisionmaking [sic] process in a different way when it is explicitly instructed to weigh aggravating and mitigating circumstances. Otherwise the distinction makes little sense -- if the thought process is the same as it is in a nonweighing state, then either no reweighing is required or the sentence must be reweighed under both types of statutory schemes.[132]

In distinguishing between two types of statutes when *Clemons* was decided, the Court may have allowed an inconsistency between the treatment of *Zant* and *Clemons* to become law.[133] The Court's somewhat backward terminology[134] further muddied the waters, giving jurisdictions two categories based on the word "weigh," while the distinction was actually based on the limitation of factors presented to the jury, rather than what it does with them.[135] In relying on this inconsistency, courts may have been lulled into using a metaphor that is unrelated to the actual statutes determining sentencing procedure.[136]

3. The Distinction Was Illusory

A concern in Breyer's dissent to *Brown* was that it is rarely possible to cleanly categorize jurisdictions as weighing or non-weighing.[137] While some states do mirror the classic weighing or non-weighing para-

---

129. *Id.* at 894.
130. *Id.*
131. *See* Widder, *supra* note 17, at 1370-71.
132. Hornbuckle, *supra* note 3, at 455.
133. *See Zant*, 462 U.S. at 873-74; *Clemons*, 494 U.S. at 743-44.
134. *See Brown*, 126 S. Ct. at 889, 898.
135. *See Id.*
136. Widder, *supra* note 17, at 1363-64.
137. *See Brown*, 126 S. Ct. at 898 (Breyer, J., dissenting).

digms of *Zant* and *Clemons*, others "fall somewhere in between."[138] California is one example, using specific factors for eligibility, and adding factors to the list to be used in sentencing.[139] Also, as noted in *Zant*, states are not required to follow sentencing schemes such as Georgia's in order to meet the *Furman v. Georgia* requirement for guided discretion.[140] If a state were to choose a different system for narrowing its class of capital offenders, it might have no place at all in the weighing/non-weighing scheme.

Breyer also noted that some jurisdictions have combined the eligibility and sentencing stages into one sentencing process.[141] In such a jurisdiction, there may be no way for an appellate review to determine what evidence is limited to eligibility and what is limited to sentence selection, and there is less chance that a jury would make such a distinction. Though the use of the weighing and non-weighing categories had been useful for analyzing several specific statutes, it does not seem that it was suited to bear the entire weight of capital sentencing. The Court used it as an explanatory tool, but it may never have been intended as a means of determining the constitutionality of all sentences. As one critic noted, "[t]he Supreme Court's weighing doctrines allow procedure to distort substance in an area of law in which it is acutely necessary that procedural rules be finely tuned to promote substantive law."[142] It is clear from *Brown's* ease in removing the distinction,[143] as well as from the complicated analysis engendered from its use[144] that the termination of classifying jurisdictions as weighing or non-weighing will not hinder substantive law in capital sentencing.

B. Unseen Risks? Where Is the Emphasis on Appellate Review?

Though *Brown* attempts to clarify sentencing decisions, it contains a flaw that could have serious repercussions. While it does not overturn any existing law on appellate reweighing or harmless-error review, the opinion fails to make this crucial element of invalid factor analysis clear for lower courts to apply along with its new rule on evidentiary analysis by trial juries. Appellate review, which may consist of reweighing of sentencing factors or a harmless-error review to determine the impact of the consideration of an invalid factor, is an important element of the decisions in invalid sentencing factor precedent.[145] *Stringer*'s discussion of harmless-error review and appellate review has led to completely different interpretations of the requirement. With the consolidation of weigh-

138. *Id.* at 900..
139. *Id.*
140. *Zant*, 462 U.S. at 874-75.
141. *Brown*, 126 S. Ct. at 900.
142. Widder, *supra* note 17, at 1346.
143. *See Brown*, 126 S. Ct. at 892.
144. Widder, *supra* note 17, at 1365; *see also Stringer*, 503 U.S. at 232-33.
145. *Clemons*, 494 U.S. at 749; *Zant*, 462 U.S. at 888, 890; *see also Stringer*, 503 U.S. at 236.

ing and non-weighing jurisdictions, jurists must exercise care that this constitutional requirement does not fall by the wayside.

