

11-1983

A Tax Policy Analysis of Bob Jones University v. United States

Charles O. Galvin

Neal Devins

Follow this and additional works at: <https://scholarship.law.vanderbilt.edu/vlr>



Part of the [Tax Law Commons](#)

Recommended Citation

Charles O. Galvin and Neal Devins, A Tax Policy Analysis of Bob Jones University v. United States, 36 *Vanderbilt Law Review* 1353 (1983)

Available at: <https://scholarship.law.vanderbilt.edu/vlr/vol36/iss6/1>

This Article is brought to you for free and open access by Scholarship@Vanderbilt Law. It has been accepted for inclusion in Vanderbilt Law Review by an authorized editor of Scholarship@Vanderbilt Law. For more information, please contact mark.j.williams@vanderbilt.edu.

A Tax Policy Analysis of *Bob Jones University v. United States*

Charles O. Galvin* and Neal Devins**

I. INTRODUCTION	1354
II. TAX POLICY AND THE THREE BRANCHES OF GOVERNMENT	1355
III. <i>Bob Jones University v. United States</i>	1356
A. <i>The Meaning of the Tax Exemption Provision</i>	1359
1. Analysis of the Public Benefit Doctrine ...	1365
2. Analysis of the Common Community Conscience Requirement	1366
3. Analysis of the Tax Exemption and the "Tax Expenditure" Budget	1368
B. <i>The Scope of IRS Rulemaking Authority</i>	1371
C. <i>Issues That the Court Did Not Resolve</i>	1374
1. Is a Tax Exemption Government Aid? ...	1374
2. Can Some of a Tax-Exempt Organization's Practices Violate Public Policy?	1376
3. Does Granting Tax-Exempt Status Trigger Constitutional Scrutiny?	1377
IV. CONCLUSION	1379

* Centennial Professor of Law, Vanderbilt University. B.S.C., Southern Methodist University, 1940; M.B.A., Northwestern University, 1941; J.D., 1947; S.J.D., Harvard, 1961. Former member, Advisory Group on Exempt Organizations to the Commissioner of Internal Revenue.

** Director, Religious Liberty and Private Education Project, Institute for Public Policy Studies, Vanderbilt University. B.A., Georgetown University, 1978; J.D., Vanderbilt University, 1982. Mr. Devins' research is supported by a grant from the Institute for Educational Affairs. The views expressed are those of the authors.

I. INTRODUCTION

In one of last term's most notable decisions, the United States Supreme Court in *Bob Jones University v. United States*¹ considered the meaning of the tax exemption provisions of the Internal Revenue Code (the Code) and the relationships among the Internal Revenue Service (IRS), Congress, and the courts in formulating tax policy. Affirming an IRS ruling² that denied tax exemptions to racially discriminatory private schools, the Court devised a model for the interaction of the three branches of government in tax policy matters. This model assigns to the IRS primary authority to develop rules governing the implementation of the tax exemption laws and assigns to the courts and Congress secondary authority to oversee the IRS. The decision seems to alter the earlier model in which the courts had retained primary authority to determine statutory intent and to review IRS actions in the tax exemption area. Furthermore, the Court interpreted the applicable Code provisions under which Bob Jones University claimed tax exemption as requiring that the institution confer a "public benefit" and have a purpose in harmony with a "common community conscience."³

This Article questions the tax policy model that the Court articulated in *Bob Jones University*. The authors believe that the Court's recognition of the primacy of IRS rulemaking is undesirable because the IRS, as an executive agency, is susceptible to the influence of the incumbent administration's policy objectives. Further, even though the life-tenured status of judges insulates the courts from external political pressures, significant problems are also present in a model in which the courts occupy a primary role in formulating tax policy. In sum, Congress is better suited than either the courts or the IRS to determine tax policy because it is institutionally organized to gather social and economic data, to define policy objectives, and to legislate to achieve these objectives, which often have repercussions beyond the circumstances of a particular case.

This Article's criticism of *Bob Jones University* also extends to the Court's "public benefit" and "common community conscience" standards for charitable organizations seeking to qualify for tax exemptions. Although restrictive standards suggest that a tax exemption is a form of government aid, the Court declined to

1. 103 S. Ct. 2017 (1983).

2. Rev. Rul. 71-447, 1971-2 C.B. 230.

3. 103 S. Ct. at 2028-29; see *infra* notes 33-56 and accompanying text.

make this holding explicitly.⁴ Moreover, the majority's recognition of the IRS's broad rulemaking authority seems inconsistent with the "community conscience" standard, for in the exercise of its broad administrative discretion the IRS need not strictly follow this standard. Finally and most significantly, the "public benefit" and "community conscience" standards may discourage organizations that provide a healthy diversity of views in a pluralistic society.

This Article begins with a general discussion in part II of the role of the courts in the development of federal tax policy.⁵ A critical analysis of the *Bob Jones University* decision—focusing upon both the specific tax exemption issue and the Court's general model for tax policy decisionmaking—follows in part III.⁶ Part IV concludes the Article with a recommendation that Congress act definitively to take the lead in formulating tax policy.⁷

II. TAX POLICY AND THE THREE BRANCHES OF GOVERNMENT

Congress, the courts, and the executive branch—through the IRS and the Treasury Department—historically have all interacted in formulating tax policy. Congress, of course, has contributed tax legislation—legislation that has evolved from the Revenue Act of 1913 to the extraordinarily complex rules of the present Internal Revenue Code. The Treasury Department and the IRS have added extensive regulations and rulings to the already unwieldy tax legislation. Courts then have attempted to work their way through the resulting murkiness—the "dank, miasmic, myxomycetous sump."⁸ Judge Learned Hand artfully described the difficulty confronting the courts:

[T]he words of such an act as the Income Tax . . . merely dance before my eyes in a meaningless procession: cross-reference to cross-reference, exception upon exception—couched in abstract terms that offer no handle to seize hold of—leave in my mind only a confused sense of some vitally important, but successfully concealed, purport. . . . I know that these monsters are the result of fabulous industry and ingenuity . . . ; one cannot help wondering whether to the reader they have any significance save that the words are

4. See *infra* notes 96-110 and accompanying text.

5. See *infra* notes 8-15 and accompanying text.

6. See *infra* notes 16-120 and accompanying text.

7. See *infra* notes 121-31 and accompanying text.

8. Galvin, *More on Boris Bittker and the Comprehensive Tax Base: The Practicalities of Tax Reform and the ABA's CSTR*, in B. BITTKER, C. GALVIN, R. MUSGRAVE & J. PECHMAN, *A COMPREHENSIVE INCOME TAX BASE?* 89 (1968) (reprints articles and additional remarks by these authors from 80 HARV. L. REV. 925 (1967) and 81 HARV. L. REV. 44, 63, 1016, and 1032 (1968)).

strung together with syntactical correctness.⁹

Notwithstanding the difficulties of interpretation, the courts profoundly have influenced the development of tax law. As one commentator noted, "Congress has the first word in [tax formulation] . . . but the courts have the last word."¹⁰ Similarly, as two former Justice Department officers observed, "Frequently, a phrase in a ruling, a court opinion, or even an argument in a government tax brief can have as much impact on tax policy as a new federal tax statute."¹¹ A court's determination of tax policy generally serves to refine congressional legislation to fit contemporary needs. Unlike tax legislation, which "prospectively formulates rules with universal applicability[,] . . . tax litigation formulates rules retrospectively, aiming principally at resolving disputes of immediate concern."¹²

Thus, under this traditional model the courts had primary authority to shape the meaning of the Code, usually after a private party's challenge of an IRS directive¹³ or procedure,¹⁴ or the government's appeal of an unfavorable lower court decision.¹⁵ *Bob Jones University* is significant because it alters the structure of decisionmaking by placing primary supervisory authority in the IRS.

III. *Bob Jones University v. United States*

Bob Jones University calls itself "the world's most unusual university."¹⁶ Although unaffiliated with any established church, the University is dedicated to the teaching and propagation of fundamentalist religious beliefs.¹⁷ In pursuit of these goals the University dictates strict rules of conduct for its students.¹⁸ To enforce

9. Hand, *Thomas Walter Swan*, 57 *YALE L.J.* 167, 169 (1947).

10. R. PAUL, *TAXATION IN THE UNITED STATES* 656 (1954 ed.).

11. Ferguson & Henzke, *The Formulation of Federal Tax Policy Through Litigation*, 1 *V.A. TAX REV.* 85, 85 (1981).

12. *Id.* at 86.

13. See, e.g., *Bob Jones Univ. v. United States*, 103 S. Ct. 2017, 2023 (1983) (discussing *Bob Jones Univ. v. Simon*, 416 U.S. 725 (1974)); see *infra* text accompanying note 22.

14. See, e.g., *Wright v. Regan*, 656 F.2d 820 (D.C. Cir. 1981), *cert. granted*, 103 S. Ct. 3109 (1983); *infra* note 95.

15. See Ferguson & Henzke, *supra* note 11, at 91-98.

16. *Bob Jones Univ. v. Simon*, 416 U.S. 725, 734 (1974).

17. *Id.*

18. For example,

The institution does not permit dancing, card playing, the use of tobacco, movie-going, and other such forms of indulgences in which worldly young people often engage; no student will release information of any kind to any local newspaper, radio station, or

one such rule forbidding interracial dating and marriage, the University denies admission to applicants engaged in or known to advocate interracial dating and marriage.¹⁹

The *Bob Jones University* controversy began in November 1970 when the United States District Court for the District of Columbia in *Green v. Kennedy*²⁰ enjoined the IRS from according tax-exempt status to racially discriminatory private schools in Mississippi. The *Green* court suggested that the IRS would not be permitted to grant tax-exempt status to institutions that violate the government's public policy of nondiscrimination. The IRS then reversed its position of granting tax exemptions to racially discriminatory institutions and notified the University that it intended to challenge the tax-exempt status of private schools with racially discriminatory admissions policies.²¹ In response, the University in 1971 sought to enjoin the IRS from revoking its tax-exempt status.

television station without first checking with the University Public Relations Director; students are expected to refrain from singing, playing, and, as far as possible, from "tuning-in" on the radio or playing on the record player jazz, rock-and-roll, folk rock, or any other types of questionable music; and, no young man may walk a girl on campus unless both of them have a legitimate reason for going in the same direction.

Note, *The Internal Revenue Service's Treatment of Religiously Motivated Racial Discrimination by Tax Exempt Organizations*, 54 NOTRE DAME LAW. 925, 925 (1979) (quoting *Bob Jones Univ. v. United States*, 468 F. Supp. 890, 894 (D.S.C. 1978)).

19. *Bob Jones Univ. v. United States*, 103 S. Ct. 2017, 2023 (1983). The sponsors of the University believe that the Bible forbids interracial dating and marriage. *Id.* at 2022.