1. Appellate Review in Capital Sentencing Precedent

The *Brown* decision highlighted the need for harmless-error review in weighing jurisdictions,[146] but the requirement for appellate review is not limited to those states. *Zant* recognized appellate review as an "important procedural safeguard."[147] It also made clear that whether the analysis of an invalid factor is a constitutional error depends on the specific circumstances of a case.[148] Though *Zant* is an example of a case in which a sentence imposed using an invalid factor stood, the sentence was only validated through appellate review.

The decision in *Clemons* hinged on the importance of appellate review in the form of reweighing of sentencing factors.[149] The case recognized that a harmless-error review of some kind took place, but the Supreme Court held that it was insufficient under the circumstances.[150] The *Clemons* decision did not outline the exact requirement for appellate review after consideration of an invalid factor, but it was made clear that reweighing or appellate review of some kind was necessary after a sentencer considered an invalid sentencing factor.[151] The Court acknowledged that the state court's reliance on one valid circumstance for sentencing[152] was "not conducting appellate reweighing as we understand the concept,"[153] and reversed the state court's decision.[154] Later analysis of this case indicates an understanding of the importance of appellate review, but case law presents no clear distinction between the general requirement for appellate review of death sentences and the specific processes of reweighing of factors[155] and harmless-error review.[156] Unfortunately, while Scalia devoted discussion to harmless error review in regard to his digression on Breyer's dissent, he did not clarify its role in the new system.[157] This may be the basis for Breyer's fear that harmless error review will no longer occur at all.[158]

146. *Brown*, 126 S. Ct. at 890.
147. *Zant*, 462 U.S. at 890.
148. *Brown*, 126 S. Ct. at 901 (Breyer, J., dissenting).
149. *See Clemons*, 494 U.S. at 749-50.
150. *Id.* at 740, 753.
151. *Id.* at 740.
152. *Id.* at 751.
153. *Id.* at 752.
154. *See* Widder, *supra* note 17, at 1353 n.133 ("The *Clemons* Court did not explain its mysterious distinction between harmess-error analysis and reweighing aggravating and mitigating factors.").
155. *See* Hornbuckle, *supra* note 3, at 453-54.
156. *See Brown*, 126 S. Ct. at 901 (Breyer,J., dissenting).
157. *See id.* at 891 n.3 (majority opinion).
158. *Id.* at 903 (Breyer, J., dissenting).

Appellate review as a constitutional requirement was set out in *Barclay*[159], and has remained an element of invalid sentencing factor decisions since. However, the precise requirements for appellate review have not been set out, and the repercussions of the vague standard are apparent in *Brown*'s lack of discussion on the issue and in Breyer's dissent.[160] In fact, the only mention of appellate review in the decision that is not dicta for the benefit of Breyer is a procedural note of the Ninth Circuit's holding that a harmless-error review was necessary.[161] Nonetheless, scholarship recognizes the appellate review requirement based on the line of cases on invalid sentencing factors.[162] What its requirements are, and how the terms used differ in meaning, remain a murky area not yet clarified by the courts.[163] The appellate review requirement is an area that has not been adequately explored, and which may be the source of the confusion and differing interpretations of *Stringer*.

2. The Confusing Heritage of *Stringer*

Stringer recognized that an appellate review of a death sentence took place at the state level.[164] It reiterated the general requirement of "appellate scrutiny of the import and effect of invalid aggravating factors"[165] and clarified the *Clemons* requirement that at least in weighing jurisdictions, there must be harmless-error review.[166] However, where *Stringer* fits into non-weighing jurisdictions and how it affects future sentences[167] is a subject of disagreement.