20. 309 F. Supp. 1127 (D.D.C.), *appeal dismissed sub nom. Cannon v. Green*, 398 U.S. 956 (1970), *appeal dismissed sub nom. Coit v. Green*, 400 U.S. 986 (1971). The IRS decision to deny tax-exempt status to discriminatory private schools was in accord with plaintiffs' argument in *Green* that to grant tax exemptions to schools that violate the important public policy objectives established in *Brown v. Board of Educ.* and the Civil Rights Act of 1964 would be improper. See IRS News Release (July 10, 1970); N.Y. Times, July 11, 1970, at A1, col. 8. The court made permanent the *Green v. Kennedy* temporary injunction in *Green v. Connally*, 330 F. Supp. 1150 (D.D.C.), *aff'd*, 404 U.S. 997 (1971). Until the Supreme Court's recent decision in *Bob Jones University*, the precedential effect of *Green* was unclear. In *Bob Jones Univ. v. Simon*, 416 U.S. 725 (1974), the Court claimed that since the IRS was no longer maintaining an adversarial position against the plaintiffs in *Green* at the time of the trial, "the [Supreme] Court's affirmance in *Green* lacks the precedential weight of a case involving a truly adversary controversy." *Id.* at 740 n.11.

21. *Bob Jones Univ.*, 103 S. Ct. at 2023. Prior to 1971, Bob Jones University completely excluded blacks. From 1971 to 1975, the University accepted applications from blacks married within their race. Since May 1975, the University has permitted unmarried blacks to enroll subject to this disciplinary rule prohibiting interracial dating and marriage. See *id.* The Fourth Circuit's decision in *McCrary v. Runyon*, 515 F.2d 1082 (4th Cir. 1975), *aff'd*, 427 U.S. 160 (1976), influenced the University's decision to open its admissions to blacks. *McCrary* held that 42 U.S.C. § 1981 prohibited private schools from denying admissions to students on the basis of race. 515 F.2d at 1086-87. *McCrary*, however, did not address whether a private school can deny admissions on the basis of race as a matter of religious belief.

That suit culminated in a 1974 Supreme Court decision that "prohibited the University from obtaining judicial review by way of injunctive action before the assessment or collection of any tax."²²

The IRS in January 1976 formally revoked the University's tax exemption.²³ After paying a portion of the federal unemployment taxes due, the University filed suit for a refund, contending that it was statutorily and constitutionally entitled to reinstatement of its tax exemption.²⁴ In April 1981 the United States Court of Appeals for the Fourth Circuit upheld the revocation of the exemption.²⁵ The Supreme Court granted certiorari in *Bob Jones University* and in *Goldsboro Christian Schools, Inc. v. United States*,²⁶ a case presenting identical issues. On January 8, 1982, the Justice Department petitioned the Court to vacate these cases as moot in light of the Reagan administration's decision to reinstate the tax-exempt status of racially discriminatory private schools.²⁷ Because of a related court order that prevented the administration

22. *Bob Jones Univ.*, 103 S. Ct. at 2023 (discussing *Bob Jones Univ. v. Simon*, 416 U.S. 725 (1974)).

23. Revenue Ruling 71-447 formally established the policy of prohibiting the granting of tax exemptions to private schools that maintained racially discriminatory policies. Rev. Rul. 71-447, 1971-2 C.B. 230. Revenue Procedure 72-54 required private schools to publicize their nondiscriminatory policies, although it demanded no particular method of publication. Rev. Proc. 72-54, 1972-2 C.B. 834. The IRS in 1975 updated its stipulations for private schools seeking tax-exempt status. Revenue Procedure 75-50 provided guidelines and mandated recordkeeping to assess whether a private school's policies were racially nondiscriminatory. Rev. Proc. 75-50, 1975-2 C.B. 587. The regulation mandates that tax-exempt institutions: (a) adopt formally nondiscriminatory policies in their charters or bylaws, (b) refer to such policies in their advertising brochures, and (c) publish annual notice of such policies in a local newspaper of general circulation. *Id.* §§ 4.01-.03, 1975-2 C.B. 587-88. See Devins, *Tax Exemptions for Racially Discriminatory Private Schools: A Legislative Proposal*, 20 HARV. J. ON LEGIS. 153, 157 (1983). The IRS in 1975 also published a revenue ruling denying tax-exempt status to any religious institution that maintained racially discriminatory policies, even if sincere religious belief was the basis of that discrimination. Rev. Rul. 75-231, 1975-1 C.B. 158. Current IRS policies rely on those two 1975 rulings. For a general description of federal governmental actions on this issue prior to the Reagan policy shift, see Devins, *supra*, at 155-61.

24. *Bob Jones Univ.*, 103 S. Ct. at 2023.

25. *Bob Jones Univ. v. United States*, 639 F.2d 147 (4th Cir. 1981). The district court had concluded that the University was entitled to tax-exempt status on both statutory and first amendment grounds. *Bob Jones Univ. v. United States*, 468 F. Supp. 890, 897-98 (D.S.C. 1978).

26. 454 U.S. 892 (1981). The Court heard both cases simultaneously.

27. Memorandum for the United States, *Goldsboro Christian Schools, Inc. v. United States and Bob Jones Univ. v. United States*, 103 S. Ct. 2017 (1983). In addition, on January 8, 1982, the IRS announced that "without further guidance from Congress, the Internal Revenue Service will no longer deny tax-exempt status for . . . organizations on the grounds that they don't conform with certain fundamental public policies." IRS News Release (Jan. 8, 1982).

from reinstating this status,²⁸ however, the administration withdrew its request that the Court declare the cases moot.²⁹ The Supreme Court denied tax exemptions to the two petitioner schools on May 24, 1983.³⁰ In its decision the Court made certain general pronouncements both on the meaning of the Code's tax exemption provision and on the IRS's authority to issue rulings in accordance with its own interpretation of the Code.³¹ The majority held that a tax-exempt institution must confer some "public benefit" and that its purpose must not be at odds with the "common community conscience."³² The Court further held that the IRS has broad authority to interpret the Internal Revenue Code and to issue rulings based on its interpretation.

A. *The Meaning of the Tax Exemption Provision*

Section 501(c)(3) of the Code provides that "[c]orporations . . . organized and operated exclusively for religious, charitable, . . . or educational purposes" are entitled to tax-exempt status.³³

28. *Wright v. Regan*, No. 80-1124 (D.C. Cir. Feb. 18, 1982).

29. See Babcock, *Administration Asks High Court to Settle School Exemption Issue*, Wash. Post, Feb. 26, 1982, at A3, col. 4; Taylor, *School Tax Issue Put to High Court in Shift by Reagan*, N.Y. Times, Feb. 26, 1982, at A1, col. 1. The government, however, did not change its position that the IRS lacked authority to promulgate procedures denying tax-exempt status to racially discriminatory private schools. Thus, the Court felt compelled to appoint "counsel adversary"—William T. Coleman, Jr.—to argue the government's side of the case. *Goldsboro Christian Schools, Inc. v. United States and Bob Jones Univ. v. United States*, 50 U.S.L.W. 3837 (U.S. Apr. 19, 1982). For a critique of this Court decision, see Devins, *Is the Supreme Court on the Reagan Team?*, Christian Sci. Monitor, Nov. 3, 1982, at 22; McCoy & Devins, *Does the Bob Jones Case Meet the Case-or-Controversy Requirement?*, Nat'l L.J., Oct. 18, 1982, at 18, col. 1.

30. *Bob Jones Univ. v. United States*, 103 S. Ct. 2017 (1983).

31. The Court also ruled against Bob Jones University on its claim that the first amendment's free-exercise-of-religion clause blocked the IRS from implementing its racial nondiscrimination policy against the University. The Court held that "the Government has a fundamental overriding interest in eradicating racial discrimination in education That governmental interest substantially outweighs whatever burden denial of tax benefits places on [the University's] exercise of [its] religious beliefs." *Id.* at 2035. For a discussion of the religious liberty issue raised in *Bob Jones University*, see Laycock, *Tax Exemptions for Racially Discriminatory Religious Schools*, 60 TEX. L. REV. 259 (1982); Neuberger & Crnmpelar, *Tax Exempt Religious Schools Under Attack: Conflicting Goals of Religious Freedom and Racial Integration*, 48 FORDHAM L. REV. 229, 258-75 (1979); Note, *The Judicial Role in Attacking Racial Discrimination in Tax-Exempt Private Schools*, 93 HARV. L. REV. 378, 401-04 (1979). For a critique of the Court's ruling on the religious liberty issue, see Devins, *Did the High Court Go Too Far to Make a Politically Popular Ruling?*, NAT'L L.J., June 20, 1983, at 13, col. 1.

32. 103 S. Ct. at 2028-29. The Court did not reach the issue of whether an organization could receive tax-exempt status if it violated some public policy but also conferred some public benefit. See *infra* notes 111-13 and accompanying text.

33. I.R.C. § 501(c)(3) (1976); see also *id.* § 501(a) (providing that an organization de-

Applying this section, the IRS had ruled that to qualify for tax exemption an institution must demonstrate that it falls within one of the categories defined in that section and that its activity is not contrary to settled public policy.³⁴ The IRS felt that the settled public policy was that the section 501(c)(3) tax exempt organization—even if not qualifying as a “charitable” organization—must be charitable in the common law sense. The common law notion of charity includes an effort to further a public purpose.³⁵ The University contested the Service’s construction of the tax exemption provision, arguing that the position of the IRS contradicted the statute’s plain language and legislative history and that the IRS had adopted the faulty reasoning of the *Green* court.³⁶

In its “plain language” argument the University emphasized the absence of any statutory language that expressly required all

scribed in § 501(c) shall be exempt from taxation unless such exemption is denied under § 502 or § 503). The complete text of § 501(c)(3), specifying those organizations qualifying for tax exemption, follows:

(3) Corporations, and any community chest, fund, or foundation, organized and operated exclusively for religious, charitable, scientific, testing for public safety, literary, or educational purposes, or to foster national or international amateur sports competition (but only if no part of its activities involve the provision of athletic facilities or equipment), or for the prevention of cruelty to children or animals, no part of the net earnings of which inures to the benefit of any private shareholder or individual, no substantial part of the activities of which is carrying on propaganda, or otherwise attempting, to influence legislation (except as otherwise provided in subsection (h)), and which does not participate in, or intervene in (including the publishing or distributing of statements), any political campaign on behalf of any candidate for public office.

Subchapter F of the Internal Revenue Code, §§ 501-28, is an elaborately detailed set of statutory rules providing tax exemptions for various qualifying organizations. Section 501(a) provides generally for exemption from tax for those organizations described in subsections (c) and (d). Subsection (c) lists 22 organizations that enjoy tax exemption, and subsection (d) describes certain religious or apostolic organizations. In particular, § 501(c)(3) describes certain types of organizations that not only enjoy tax-exempt status but also afford generally deductible contributions under § 170. See I.R.C. § 170(c)(2) (1976 & Supp. V 1981).

34. *Bob Jones Univ. v. United States*, 103 S. Ct. at 2025; see Brief of Amicus Curiae William T. Coleman, Jr. at 11-44, *Bob Jones Univ.*

35. 103 S. Ct. at 2025-26.

36. Brief for Petitioner at 11, *Bob Jones Univ.* For a discussion of *Green*, see *infra* notes 55-56 and accompanying text.