Some instances within *Stringer* seem to apply to invalid factor sentencing in general.[168] Consider the final statement of the *Stringer* court before its reversal order: "the precedents even before *Maynard* and *Clemons* yield a well-settled principle: Use of a vague or imprecise aggravating factor in the weighing process invalidates the sentence and at the very least requires constitutional harmless-error analysis or reweighing in the state judicial system."[169] There is no question that *Stringer* is an example of a weighing state.[170] Nonetheless, as *Clemons* was the first example of a weighing jurisdiction to be presented to the United States Supreme Court,[171] any precedent prior to it must have referred to a non-weighing jurisdiction. Thus, it would seem that the proposition in

159. Barclay v. Florida, 463 U.S. 939, 1941 (1983).
160. *See Brown*, 126 S. Ct. at 901.
161. *Id.* at 889.
162. Hornbuckle, *supra* note 3, at 442; Widder, *supra* note 17, at 1344, 1354.
163. *See* Widder, *supra* note 17, at 1370 n.133.
164. *Stringer*, 503 U.S. at 234.
165. *Id.* at 230.
166. Srinivasan, *supra* note 2, at 1368.
167. *Brown*, 126 S. Ct. at 889.
168. *See Stringer*, 503 U.S. at 230-31, 236.
169. *Id.* at 237.
170. *See id.* at 232.
171. *Clemons*, 494 U.S. at 744-45.

Stringer requiring appellate reweighing or harmless-error review[172] applies to both types of jurisdictions.

Justice Scalia, in his *Brown* opinion, vehemently disagreed with this point: "Justice Breyer contends that harmless-error review applies in *both* weighing and non-weighing States. It would be strange indeed to discover at this late stage that our long-held distinction between the two sorts of States for purposes of reviewing invalid eligibility factors in fact made no difference."[173] Justice Scalia did not mention in this note the other purported differences between the types of jurisdictions, such as the use of limited sentencing factors in weighing states versus broad or unlimited factors or circumstances in sentencing in non-weighing states.[174] Scalia, in recognizing that harmless error review is an important element at least in weighing jurisdictions, seemed to recognize that there is a role for appellate review of sentences rendered using invalid factors even under the new system.[175] Unfortunately, Scalia made no mention of this role in the body of the case.[176]

Justice Scalia perhaps overstated Justice Breyer's interpretation of *Stringer*. Breyer's contention that harmless-error review is necessary in all jurisdictions is based on *Zant* and *Clemons*.[177] As to *Stringer*, Breyer discussed its reference to non-weighing cases specifically as a "single ambiguous sentence of dicta" that should not be a basis for future law.[178]

Other scholarship indicates the same divergent interpretations of *Stringer*, some interpretations indicating that reweighing and harmless-error review apply only to weighing states,[179] and others applying it to both types.[180] There have also been no definitions of "appellate reweighing" or "harmless-error review" beyond the references in cases such as *Clemons* and *Stringer*. The terms may or may not be interchangeable: "Regardless of the validity of assessing capital sentencing errors under a harmless-error analysis, it remains unclear how harmless-error review is functionally different from appellate reweighing of aggravating and mitigating factors in evaluating the effect of invalid aggravating factors on a death sentence."[181] With no consensus on terms, and no consensus on what is a rule and what is dicta, or to which jurisdictions rules apply to, it is no small wonder that *Stringer* has been the source of confusion and discord. Nonetheless, *Brown* and its followers must still apply the requirement of appellate review to invalid factor sentencing.

---

172. *Stringer*, 503 U.S. at 230.
173. *Brown*, 126 S. Ct. at 891 n.3 (citation omitted).
174. *Id.* at 890.
175. *See id.* at 891 n.4.
176. *See id.* at 892.
177. *Brown*, 126 S. Ct. at 900-02 (Breyer, J., dissenting).
178. *Brown*, 126 S. Ct. at 902 (Breyer, J., dissenting).
179. Widder, *supra* note 17, at 1344.
180. Srinivasan, *supra* note 2, at 1367.
181. Widder, *supra* note 17, at 1371 n.133.