Goldsboro Christian Schools, Inc., concerned similar arguments. *Goldsboro* is a private elementary and secondary school that denied admissions to blacks, supposedly for religious reasons. *Goldsboro* challenged the denial of its tax-exempt status in much the same way that *Bob Jones University* had done. The two cases were joined before the Supreme Court. 454 U.S. 892 (1981). Because the statutory and tax policy issues raised in *Goldsboro* are identical to those issues in *Bob Jones University*, the two cases can be treated as identical for the purposes of this Article. The two cases, however, do differ on the religious liberty issue. The government has a stronger interest in preventing a school from denying admissions to blacks than in limiting a school’s interracial social relationships. See Weeks & Devins, *First Amendment Free Exercise Protections*, 6 LEX. COLLEGE 1 (Summer 1982).

exempt organizations to be "charitable" in the common law sense of providing some public benefit. The University maintained that the disjunctive "or" separating the categories in section 501(c)(3) precluded this reading of the statute and that any organization falling within the specified categories automatically qualified for an exemption.³⁷ This "plain language" argument relied on decisions in which the Court refused to transfer the meaning of one statutory term to another employed in the disjunctive.³⁸ Similarly, the University condemned the IRS's attempt "to make 'religious' an adjective modifying 'charitable'" as equally untenable.³⁹

The majority, however, rejected the "plain language" argument:

It is a well-established canon of statutory construction that a court should go beyond the literal language of a statute if reliance on that language would defeat the plain purpose of the statute . . . "[I]n interpreting a statute, the court will not look merely to a particular clause in which general words may be used, *but will take in connection with it the whole statute* . . .

37. 103 S. Ct. at 2025; *see also* Brief for Petitioner at 10-23, *Bob Jones Univ.*

38. *See, e.g., Reiter v. Sonotone Corp.*, 442 U.S. 330 (1979). Defendants in *Reiter* contended that Clayton Act coverage of "[a]ny person who shall be injured in his business or property" should extend only to "business activity or property related to one's business." *Id.* at 337-38. The Court, however, disagreed:

That strained construction would have us ignore the disjunctive "or" and rob the term "property" of its independent and ordinary significance; moreover, it would convert the noun "business" into an adjective. In construing a statute we are obliged to give effect, if possible, to every word Congress used. . . . Canons of construction ordinarily suggest that terms connected by a disjunctive be given separate meanings, unless the context dictates otherwise; here it does not. . . . Congress' use of the word "or" makes plain that "business" was not intended to modify "property," nor was "property" intended to modify "business".

Id. at 338-39 (citations omitted).

39. Brief for Petitioner at 12, *Bob Jones Univ.* The University also contended that "where substantial constitutional issues under the Religion Clauses would arise by virtue of the extension to religious institutions of a governmental requirement, this Court has held that the extension may not be left to implication, but instead 'there must be present the affirmative intention of the Congress clearly expressed.'" *Id.* (relying upon *NLRB v. Catholic Bishop of Chicago*, 440 U.S. 490, 500 (1979)). This contention is without merit. First, the *Catholic Bishop* decision, upon which the University based this argument, presented the Court with an opportunity to vindicate the Bishop's position on either statutory or constitutional grounds. *See Catholic Bishop of Chicago v. NLRB*, 559 F.2d 1112 (7th Cir. 1977) (NLRB cannot order representation elections for lay teachers in Catholic high schools on both statutory and first amendment grounds). *Bob Jones University* did not present the Court with such a choice because the Court viewed the school's religious liberty claim as subsidiary to the government's interest in racial nondiscrimination. Second, the Court in *Catholic Bishop* recognized that "the amendment may not be substituted for construction and that a court may not exercise legislative functions to save the law from conflict with constitutional limitation." *Yu Cong Eng. v. Trinidad*, 271 U.S. 500, 518 (1926). In short, the Supreme Court in *Bob Jones University* did not address the *Catholic Bishop* issue.

and the objects and policy of the law”⁴⁰

The majority adopted the view that the common law of charitable trusts had guided the enactment of section 501(c)(3)⁴¹ and that Congress had expressly adopted the common law’s public benefit rationale for charitable exemptions:

“The exemption from taxation of money and property devoted to charitable and other purposes is based on the theory that the Government is compensated for the loss of revenue by its relief from financial burdens which would otherwise have to be met by appropriations from other public funds, and by the benefits resulting from the promotion of the general welfare.”⁴²

In unusually sweeping language, the majority then articulated its standards for exemptions:

Charitable exemptions are justified on the basis that the exempt entity confers a *public benefit*—a benefit which the society or the community may not itself choose or be able to provide, or which supplements and advances the work of public institutions already supported by tax revenues. . . . The institution’s purpose must not be so at odds with the *common community conscience* as to undermine any public benefit that might otherwise be conferred.⁴³

The University was not entitled to tax-exempt status under these standards because “an educational institution engaging in practices affirmatively at odds with [the government’s] declared position [on racial nondiscrimination] . . . cannot be seen as exercising a ‘beneficial and stabilizing influence in community life.’”⁴⁴ Further, “In determining what purposes may benefit the community and what organizations are therefore exempt, public benefit must be measured by present social and governmental legal and moral standards, rather than those in existence at the time section 501 was enacted.”⁴⁵ Under such contemporary standards⁴⁶ the Court recognized that the “‘*legitimate educational function* [of a racially discriminatory private school] *cannot be isolated from discriminatory practices. . . . [D]iscriminatory treatment exerts a pervasive influence on the entire educational process.*’”⁴⁷ The

40. 103 S. Ct. at 2025-26 (quoting *Brown v. Duchesne*, 60 U.S. (19 How.) 183, 194 (1857)) (emphasis supplied by the Court).

41. 103 S. Ct. at 2026 (citing Simon, *The Tax-Exempt Status of Racially Discriminatory Religious Schools*, 36 TAX L. REV. 477, 485-99 (1981)).

42. 103 S. Ct. at 2028 (quoting H.R. REP. NO. 1860, 75th Cong., 3d Sess. 19 (1938)).

43. 103 S. Ct. at 2028-29 (emphasis supplied).

44. *Id.* at 2032 (quoting *Walz v. Tax Com’n*, 397 U.S. 664, 673 (1970)).

45. Simon, *supra* note 41, at 488; see 103 S. Ct. at 2029 n.20.

46. Racial discrimination in education officially did not become a public wrong until the Supreme Court’s 1954 decision in *Brown v. Board of Education*, 347 U.S. 483 (1954).

47. 103 S. Ct. at 2030 (quoting *Norwood v. Harrison*, 413 U.S. 455, 468-69 (1973))

majority thus concluded that “[i]t would be wholly incompatible with the concepts underlying tax exemption to grant the benefit of tax-exempt status to racially discriminatory educational entities.”⁴⁸

Justice Rehnquist, the sole dissenter, endorsed the University’s argument that the legislative history militated against a finding that section 501(c)(3) required common law charitability. His dissent traced the evolution of the tax exemption provision and rejected finding “some additional, undefined public policy requirement.”⁴⁹ He concluded “that the legislative history of § 501(c)(3) unmistakably makes clear that *Congress has decided* what organizations are serving a public purpose and providing a public benefit within the meaning of § 501(c)(3) and has clearly set forth in § 501(c)(3) the characteristics of such organizations.”⁵⁰ Nonetheless, Justice Rehnquist could not persuade the majority of the absence of the common law notion of charity in the section.

Justice Powell, concurring, avoided making broad pronouncements on the meaning and purpose of the tax exemption provision. Instead, he confined his analysis to the narrow issue of whether “there are now sufficient reasons for accepting the IRS’s construction of the Code as proscribing tax exemptions for schools that discriminate on the basis of race as a matter of policy.”⁵¹ In trying to discern the legislative intent behind the provision, Justice Powell attributed great weight to the refusal of Congress to act on proposals that would have overturned the IRS’s nondiscrimination policy⁵² and to the amendment of the Code that denies tax exemp-

(emphasis supplied by the Court).

48. 103 S. Ct. at 2030.

49. 103 S. Ct. at 2040 (Rehnquist, J., dissenting).

50. *Id.* at 2041 (Rehnquist, J., dissenting) (emphasis supplied by the dissent). Justice Powell’s concurring opinion also was sympathetic to this reading:

It also is clear that the language [of § 501(c)(3)] itself does not mandate refusal of tax-exempt status to any private school that maintains a racially discriminatory admissions policy. Accordingly, there is force in Justice Rehnquist’s argument Indeed, were we writing prior to the history detailed in the Court’s opinion, this could well be the construction that I would adopt.

Id. at 2036 (Powell, J., concurring).

51. *Id.* at 2036 (Powell, J., concurring).

52. As the majority noted, “During the past 12 years there have been no fewer than 13 bills introduced to overturn the IRS interpretation of § 501(c)(3). Not one of these bills has emerged from any committee, although Congress had enacted numerous other amendments to § 501 during the same period” *Id.* at 2033. Justice Rehnquist, however, thought this evidence irrelevant: “[W]e have said before, and it is equally applicable here, that this type of congressional inaction is of virtually no weight in determining legislative intent. . . . These bills and related hearings indicate little more than that a vigorous debate has existed in Congress concerning the new IRS position.” *Id.* at 2043 (Rehnquist, J., dissenting) (cita-

tions to racially discriminatory private clubs.⁵³ Consequently, although he read the disjunctive "or" in the statute to indicate that Congress did not intend a common law "charitable" gloss to apply to each category, he concurred with the majority's result because "there has been a decade of acceptance [by Congress] that is persuasive in the circumstances of this case."⁵⁴

The majority did not address directly the University's argument that *Green* was a bad decision.⁵⁵ Although the Court reached its decision on the "public benefit" theory, it did suggest that the *Green* "public policy" doctrine was good law.⁵⁶

tions omitted). As it relates to the relevance of proposed legislation, this position does not seem unreasonable. Still, these affirmative acts by Congress indicate that it both recognized and supported the racial nondiscrimination requirement. See *infra* note 53.

53. See 26 U.S.C. § 501(c) (1976). Congress amended the Code in response to the District Court for the District of Columbia's decision in *McGlotten v. Connally*, which held, in part, that nonprofit private clubs that excluded nonwhites from membership were entitled to tax-exempt status. 338 F. Supp. 448, 457-59 (D.D.C. 1972). This legislation indicates that Congress supports nondiscrimination as a social policy, as the Senate Committee Report on the amendment illustrates: "[I]t is believed that it is inappropriate for a social club . . . to be exempt from taxation if its written policy is to discriminate on account of race, color or religion." S. REP. NO. 1318, 94th Cong., 2d Sess. 8, reprinted in 1976 U.S. CODE CONG. & AD. NEWS 6051, 6058. Further, as the majority in *Bob Jones University* noted, Congress had demonstrated its approval of racial nondiscrimination for tax-exempt private schools: "Congress' awareness of the denial of tax-exempt status for racially discriminatory schools when enacting other and related legislation make out an unusually strong case of legislative acquiescence in and ratification by implication of the 1970 and 1971 rulings [including Rev. Rul. 41-477]." 103 S. Ct. at 2033. For a general discussion of Congress' recognition of the nondiscrimination requirement, see Devins, *supra* note 23, at 161-63. See also *Haig v. Agee*, 453 U.S. 280, 300 (1981) (Congress' failure to change an agency ruling is an implicit acceptance of that ruling).

54. 103 S. Ct. at 2036 (Powell, J., concurring).

55. The University had contended that *Green* represented "an elaborate, but insupportable, effort to write a provision into the Internal Revenue Code which the Congress did not write and did not imply." Brief for Petitioner, *supra* note 36, at 17-18. Although *Green* discussed the public benefit theory, that decision's primary basis was the doctrine that the tax exemption provision must be construed to avoid frustrations of public policy. See *Green v. Connally*, 330 F. Supp. 1150, 1161 (D.D.C. 1971); Note, *The Revocation of Tax Exemptions and Tax Deductions for Donations to 501(c)(3) Organizations on Statutory and Constitutional Grounds*, 30 U.C.L.A. L. REV. 156, 160, 164 (1982).

56. "A corollary to the public benefit principle is the requirement, long recognized in the law of the trusts, that the purpose of a charitable trust may not be illegal or violate established public policy." 103 S. Ct. at 2028.

The *Green* court's "public policy" rule relied on cases concerned with the deduction of "ordinary and necessary" business expenses under § 162 of the Code, particularly on *Tank Truck Rentals, Inc. v. Commissioner*, 356 U.S. 30 (1958). In *Tank Truck* the Court disallowed the deduction of fines paid by truck owners who had violated the state's maximum weight laws. The Court declared that it would deny deductions that would "frustrate sharply defined national or state policies proscribing particular types of conduct, evidenced by some governmental declaration thereof." 356 U.S. at 33-34. *Tank Truck*, however, may be of limited value in the tax exemption context. Its holding applies only to situations "in

1. Analysis of the Public Benefit Doctrine

The legislative histories of tax laws throughout this century support the majority's holding that tax exempt organizations must provide some public benefit. In the floor debate over the Tariff Act of 1894,⁵⁷ which provided tax exemptions for organizations "organized and conducted solely for charitable, religious, or educational purposes,"⁵⁸ Congress made clear that these tax benefits were available because the organizations served desirable public purposes.⁵⁹ The legislative histories of subsequent taxing acts have reaffirmed this rationale,⁶⁰ as the following excerpt from a 1938 Congressional report illustrates.

The exemption from taxation of money or property devoted to charitable and other purposes is based upon the theory that the Government is compensated for the loss of revenue by its relief from financial burdens which would other-

which an allowance of the deductions would amount to 'a device to avoid the consequence of violations of a law.'" Simon, *supra* note 41, at 497 (quoting *Commissioner v. Sullivan*, 356 U.S. 28, 29 (1958)). See also Brief for Petitioner, *supra* note 36, at 19.

In contrast, although tax-exempt status may be important to an organization's very existence, "to the extent that the organization's alleged public policy violation violates a federal or state statute, the granting of the exemption does not mitigate the consequence of the violation." Note, *supra* note 55, at 168. See also Brief for Petitioner, *supra* note 36, at 19. But see Simon, *supra* note 41, at 497-500 (suggesting that the "public policy" doctrine applies to tax-exempt private schools, despite the *Tank Truck* limitation). Moreover, what would happen if an organization partially deviated from public policy? "[I]f the *Green* [public policy] rationale is accepted, that if any organization, otherwise exempt under § 501(c)(3), were to discriminate on account of age, maintain unsafe or unhealthful working conditions, create any financial barrier to education based on sex, or create any environmental disharmony, that organization's tax exemption would have to be denied." Brief for Petitioner, *supra* note 36, at 20 (footnote omitted). See also Neuberger & Crumplar, *supra* note 31, at 272-73. Cf. *Tax-Exempt Status of Private Schools: Hearings Before the Subcomm. on Oversight of the House Comm. on Ways and Means*, 96th Cong., 1st Sess. 1229 (1979) (statement of Charles A. Bane, Co-Chairman, Lawyer's Comm. for Civil Rights Under Law [hereinafter cited as *Hearings*]); *id.* at 470 (statement of Bill Lann Lee, Assistant Counsel, NAACP Legal Defense and Educational Fund).

The *Green* court incorrectly relied on the business deduction cases to support the proposition that tax-exempt organizations must conform with public policy before individuals can take a deduction for their contributions to these organizations. These cases are relevant only to the issue of defining "ordinary and necessary" expenses for purposes of determining taxable income. It is pure conjecture to suggest that judicial rulings on what is a permissible deduction for purposes of determining taxable income carry over to the charitable deduction issue. Even more egregious than this reasoning by conjecture, the income tax deduction cases do not support the *Green* public policy formulation. For a general discussion of the *Green* ruling, see also McCoy & Devins, *Standing and Adverseness on the Issue of Tax Exemptions for Discriminatory Private Schools*, 52 *FORDHAM L. REV.* ____ (1984) (forthcoming).

57. Ch. 349, 28 Stat. 509 (1894).

58. *Id.* at 556.

59. 26 CONG. REC. 585-86 (1894).

60. See, e.g., 50 CONG. REC. 1305 (1913); 44 CONG. REC. 4147, 4150 (1909).

wise have to be met by appropriations from public funds, and by the benefits resulting from the promotion of the general welfare.⁶¹

The same rationale—the service of desirable public purposes—underlies the allowance of deductions for contributions to charitable organizations. When Congress in 1917 enacted a companion provision to section 501(c)(3) providing for these deductions (under what is now section 170(c)(2)), the debate emphasized the requirement of a public benefit: “For every dollar that a man contributes for these public charities, educational, scientific, or otherwise, the public gets 100 per cent.”⁶² In a similar vein, the legislative history included the following:

[The charitable deduction] would remove the absurdity of exacting a tax even on that share of a man's income which he devotes not at all to himself, but to the pressing needs of educational and charitable institutions which operate without private profit. The exaction of such a tax is, at this time, worse than an absurdity. . . . It passes beyond individuals and strikes at America's whole organization for social progress and education, the relief of distress, and the remedy of evils.⁶³

Relying on these legislative histories, the majority correctly concluded that the IRS acted properly under its rulemaking authority when it determined in 1971 to deny tax-exempt status unless an organization both fell within one of the categories described in section 501(c)(3) and did not engage in activities contrary to settled public policy.⁶⁴

2. Analysis of the Common Community Conscience Requirement

Although a tax exemption standard properly may require that an organization have a public purpose or confer a public benefit, the majority's requirement of conformity to the common community conscience goes too far. Even more significantly, this requirement could stifle the development of new ideas.⁶⁵ A common com-

61. HOUSE COMM. ON WAYS AND MEANS, THE REVENUE BILL OF 1938, H.R. REP. NO. 1860, 75th Cong., 3d Sess. 19 (1938). For an excellent early analysis, see Reiling, *Federal Taxation: What is a Charitable Organization?*, 44 A.B.A. J. 525 (1958) (in accord with *Bob Jones University*).

62. 55 CONG. REC. 6714, 6728 (1917) (statement of Senator Hollis).

63. *Id.* at 6729 (reprinting “Do Not Penalize Generosity,” Boston Transcript, June 29, 1917) (emphasis supplied).

64. Rev. Rul. 71-447, 1971-2 C.B. 230.

65. A good example is private schools. Private schools are often a desirable educational alternative precisely because they are free of many of the governmental constraints on public schools. Private schools can impart values, teach religion, enforce different disciplinary standards, select and dismiss teachers, and insist on sustained academic achievement

munity conscience may reflect an ever-changing set of values that all exempt organizations continually and painstakingly would have to satisfy. As Justice Powell argued in his concurrence, the *Bob Jones* majority "ignores the important role played by tax exemptions in encouraging diverse, indeed often sharply conflicting, activities and viewpoints."⁶⁶ Similarly, Justice Brennan observed in a case upholding the constitutionality of property tax exemptions for religious organizations that "private, nonprofit groups . . . receive tax exemptions . . . [because] each group contributes to the diversity of association, viewpoint, and enterprise essential to a vigorous, pluralistic society."⁶⁷ In short, the majority in *Bob Jones University* ignored the public benefits of a heterogeneous society.

The Court could have achieved its desired result by using the public benefit test alone, without the common community conscience element. An organization's racially discriminatory practices may be so odious and contrary to the fundamental value of equal treatment under the law that the organization could never assert convincingly that it serves any public purpose. Regrettably, however, the vagueness of both the common community conscience and the public benefit standards creates the danger that the IRS may overzealously enforce the standards, resulting in unwanted social homogeneity.

in ways that public schools cannot. This freedom is the essence of their appeal. See Finn & Devins, *Reagan, Discrimination, and Private Schools*, Wall St. J., Feb. 2, 1982, at 30, col. 3; Finn, *Public Support for Private Education*, pt. 1, AM. EDUC. (May 1982). Tax exemptions are essential to ensuring this diversity and autonomy. Private schools derive 23% of their revenues from their tax-exempt status or the related charitable deduction. See *Hearings*, *supra* note 56, at 400 (testimony of John Esty, Jr., President, National Association of Independent Schools).

Mr. Justice Powell pointed out in his concurring opinion that over 106,000 organizations filed § 501(c)(3) returns in 1981. He found "it impossible to believe that all or even most of those organizations could prove that they 'demonstrably serve and [are] in harmony with the public interest' or that they are 'beneficial and stabilizing influences in community life.'" 103 S. Ct. at 2038.

66. 103 S. Ct. at 2038.

67. *Walz v. Tax Comm'n*, 397 U.S. 664, 689 (1970) (Brennan, J., concurring). Moreover, the majority in *Walz* noted that "the use of a social welfare yardstick as a significant element to qualify for tax exemption could conceivably give rise to confrontations that could escalate to constitutional dimensions." *Id.* at 674. For a general discussion of constitutional limitations on government in the promotion of behavior or ideology, see Kamenshine, *The First Amendment's Implied Political Establishment Clause*, 67 CALIF. L. REV. 1104 (1979).

If the function of tax exemptions in encouraging diversity and conflicting views celebrated by Mr. Justice Brennan in *Walz* now gives way to one of ensuring harmony with the public interest and community conscience, then only a narrower range of organizations likely will qualify for tax exemptions.

3. Analysis of the Tax Exemption and the "Tax Expenditure" Budget

The *Bob Jones University* decision revived the unsettled debate over whether tax exemptions are really a form of government aid.⁶⁸ Because the sixteenth amendment authorizes the taxation of all income "from whatever source derived,"⁶⁹ Congress could have constructed a comprehensive tax base,⁷⁰ requiring organizations that presently are exempt to conform to the same rules as taxable entities. Instead, Congress from the outset has chosen not to tax all possible entities but rather to tax selectively, often in pursuit of various social objectives.

Not surprisingly, commentators have disagreed on the use of exemptions, deductions, and credits to accomplish social goals. According to Professor Stanley Surrey,⁷¹ numerous tax incentive provisions have evolved in the Internal Revenue Code to assist particular industries, business activities, and financial transactions, or to encourage certain social activities, such as contributions to charity. He contends that a tax incentive is a cost to the government of the tax revenue that it would have collected if the law did not provide for that particular deduction, exemption, or credit.⁷² Therefore, a tax incentive is an indirect *expenditure* of government funds to support the particular purpose behind the incentive.⁷³ Professor

68. See *infra* notes 96-110 and accompanying text.

69. U.S. CONST. amend. XVI. A comprehensive tax base would conform to the classical definition of income as the increase in net market value of assets between the beginning and end of the taxable period, plus the market value of consumption (personal and living outlays) during the period, including gifts. For corporations with no consumption expenditures, the measure would be the net accretion in asset values between the beginning and end of the taxable period, without regard to distributions to shareholders.

70. HAIG, *THE CONCEPT OF INCOME—ECONOMIC AND LEGAL ASPECTS IN THE FEDERAL INCOME TAX* (1921); H. SIMONS, *PERSONAL INCOME TAXATION* 61-62 (1938).