3. What Is Required for Appellate Review after *Brown*?

Despite the uniformity of capital sentencing jurisprudence at the Supreme Court level, there are varied opinions on what is required as far as appellate review of sentences based on invalid factors. One extreme holds that an erring sentence cannot ever stand: "When a sentence of death has been based in part on invalid aggravating factors, the only remedy faithful to the Eighth Amendment is to reverse the sentence and remand for a new sentencing proceeding."[182] An approach foreseen by Justice Breyer based on the majority opinion in *Brown* is a limiting of the appellate review requirement.[183] Breyer's favored approach stands between these points, more in line with prior cases: "A reviewing court must find that the jury's consideration of an invalid aggravator was harmless beyond a reasonable doubt, regardless of the form a State's death penalty law takes."[184] Based on sentencing precedent, it seems unlikely that there will be any changes in the requirements for a constitutional death sentence. But these alternative approaches to sentencing indicate a continued recognition of the vital role appellate sentencing takes.

Future cases will rely on *Brown* as well as on its predecessors, and one could wish that the opinion had made the Court's stance on appellate review clear within the decision that changed standing law from a two-pronged to a unified system. Without the distinctions of weighing or non-weighing states to rely on, lower courts may find themselves at a loss for what rule of appellate review to apply. Justice Scalia, in devoting space to arguing with Justice Breyer over what is not required under a now obsolete classification, failed to address the standard of review required in all jurisdictions under the new rule.[185] Even if *Stringer* clearly outlined the requirements for weighing states[186] (which is a debatable proposition in itself), there is no new law as to what is required for the jurisdictions as they are now unified. So, it seems that lower courts must continue to rely on the law as it was defined in *Zant*, "the mandatory appellate review of each death sentence . . . ."[187]

If the Court chooses to clarify its stance on appellate review in a future decision, the prudent course may be an approach which applies the most stringent standard. If weighing and non-weighing jurisdictions are no longer distinct, the level of review required should be that required for weighing jurisdictions, which previously held a greater risk of harmful error.[188] According to the most recent appellate review precedent,

182. *Id.* at 1346.
183. *Brown*, 126 S. Ct. at 903 (Breyer, J., dissenting).
184. *Id.* at 896.
185. *See id.* at 891 n.3 (majority opinion).
186. *Stringer*, 503 U.S. at 232.
187. *Zant*, 462 U.S. at 890.
188. *See Brown*, 126 S. Ct. at 890; *see also Stringer*, 503 U.S. at 232.

Stringer, the new rule for all invalid factor sentencing cases would be that "[w]hen the weighing process itself has been skewed, only constitutional harmless-error analysis or reweighing at the trial or appellate level suffices to guarantee that the defendant received an individualized sentence."[189] If, as Justice Scalia stated, all jurisdictions require a weighing of aggravating and mitigating factors,[190] it seems only appropriate that all jurisdictions now be considered weighing jurisdictions.

CONCLUSION

Brown v. Sanders represents forward progress in its elimination of the unwieldy distinction between weighing and non-weighing jurisdictions in the treatment of invalid sentencing factors. However, it fails to bring up to date the law on the appellate review requirement of erring capital sentencing cases, and thus could present future problems. The most prudent course would be for the Court to clearly outline the requirement for all cases sentenced under invalid factors to be subjected to reweighing of the sentencing factors or a harmless-error review. The necessity for meaningful appellate review cannot be understated in the case of sentencing error. As the Court exhorted in *Zant v. Stephens*: "because there is a qualitative difference between death and any other permissible form of punishment, 'there is a corresponding difference in the need for reliability in the determination that death is the appropriate punishment . . . .'"[191] Precedent provides a foundation for reliability through appellate review, but it is the responsibility of the Court to make it law.

Karen Lamprey[*]

189. *Stringer*, 503 U.S. at 232.

190. *Brown*, 126 S. Ct. at 889.

191. *Zant*, 462 U.S. at 884-85 (quoting Woodson v. North Carolina, 428 U.S. 280, 305 (1976)).

*. J.D. Candidate, 2008, University of Denver Sturm College of Law. Thanks to Professor Sam Kamin and Steve Arvin for their editorial assistance, and to Nate Obee for his unfading and constant support.