71. See Surrey, *Our Troubled Tax Policy: False Routes and Proper Paths to Change*, 12 *TAX NOTES* 179 (1981); Surrey, *Tax Incentives as a Device for Implementing Government Policy: A Comparison with Direct Government Expenditures*, 83 *HARV. L. REV.* 705 (1970) [hereinafter cited as Surrey, *Tax Incentives*].

72. For the fiscal year 1983, the Office of Management and Budget estimates that the deduction for charitable contributions to education results in a revenue loss to the Treasury of \$925 million and that the "outlay equivalent," which would be the amount required to be spent by government to accomplish the same objective, would be \$940 million. OFFICE OF MANAGEMENT AND BUDGET, *SPECIAL ANALYSIS G* 29, 35 (1983).

73. The following equation illustrates Surrey's position: if X represents what Congress could tax as income, and if Y represents what in fact Congress taxes after allowing for exclusions, exemptions, deductions and credits, then X-Y is the aggregate cost of these special provisions, or the indirect expenditure by the Government attributable to these provisions. The tax expenditure budget breaks out this aggregate cost into the cost of each exclusion, exemption, deduction, or credit.

Surrey argues that the government could administer these incentives or subsidies more efficiently through direct governmental assistance in the form of grants, loans, interest subsidies, loan guarantees, and the like.⁷⁴ Moreover, direct governmental assistance would relieve the present inequity of forcing taxpayers who do not benefit from these tax incentives to bear a greater share of the tax burden.⁷⁵

On the other hand, Professor Boris Bittker argues that exemptions, and credits are *not* necessarily costs to the government. Instead, they reflect a legislative choice to omit certain transactions, entities, or activities from the tax base.⁷⁶

There is no way to tax *everything*; a legislative body, no matter how avid for revenue, can do no more than pick out from the universe of people, entities, and events over which it has jurisdiction those that, in its view, are appropriate objects of taxation. In specifying the ambit of any tax, the legislature cannot avoid "exempting" those persons, events, activities, or entities that are outside the territory of the proposed tax. In describing a tax's boundaries, the draftsman may choose to make the exclusions explicit ("all property except that owned by nonprofit organizations"), or implicit ("all property owned by organizations operated for profit"), but either way, the result is the

Tax expenditures increased from 24.8% of federal revenues in 1971 to 40.8% in 1982. See CONGRESSIONAL BUDGET OFFICE, TAX EXPENDITURES: BUDGET CONTROL OPTIONS AND FIVE-YEAR BUDGET PROJECTIONS FOR FISCAL YEARS 1983-1987 12-13 (1982). In the 1984 budget the estimated total is \$388.4 billion. See PECHMAN, SETTING NATIONAL PRIORITIES, THE 1984 BUDGET 178 (Brookings Institution 1983). Whatever one's position on tax expenditures, the enormity of the amount must be a factor in any discussion of tax policy. For a discussion of the "tax exemption as aid" issue, see *infra* notes 96-110 and accompanying text.

74. [T]he deduction for charitable contributions is sometimes cited as a method of government assistance that promotes private decisionmaking—the taxpayer, and not the Government, selects the charity and determines how much to give. But a direct expenditure program under which the Government matched with its grants, on a no-question-asked and no-second-thoughts basis, the gifts of private individuals to the charities they selected, would equally preserve private decisionmaking.

Surrey, *Tax Incentives*, *supra* note 71, at 719.

75. See *id.* at 713-38.

In other words, Professor Surrey maintains that if the government taxed all income at a given percentage, a taxpayer would retain his income less the tax. But if the government selectively taxes only a part of a taxpayer's income, then it affords him preferential treatment over other taxpayers who have income not so preferred. This result violates what economists call vertical and horizontal equity: all those with different levels of real economic income should pay proportionately different taxes (vertical equity); all those with the same levels of income should pay the same tax (horizontal equity). The Tax Expenditure Budget proves that our present tax system has neither.

76. Bittker, *Income Tax "Loopholes" and Political Rhetoric*, 71 MICH. L. REV. 1099, 1102-28 (1973); see also Bittker, *The Tax Expenditure Budget—A Reply to Professors Surrey & Hellmuth*, 22 NAT'L TAX J. 538, 539 (1969); Bittker, *A "Comprehensive Tax Base" as a Goal of Income Tax Reform*, 80 HARV. L. REV. 925, 934-58 (1967).

same—taxpayers are separated from non-taxpayers.⁷⁷

Irving Kristol has criticized the tax expenditure concept of exemptions for similar reasons:

You are implicitly asserting that all income covered by the general provisions of the tax laws belongs of right to the government, and that what the government decides, by exemptions or qualification, not to collect in taxes constitutes a subsidy. Whereas a subsidy used to mean a governmental expenditure for a certain purpose, it now acquires quite another meaning—i.e., a generous decision by government not to take your money.⁷⁸

Although the *Bob Jones University* Court did not explicitly adopt Surrey's view that a tax exemption is government aid, its narrow "common community conscience" standard could portend a tax policy that would grant exemptions only to those organizations that pander to community or majority sentiment. This system for awarding tax exemptions would be much like the present system for awarding federal grants, in that both would condition governmental benefits on community assent. This system, however, could stray far from past policies that have encouraged the diverse and conflicting views of a pluralistic society because it would inhibit thousands of presently exempt entities from venturing into ideologically rough waters.

In sum, the majority opinion on tax exemptions—even though the Court seems to favor Surrey's viewpoint—falls into a penumbra between the Surrey and Bittker positions. If the Court had pursued its inclination and held that the tax exemption was government aid, then the IRS could develop a series of regulations, similar to the system of rules associated with a government subsidy program, to assure consistency and parity in the benefits granted. By refusing to hold that a tax exemption is aid—a holding that would have forced tax-exempt institutions to comply with a host of government regulations associated with governmental subsidy programs⁷⁹—while at the same time recognizing broad IRS authority to develop rules governing tax-exempt status—an approach that might result in the granting of tax exemptions on a toothless pro forma basis⁸⁰—the *Bob Jones University* Court established a

77. Bittker, *Churches, Taxes and the Constitution*, 78 YALE L.J. 1285, 1288 (1969).

78. Kristol, *Taxes, Property and Equality*, 37 PUB. INT. 3, 14-15 (1974). The proponents of the Tax Expenditure Budget would contend that income may be defined as annual accretions (or decrections) in wealth, plus consumption. To whatever extent this base is reduced by a deduction, exclusion, or credit, this amount becomes an indirect expenditure. See *supra* note 71. See also Galvin, *It's VAT Time Again*, 21 TAX NOTES 275 (1983).

79. See *infra* text accompanying notes 96-110.

80. For example, current IRS enforcement procedures may not effectuate the goals

rulemaking model that is inconsistent with its interpretation of the tax exemption provision. In short, the majority failed to ensure the enforcement of its view that tax-exempt institutions must reflect community values because it failed to hold that tax exemptions are government aid or that the IRS must enforce the public policy requirement strictly.

B. *The Scope of IRS Rulemaking Authority*

The *Bob Jones University* Court recognized broad IRS authority to determine what activities are “at odds with the common community conscience.”⁸¹ The majority noted that “ever since the inception of the tax code, Congress has seen fit to vest in those administering the tax laws very broad authority to interpret those laws.”⁸² The Court thus rejected petitioners’ argument that IRS rulemaking on tax exemptions is “a plain usurpation of Congressional law-making powers by the non-elected public servants of the Internal Revenue Service.”⁸³

An analysis of the proper scope of the Service’s rulemaking authority must begin with the principle that Congress enacts the tax laws and the IRS has the responsibility of interpreting and enforcing them. The IRS cannot legislate.⁸⁴ The issue, then, is to determine the permissible boundaries of IRS rulemaking. In *Manhattan General Equipment Co. v. Commissioner*⁸⁵ the Court described this power:

The power of an administrative officer or board to administer a federal statute and to prescribe rules and regulations to that end is not the power to make law—for no such power can be delegated by Congress—but the power to adopt regulations to carry into effect the will of Congress as expressed by

underlying the *Bob Jones University* decision because some schools that have been adjudicated as racially discriminatory under the 1964 Civil Rights Act have received tax exemptions anyway. See Devins, *The Bob Jones Case—Over to Congress*, Christian Sci. Monitor, June 29, 1983, at 23, col. 1.

81. 103 S. Ct. at 2029.

82. *Id.* at 2031.

83. Brief for Petitioner, *supra* note 36, at 22.

84. Congress has charged the Internal Revenue Service with the administration of the tax laws. Because the language of this legislation is general, the Service issues regulations and rulings to implement and explain its position on the law. Under the Administrative Procedure Act, the Treasury gives notice of proposed rulemaking and publishes proposed regulations in the Federal Register. 5 U.S.C. § 553 (1982); Schmid, *The Tax Regulations Making Process—Then and Now*, 24 TAX LAW. 541, 541-42 (1971). After a period of receiving written comments, the Treasury then promulgates the final regulations. The Service issues revenue rulings, which usually deal with particular transactions or problems. See generally Note, *Federal Tax Rulings: Procedure and Policy*, 21 VAND. L. REV. 78 (1967).

85. 297 U.S. 129 (1936).

the statute. . . . The statute defines the rights of the taxpayer and fixes a standard by which such rights are to be measured.⁸⁶

In *Bob Jones University* the Court granted the Service almost plenary rulemaking authority:

In the first instance . . . the responsibility for construing the Code falls to the IRS. Since Congress cannot be expected to anticipate every conceivable problem that can arise or to carry out day-to-day oversight, it relies on the administrators and on the courts to implement the legislative will. Administrators, like judges, are under oath to do so.⁸⁷

The majority's interpretation, however, poses several problems. First, the danger exists that the Service may selectively enforce its regulations.⁸⁸ Justice Blackmun commented on this threat as follows:

[W]here the philanthropic organization is concerned, there appears to be little to circumscribe the almost unfettered power of the Commissioner. This may be very well so long as one subscribes to the particular brand of social policy the Commissioner happens to be advocating at the time . . . , but application of our tax laws should not operate in so fickle a fashion.⁸⁹

86. *Id.* at 134-35. See also *Mobil Oil Corp. v. Higginbotham*, 436 U.S. 618, 625 (1978) ("There is a basic difference between filling a gap left by Congress' silence and rewriting rules that Congress has affirmatively and specifically enacted.")

87. 103 S. Ct. at 2031. Curiously, in the first of the tax exemption cases, *Green v. Kennedy*, the district court suggested that the IRS was blameless for not having changed the policy of granting tax exemptions to discriminatory private schools. The court believed that "[w]hat stops [the Commissioner of Internal Revenue] from extending disallowance to the schools . . . is not unawareness of the significance of deductions, but rather certain legal conclusions, including conclusions as to the scope of his authority under the Code." 309 F. Supp. 1127, 1135 (D.D.C. 1970). See *supra* notes 20, 55-56. The court apparently felt that the Commissioner should act only on an explicit congressional directive or a binding court determination. In other words, the court envisioned a scheme in which the judiciary—not the Service—would have primary authority in interpreting the meaning of the congressionally enacted Internal Revenue Code. Ironically, in *Bob Jones University* the Supreme Court recognized the IRS's authority to promulgate regulations established through judicial initiatives directed at the IRS. For a general discussion of the judiciary's usurpation of legislative authority on the tax-exempt school issue, see *McCoy & Devins*, *supra* note 56.

88. See Note, *supra* note 55, at 172-74.

89. *Commissioner v. "Americans United" Inc.*, 416 U.S. 752, 774-75 (1974) (Blackmun, J., dissenting). The D.C. Circuit echoed Justice Blackmun's concerns in *Big Mama Rag, Inc. v. United States*, 631 F.2d 1030, 1040 (D.C. Cir. 1980): "The standards [used by the IRS to grant or deny tax exemptions] may not be so imprecise that they afford latitude to individual IRS officials to pass judgment on the content and quality of an applicant's views and goals . . ." Under this standard, the court held that the definition of "educational" contained in the IRS regulations under § 501(c)(3) was unconstitutionally vague. *Id.* at 1037. See also *Center on Corp. Responsibility, Inc. v. Shultz*, 368 F. Supp. 863 (D.D.C. 1973) (politically motivated denial of educational exemption by IRS is null and void); Comment, *Tax Exemptions for Educational Institutions: Discretion and Discrimination*, 128 U. PA. L. REV. 849 (1980). *But cf.* *National Alliance v. United States*, 710 F.2d 868 (D.C. Cir. 1983) (neither § 501(c)(3) nor the first amendment compels granting educational exemption to organization whose publications could not be found "educational" under any reasonable in-

Nonetheless, the *Bob Jones University* majority believed that the Service could be trusted not to breach this authority.⁹⁰ The Court sought to minimize the danger of selective enforcement by stressing that “these sensitive determinations should be made only where there is no doubt that the organization’s activities violate fundamental public policy.”⁹¹

Second, the IRS may go either too far or not far enough in regulating the wide array of tax-exempt organizations. The Service now may examine the tax-exempt status of many organizations against the stricter *Bob Jones* standard of “harmony with the public interest” and “common community conscience.” This prospect is disquieting because the standard, although strict, is open-ended and beclouded. For example, the fate of an organization that does not violate *fundamental* public policy but may not be clearly in full compliance with the *Bob Jones University* standard is unclear. Justice Blackmun’s observation that too much administrative discretion may permit the IRS to administer tax laws in “so fickle a fashion” continues to be a concern because the Internal Revenue Code is so pervasive in its application and the opportunity for abuse is so great. In the tax exemption area the stakes are too high to tolerate a system that sometimes functions haphazardly or desultorily.

Third, because of the pecuniary value of tax-exempt status, an organization’s survival may depend on the views of the particular administration in office. For example, President Carter had sought to impose racial quotas on tax-exempt private schools to further

terpretation of that term).

90. Contrary to this view, Justice Powell emphasized the following:

[T]he balancing of these substantial interests [of racial nondiscrimination in education and of permitting unorthodox private behavior] is for *Congress* to perform. I am unwilling to join any suggestion that the Internal Revenue Service is invested with authority to decide which public policies are sufficiently “fundamental” to require denial of tax exemptions. Its business is to administer laws designed to produce revenue for the Government, not to promote “public policy.”

103 S. Ct. at 2039 (Powell, J., concurring) (emphasis in original). The majority responded by contending that “The Court’s opinion does not warrant that interpretation . . . [because] the policy against racial discrimination in education . . . is sufficiently clear to warrant Justice Powell’s . . . support [for] our finding of longstanding Congressional acquiescence. . . .” *Id.* at 2032 n.23. This contention, however, lacks any merit. Justice Powell’s concurrence addressed the narrow issue of whether § 501(c)(3) prohibited racial discrimination in education. The majority, on the other hand, made broad pronouncements as to both the meaning of § 501(c)(3) and the authority of the IRS to determine that meaning. See *supra* text accompanying notes 41-43.

91. 103 S. Ct. at 2032.

the nondiscrimination requirement.⁹² In contrast, President Reagan has attempted to lift nonstatutory regulations governing tax-exempt private schools.⁹³ Although neither president succeeded,⁹⁴ the threat remains that similar acts by the Executive branch could bankrupt organizations whose existences depend on tax exemptions.⁹⁵

C. Issues That the Court Did Not Resolve

1. Is a Tax Exemption Government Aid?⁹⁶

Although the Court recognized that tax-exempt status was a governmentally conferred benefit,⁹⁷ it did not say that tax-exempt status is public aid. Whether a tax exemption falls within this classification raises issues under both the Civil Rights Act of 1964, which forbids granting federal aid to institutions that discriminate on the basis of "race, color or national origin,"⁹⁸ and under the

92. 44 Fed. Reg. 9,451 (1979); 43 Fed. Reg. 37,296 (1978). See Wilson, *An Overview of the IRS's Revised Proposed Revenue Procedure on Private Schools as Tax-Exempt Organizations*, 57 TAXES 515 (1979).

93. IRS News Release (Jan. 8, 1982).

94. Congress responded to the Carter proposal by passing riders to the Treasury Appropriations Act of 1980 that prohibited the IRS from spending money to implement it. See *supra* note 53. The Court of Appeals for the District of Columbia enjoined the Reagan administration from granting tax-exempt status to any school that discriminates on the basis of race. *Wright v. Regan*, No. 80-1124 (D.C. Cir. Feb. 18, 1982), *order continued in force*, No. 80-1124 (D.C. Cir. Mar. 24, 1982), *cert. granted*, 103 S. Ct. 3109 (1983).

95. Another troubling possibility of allowing the IRS this broad power is that the IRS might be too deferential and grant tax exemptions to organizations that clearly violate fundamental public policy. This threat is particularly significant because civil rights proponents might not be able to utilize the courts to ensure that IRS procedures are sufficient. Civil rights advocates currently are seeking judicial adoption of stringent enforcement standards through the *Wright v. Regan* case that is now before the Supreme Court. See *supra* note 94. *Wright*, however, ultimately may prove that civil rights groups lack a sufficiently particularized and identifiable harm to bring a lawsuit. See McCoy & Devins, *supra* note 56. The Court's recognition of broad IRS rulemaking authority in *Bob Jones University* actually suggests that the courts will limit their substantive intervention in this area. See Devins, *A Political Analysis of Bob Jones University v. United States*, — J. OF L. & POL. — () (forthcoming). Congress likewise cannot be trusted for satisfactory guidance, as demonstrated both by its failure to make any sort of response to President Reagan's policy shift and by its inability to pass any affirmative legislation on this matter.

96. Portions of this section are adopted from Devins, *supra* note 23, at 163-65.

97. See 103 S. Ct. at 2026-28. "When the Government grants exemptions or allows deductions all taxpayers are affected; the very fact of the exemption or deduction for the donor means that other taxpayers can be said to be indirect and vicarious 'donors.'" *Id.* at 2028. "It would be wholly incompatible with the concepts underlying tax exemption to grant the benefit of tax-exempt status to racially discriminatory educational entities . . ." *Id.* at 2030.

98. 42 U.S.C. §§ 2000-2001 (1964).

establishment clause of the first amendment, which forbids government establishment of religion and limits federal aid to religiously affiliated private schools.⁹⁹

The Civil Rights Act's total prohibition of governmental assistance to discriminatory institutions suggests that its coverage should extend to the granting of tax exemptions to private schools.¹⁰⁰ The United States District Court for the District of Columbia reached this conclusion in *McGlotten v. Connally*,¹⁰¹ in which it decided that a tax exemption to a racially discriminatory fraternal order is federal aid for purposes of the Civil Rights Act.¹⁰² Part of the basis of this holding was the court's recognition that regulations promulgated pursuant to the Act recognized various forms of indirect assistance as federal aid, such as the sale of government property at a reduced price.¹⁰³ Furthermore, the court found that the purpose of the Act "is clearly to eliminate discrimination in programs or activities benefitting from federal financial assistance."¹⁰⁴ The *McGlotten* decision may contribute to what has been described as a "constitutionalizing" of the Internal Revenue Code because it subsumes the revenue collecting function of the Code under broader social policies derived from the Constitution.¹⁰⁵ Professors Bittker and Kaufman see an even broader impact of the court's reasoning:

[T]he "tax-subsidy" rationale of the *McGlotten* case has implications beyond the area of racial restrictions. . . . *McGlotten's* logic apparently prohibits the granting of tax allowances to a fraternal order that imposes such restrictions (based on its customers' religion, national or ethnic origin, political allegiance, sex, and perhaps other characteristics). . . . Finally, nothing in *McGlotten* limits its reach to income, estate and gift taxes; "subsidies" in the form of exemptions, deductions, special rates, and similar allowances may be found in other federal taxes, as well as in state and local taxes.¹⁰⁶

The establishment clause requires a different analysis of tax exemptions. In *Walz v. Tax Commission of New York*,¹⁰⁷ the Supreme Court held that a tax exemption is not governmental aid

99. U.S. CONST. amend. I ("Congress shall make no law respecting an establishment of religion . . ."). See generally L. TRIBE, *AMERICAN CONSTITUTIONAL LAW* §§ 14-8, -9 (1978).

100. Cf. *Tank Truck Rentals v. Comm'r*, 356 U.S. 30 (1958).

101. 338 F. Supp. 448 (D.D.C. 1972).

102. *Id.* at 461.

103. See *id.* at 461 & n.7.

104. *Id.*

105. See Bittker & Kaufman, *Taxes and Civil Rights: "Constitutionalizing" the Internal Revenue Code*, 82 *YALE L.J.* 51 (1972).

106. *Id.* at 62; see *supra* note 56.

107. 397 U.S. 664 (1970).

under the establishment clause. The majority opinion explained that "[t]he grant of a tax exemption is not sponsorship since the government does not transfer part of its revenue to churches but simply abstains from demanding that the church support the state."¹⁰⁸ Although the majority's recognition in *Walz* that a religious institution benefits through a tax exemption seems inconsistent with its principal holding,¹⁰⁹ establishment clause analysis focuses upon whether the "primary effect" of the exemption is to aid the institution, not upon whether some benefit might accrue to the institution.¹¹⁰ Thus, a tax exemption might be permissible under the establishment clause but impermissible under the Civil Rights Act of 1964.

2. Can Some of a Tax-Exempt Organization's Practices Violate Public Policy?

The Court in *Bob Jones University* avoided the question of whether an organization providing a public benefit and satisfying the other requirements of section 501(c)(3) could nevertheless be denied tax-exempt status if certain of its activities violated a law or public policy by holding that racially discriminatory private schools confer no public benefit.¹¹¹ By limiting its holding to racial discrimination in education, the Court neither encouraged nor discouraged the Service from adopting regulations to ensure compliance with other fundamental public policies. Of course, a requirement that an organization desiring tax-exempt status must comply with all laws and all public policies could make the attainment of tax-exempt status incredibly difficult. The amicus curiae Independent Sector¹¹² speculated on the burden of forcing tax-exempt or-

108. *Id.* at 675.

109. *Id.* at 674-75.

110. Supreme Court precedents before 1977 had suggested that almost no form of aid from the state, either to nonpublic schools or to families of nonpublic school students, would be constitutionally permissible. The Court, however, has relaxed this restriction in recent years. Compare *Committee for Pub. Educ. v. Nyquist*, 413 U.S. 756, 780-81 (1973) (tuition grants and deductions and maintenance reimbursements declared unconstitutional), and *Meeke v. Pittenger*, 421 U.S. 349 (1975) (broad range of direct and indirect aid declared unconstitutional), with *Wolman v. Walter*, 433 U.S. 229 (1977) (funding upheld for therapeutic and diagnostic tests but prohibited for field trips and instructional materials), and *Committee for Pub. Educ. v. Regan*, 444 U.S. 646, 658-61 (1980) (direct reimbursement to private schools for state mandated testing upheld).

111. 103 S. Ct. at 2031 n.21.

112. Independent Sector is a coalition of over 400 national voluntary organizations. Its brief, which argued in favor of statutory affirmation of the Fourth Circuit decision in *Bob Jones University* and *Goldsboro Christian Schools, Inc.*, focused on the need for the Court

ganizations to comply with *all* fundamental public policies:

And what about the exempt organizations themselves? Would they become subject to myriad regulations and legal obligations . . . ? If so, retirement homes operated by particular religious charities might be forced to admit persons of any creed; private schools or organizations for girls or boys might have to become coeducational; senior citizens groups might be forbidden to discriminate on the basis of age; community centers designed to serve particular ethnic groups might have to open their doors to all comers; and any exempt organization might be required to modify its physical facilities to provide access to the handicapped. Resolving these and similar questions that might be raised by an overbroad holding in this case could occupy exempt organizations, their benefactors, the courts, and Congress for years to come.¹¹³

The Court's stated limitation of its holding to the racial discrimination issue is no satisfactory answer to these concerns.

3. Does Granting of Tax-Exempt Status Trigger Constitutional Scrutiny?

Several amici in *Bob Jones University* urged the Supreme Court to hold that granting tax-exempt status to racially discriminatory private schools would violate the fifth amendment's equal protection component.¹¹⁴ These amici argued that the government's grant of a tax exemption is "state action" subject to constitutional restraints that prohibit government from providing tangible financial aid to racially discriminatory private schools.¹¹⁵ The Court did not address this issue because it was able to decide the case on statutory grounds.¹¹⁶

Although the decisions on whether tax exemptions constitute state action are inconsistent, at least some trends emerge. Courts tend to find state action more often when racial discrimination is at issue¹¹⁷ and when the action sought to be stopped is the govern-

to limit its holding to the issue of racial discrimination in education. *See infra* note 113 and accompanying text.

113. Brief for Independent Sector at 20, *Bob Jones Univ.*

114. *See* Brief of William T. Coleman, Jr. at 57-62; Brief for Lawyers' Committee for Civil Rights Under Law at 30; Brief for American Civil Liberties Union at 17-32; Brief for North Carolina Association of Black Lawyers at 5-7.

115. *See, e.g.,* *Norwood v. Harrison*, 413 U.S. 455 (1973) (Mississippi statute that would have provided textbook assistance to racially discriminatory private schools held unconstitutional). In *Norwood*, the Court recognized that "discriminatory treatment exerts a pervasive influence on the entire educational process." *Id.* at 469.

116. 103 S. Ct. at 2032 n.24.

117. Courts generally apply a less stringent test for state action in race discrimination cases. *See, e.g.,* *Spark v. Catholic Univ. of Am.*, 510 F.2d 1277, 1281-82 (D.C. Cir. 1975) (*per curiam*); *Greco v. Orange Memorial Hosp. Corp.*, 513 F.2d 873, 879 (5th Cir.), *cert. denied*, 423 U.S. 1000 (1975); *Greenya v. George Washington Univ.*, 512 F.2d 556, 560 (D.C. Cir.), *cert. denied*, 423 U.S. 995 (1975). This principle extends to tax exemption cases. For exam-

mental grant of tax-exempt status rather than private discriminatory conduct.¹¹⁸ Existing doctrine thus suggests that the Constitu-

ple, in *Pitts v. Department of Revenue*, 333 F. Supp. 662 (E.D. Wis. 1971) (three judge court), a federal district court in Wisconsin held that "the 'state action' doctrine was developed in response to efforts to eliminate private racial discrimination Accordingly[,] . . . there might be a less demanding standard of what constitutes sufficient state involvement where there are allegations of racial discrimination." *Id.* at 667 (quoting *Bright v. Isenbarger*, 314 F. Supp. 1382, 1392-94 (N.D. Ind. 1970)). Similarly, in *Jackson v. Statler Found.* the Second Circuit noted the following:

Where racial discrimination is involved, the courts have found "state action" to exist; where other claims are at issue (due process, freedom of speech), the courts have generally concluded that no "state action" has occurred. . . . [Thus, there is] a less onerous test for cases involving racial discrimination, and a more rigorous standard for other claims.

496 F.2d 623, 628-29 (2d Cir. 1974) (citations omitted); see also Note, *The Judicial Role in Attacking Racial Discrimination in Tax-Exempt Private Schools*, 93 HARV. L. REV. 378, 399 (1979). Cf. Neuberger & Crumplar, *supra* note 31, at 246-48 (arguing that religious schools might be judged under a different "state action" standard since an important purpose of granting tax exemptions to such institutions is the avoidance of impermissible government entanglement with religion).

118. One commentator articulated the rationale for this approach as follows:

Because enjoining the private party necessarily will affect private interests, some weight must be accorded these interests during the course of judicial inquiry. In such a case, the court should inquire whether there is a sufficient nexus of governmental and private action to transform the private actor into an agent of the government. When the litigation is directed at the conduct of the government, however, the issue becomes whether the government's actions encourage or support violations of constitutional guarantees. Thus, the different implications of the remedies sought suggest that different constitutional standards should be formulated for the two types of cases.

Comment, *The Tax-Exempt Status of Sectarian Educational Institutions that Discriminate on the Basis of Race*, 65 IOWA L. REV. 258, 265-66 (1979). See also Brown, *State Action Analysis of Tax Expenditures*, 11 HARV. C.R.-C.L. L. REV. 97, 116-22 (1976). Case law also supports this view. Compare *Falkenstein v. Dep't of Revenue*, 350 F. Supp. 887 (D. Or. 1972) (three judge court) (state tax exemptions for racially exclusive fraternal organization held unconstitutional), *appeal dismissed for want of jurisdiction*, 409 U.S. 1099 (1973), and *Pitts v. Department of Revenue*, 333 F. Supp. at 662 (state tax exemptions for racially discriminatory organizations enjoined), with *New York City Jaycees, Inc. v. United States Jaycees, Inc.*, 512 F.2d 856 (2d Cir. 1975) (sex discrimination by exempt organization held not state action); *Junior Chamber of Com. v. United States Jaycees*, 495 F.2d 883 (10th Cir.) (same), *cert. denied*, 419 U.S. 1026 (1974); and *Stearns v. Veterans of Foreign Wars*, 394 F. Supp. 138 (D.D.C. 1975) (memorandum opinion) (sex discrimination by congressionally chartered organization held not state action), *cert. denied*, 429 U.S. 822 (1976).

Even if a court finds state action, it also must determine whether tax-exempt status is a significant governmental benefit and whether the granting of tax-exempt status to racially discriminatory institutions constitutes intentional discrimination. This Article has adopted the view that tax exemptions are government aid for purposes of the 1964 Civil Rights Act. See *supra* notes 96-110; see also Comment, *Tax Incentives as State Action*, 122 U. PA. L. REV. 414, 421-23, 453-55 (1973). But see Bittker & Kaufman, *supra* note 105, at 63-68; Note, *supra* note 55, at 180-84. The intentional discrimination requirement poses no problem in the context of tax exemptions to private schools with stated policies of racial discrimination, such as Bob Jones University (interracial dating) and Goldsboro Christian Schools (admissions). IRS knowledge of such policies satisfies the intent requirement established in Wash-

tion may prevent the government from granting tax-exempt status to racially discriminatory institutions.¹¹⁹ Fortunately, the *Bob Jones University* Court was able to avoid this issue by deciding on statutory grounds. Constitutional prohibitions are broader than 1964 Civil Rights Act standards. Consequently, “[a] broad holding that tax exemptions are [state] action for constitutional purposes could leave exempt organizations vulnerable to legal challenges on a variety of theories having nothing to do with racial discrimination. That result would be contrary to the public interest in encouraging philanthropic activity”¹²⁰

IV. CONCLUSION

The *Bob Jones University* decision exemplifies the risks of leaving tax policy determinations to the courts. On one hand, the Court went too far in interposing into the Code its own standards of “common community conscience” and “public purpose.”¹²¹ On the other hand, the Court appeared to abdicate its supervisory powers to the IRS.¹²² This Article recommends that in the formulation of tax policy, courts should not supplant the role of Congress as lawmaker by making broad tax policy pronouncements, but should oversee the IRS to ensure that it properly implements and enforces the tax laws.

Thus far Congress has relied on judicial and IRS initiatives to define the tax exemption requirements and the parameters of the IRS’s rulemaking authority; it has interceded only when dissatisfied with the actions of the other branches.¹²³ Congress, however, must take the lead in resolving the tax-exemption issue, as Justice

ington v. Davis, 426 U.S. 229 (1976) (qualifying test for police officers with racially disproportionate impact held constitutional). If the IRS was unaware of a school’s racially discriminatory practices, however, under *Washington* that school would retain its tax-exempt status. The solution here would be to impute to the IRS notice of all schools that fail to satisfy racial nondiscrimination enforcement standards.

119. The *Green* decisions all but concluded that tax-exempt status constituted state action. In *Green v. Kennedy* the court issued the injunction, in part, because “of the substantiality of the grave constitutional questions presented by plaintiffs.” 309 F. Supp. at 1133. See *supra* note 20 and accompanying text. In *Green v. Connally* the court finalized that temporary injunction in the form of a permanent injunction. 330 F. Supp. at 1150. Although it based its holding on statutory grounds, the court noted: “We are fortified in our view of the correctness of the IRS construction by the consideration that a contrary interpretation of the tax laws would raise serious constitutional questions, such as those we ventilated in [*Green v. Kennedy*].” *Id.* at 1164.

120. Brief for Independent Sector, *supra* note 113, at 27-28.

121. See *supra* notes 65-67 and accompanying text.

122. See *supra* notes 84-95 and accompanying text.

123. See generally McCoy & Devins, *supra* note 56.

Powell emphasized in his concurrence in *Bob Jones University*:

There no longer is any justification for Congress to hesitate—as it apparently has—in articulating and codifying its desired policy. . . . Many questions remain, such as whether organizations that violate other policies should receive tax-exempt status under § 501(c)(3). These should be legislative policy choices. It is not appropriate to leave the IRS “on the cutting edge of developing national policy.” . . . The contours of public policy should be determined by Congress, not by judges or the IRS.¹²⁴

Congress, as a more capable legislator than the courts or the IRS, is better equipped to formulate a tax exemption policy. The complexity of an indefinite and unidentified class of potential plaintiffs with perhaps varying levels of grievance, an open-ended class of defendant institutions neither entirely similar nor dissimilar to the particular institutions in litigation, and the economic costs and

124. 103 S. Ct. at 2039 (Powell, J., concurring). According to Department of Justice attorneys who have worked on this matter, Congress' passivity has been costly: “[T]he continued litigation of the issue in open-ended injunction suits, coupled with Congress' decision to prohibit new policy shifts by the Treasury, has caused a paralysis among the three branches of government. This paralysis has prevented the establishment of further guidelines to meet changing conditions.” Ferguson & Henzken, *supra* note 11, at 103. Congress should take the lead in resolving the tax exemption issue because Congress is a better legislator than are the courts. In a Brookings Institution study of court efforts to develop and implement social policy, Donald Horowitz drew this conclusion:

The distinctiveness of the judicial process—its expenditure of social resources on individual complaints, one at a time—is what unfits the courts for much of the important work of government. Retooling the judicial process to cope with the new responsibilities of the courts means enhancing their capacity to function more systematically in terms of general categories that transcend individual cases. Some such innovations are required. And yet, it would seem, there is a limit to the changes of this kind that courts can absorb and still remain courts. Heightened attention to recurrent patterns of behavior risks inattention to individual cases. Over the long run, augmenting judicial capacity may erode the distinctive contribution the courts make to the social order. The danger is that courts, in developing a capacity to improve on the work of other institutions, may become altogether too much like them.

D. HOROWITZ, *THE COURTS AND SOCIAL POLICY* 298 (1977). Professor Alexander Bickel similarly noted that the Court's institutional survival hinged on its ability to abide by the constraints of our system of divided powers. A. BICKEL, *THE SUPREME COURT AND THE IDEA OF PROGRESS* (1970). Considering recent congressional efforts to limit federal court jurisdiction, Professor Bickel's fears seem well founded. See generally Sager, *The Supreme Court, 1980 Term—Foreword: Constitutional Limitations on Congress' Authority to Regulate the Jurisdiction of the Federal Courts*, 95 HARV. L. REV. 17 (1981).

Many jurists, however, feel that the courts do have a place in the shaping of social policy. As Justice Cardozo suggested, “[W]hen the social needs demand one settlement rather than another, there are times when we must bend symmetry, ignore history and sacrifice custom in the pursuit of other and larger ends.” B. CARDOZO, *THE NATURE OF THE JUDICIAL PROCESS* 65 (1921). Similarly, Archibald Cox and James Hart Ely have suggested that the Supreme Court has been at its best when it has introduced universalistic normative principles in an effort to set the parameters of acceptable social behavior. A. COX, *THE ROLE OF THE SUPREME COURT IN AMERICAN GOVERNMENT* (1976); J. ELY, *DEMOCRACY AND DISTRUST* (1980).

budgetary concerns affecting not only the parties but also others demand that Congress glean and sift the social facts, weigh the costs, and determine time periods for phasing in and phasing out particular practices.¹²⁵

Given that Congress should take the lead in clarifying the requirements private schools must meet to attain tax-exempt status, the issue becomes how such legislation should be formulated. As a starting point, Congress should assess current enforcement procedures, which require that a tax-exempt private school "show affirmatively both that it has adopted a racially nondiscriminatory policy as to students that is made known to the general public and that since the adoption of that policy it has operated in a bona fide manner in accordance therewith."¹²⁶ A school can comply with these current requirements if it (1) adopts formally nondiscriminatory policies in its charter or by-laws, (2) refers to these policies in its advertising brochures, and (3) publishes annual notice of these policies in a local newspaper of general circulation.¹²⁷ These procedures are insufficient; private schools adjudicated as discriminatory under the fourteenth amendment and thus ineligible to receive direct government assistance often qualify for tax-exempt status.¹²⁸ The tax exemption rules ought to conform with the 1964 Civil Rights Act, which prohibits government aid to private schools that have no minority students or staff and that were formed or substantially expanded at or about the time of area-wide public school desegregation.¹²⁹

*Brown v. Board of Education*¹³⁰ is now almost thirty years old; its call for due deliberate speed in the elimination of segregation in the nation's school system still goes unheeded in many quarters. The current emphasis on the need for educational excellence in primary and secondary education is of no greater urgency than the need to achieve equality of educational opportunity for every child. Congress has the capability and the competency to act—and surely

125. See D. HOROWITZ, *supra* note 124, at 255-74.

126. Rev. Proc. 75-50, 1975-2 C.B. 587.

127. *Id.* at 587-88.

128. Under 1964 Civil Rights Act standards, the government cannot grant aid to a school if a judicial or administrative proceeding has determined that school to be discriminatory, or if the school was established at a time when public schools in its area were desegregating and the school cannot demonstrate that it was nondiscriminatory. See, e.g., *Norwood v. Harrison*, 413 U.S. 455 (1973).

129. For a similar proposal, in the form of a Model Statute, see Devins, *supra* note 23, at 176-78.

130. 347 U.S. 483 (1954).

must act—in formulating effective national tax policy that incorporates *Brown's* objectives.

The Parol Evidence Rule: Promissory Estoppel's Next Conquest?

Michael B. Metzger*

I. INTRODUCTION	1383
II. THE PAROL EVIDENCE RULE	1385
A. <i>The Policy Bases of the Rule</i>	1386
B. <i>The Parol Evidence Rule: Formulations, Ex-</i> <i>ceptions, and Applications</i>	1391
C. <i>The Parol Evidence Rule: Criticisms and Sug-</i> <i>gestions for Improvement</i>	1403
III. PROMISSORY ESTOPPEL—THE EVOLUTION OF THE DOCTRINE	1408
IV. PROMISSORY ESTOPPEL AND THE STATUTE OF FRAUDS	1422
A. <i>The Statute of Frauds</i>	1422
B. <i>Promissory Estoppel and the Statute of</i> <i>Frauds</i>	1425
V. PROMISSORY ESTOPPEL AND THE PAROL EVIDENCE RULE	1437
A. <i>The Cases</i>	1437
B. <i>Should Promissory Estoppel Apply to Parol</i> <i>Evidence Cases?</i>	1454
VI. CONCLUSION	1466

I. INTRODUCTION

A specter is haunting the law of contracts. The doctrine of promissory estoppel has evolved from relatively modest beginnings as a "consideration substitute" in donative promise cases¹ to a force that threatens to engulf a major portion of contract law.² In-

* Associate Professor and Chairman of Business Law, Indiana University Graduate School of Business. B.A., 1966, Indiana University; J.D., 1969, Indiana University School of Law.

1. See *infra* notes 251-60 and accompanying text.
2. See *infra* notes 306-11 and accompanying text.

deed, some commentators have suggested that promissory estoppel is evolving into a separate theory of recovery, independent of contract law.³ To the extent that this evolution continues, the future of many traditional contract rules, such as the parol evidence rule, is doubtful.⁴

Regardless of whether promissory estoppel ultimately achieves the status of an independent theory or merely continues its evolution within the parameters of contract law, the historical course of the reliance principle's expansion exhibits such significant momentum that the thoughtful observer must speculate about likely future candidates for the doctrine's debilitating ministrations.⁵ The parol evidence rule, at first glance, seems to be such a candidate for many reasons. The parol evidence rule has confused⁶ and dissatisfied⁷ legal scholars for a long time; for example, Professor Wigmore condemned the rule as "the most discouraging subject in the whole field of evidence."⁸ Bringing the rule within estoppel's domain could simplify the application of the rule,⁹ and legal scholars should appreciate anything that could clarify and rationalize its application. Furthermore, that promissory estoppel already has made substantial incursions into the province of the Statute of Frauds¹⁰ may portend a similar role for promissory estoppel in the parol evidence context. Many similarities exist between the Statute and the rule¹¹ and many of the same arguments that justify the use of estoppel to circumvent the Statute¹² are equally applicable to justify the use of estoppel to bypass the parol evidence rule.¹³

Few cases to date have explored the possible interaction between promissory estoppel and the parol evidence rule.¹⁴ This Arti-

3. For an exhaustive discussion of this possibility, see Metzger & Phillips, *The Emergence of Promissory Estoppel as an Independent Theory of Recovery*, 35 *RUTGERS L. REV.* 472 (1983). See also *infra* notes 295-308 and accompanying text.

4. See *infra* note 434 and accompanying text.

5. One commentator describes promissory estoppel as "perhaps the most radical and expansive development of this century in the law of promissory liability." Knapp, *Reliance in the Revised Restatement: The Proliferation of Promissory Estoppel*, 81 *COLUM. L. REV.* 52, 53 (1981).

6. See *infra* notes 15-25 and accompanying text.

7. See *infra* notes 162-202 and accompanying text.

8. 9 J. WIGMORE, *A TREATISE ON THE ANGLO-AMERICAN SYSTEM OF EVIDENCE IN TRIALS AT COMMON LAW* § 2400, at 3 (3d ed. 1940).

9. See *infra* notes 583-87 and accompanying text.

10. See *infra* notes 354-435 and accompanying text.

11. See *infra* notes 521-30 and accompanying text.

12. See *infra* notes 347-48 & 377-419 and accompanying text.

13. See *infra* notes 532-37 & 545-64 and accompanying text.

14. See *infra* notes 438-518 and accompanying text.

cle attempts to explain why so few cases have explored this interaction, and discusses both the likelihood and the desirability of a significant intrusion by promissory estoppel into the parol evidence rule's domain. First, however, this Article necessarily focuses upon the parol evidence rule itself, the evolution of the promissory estoppel doctrine in general, and the doctrine's development as a device for circumventing the strictures of the Statute of Frauds.

II. THE PAROL EVIDENCE RULE

The parol evidence rule long has been "the source of endless confusion in contract law."¹⁵ As Thayer noted eight decades ago, "[f]ew things are darker than this, or fuller of subtle difficulties."¹⁶ Little has occurred in the ensuing years to gainsay the veracity of Thayer's observation. Even the rule's name is somewhat misleading.¹⁷ "Parol" misleads because most formulations of the rule include within its scope written as well as oral terms.¹⁸ While the use of the word "evidence" in the rule's title and in some judicial statements about the rule¹⁹ may suggest that the rule is an evidentiary rule, legal scholars generally agree that the parol evidence rule is actually a rule of law.²⁰ Finally, some commentators have questioned whether "a maze of conflicting tests, subrules, and exceptions"²¹ is a "rule" at all, especially in light of the number of exceptions to its operation²² and the uncertainties surrounding

15. Note, *The Parol Evidence Rule: Is It Necessary?*, 44 N.Y.U. L. Rev. 972, 972 (1969).

16. J. THAYER, A PRELIMINARY TREATISE ON EVIDENCE AT THE COMMON LAW 390 (1898).

17. "The term 'parol evidence rule' is something of a misnomer for it does not deal exclusively with parol, is not a rule of evidence, and at least in the minds of some, is not even a rule." Wallach, *The Declining "Sanctity" of Written Contracts—The Impact of the Uniform Commercial Code on the Parol Evidence Rule*, 44 Mo. L. Rev. 651, 651 (1979). See also Murray, *The Parol Evidence Rule: A Clarification*, 4 Duq. L. Rev. 337, 343 (1965-1966).

18. Murray, *supra* note 17, at 343. For a discussion of various formulations of the rule, see *infra* notes 61-69 and accompanying text.

19. "Because judges, in reserving a question of fact to themselves couched the process in rules which sounded like rules of evidence, the term 'evidence' is part of the title of the rule which only serves to add to the mystery surrounding it." Murray, *supra* note 17, at 343.

20. See, e.g., J. CALAMARI & J. PERILLO, *CONTRACTS* § 3-5 (2d ed. 1977); 9 J. WIGMORE, *supra* note 8, § 2400, at 4 (3d ed. 1940); Corbin, *The Interpretation of Words and the Parol Evidence Rule*, 50 CORNELL L.Q. 161, 189 (1965).

21. Sweet, *Contract Making and Parol Evidence: Diagnosis and Treatment of a Sick Rule*, 53 CORNELL L. REV. 1036, 1036 (1968).

22. The number of recognized exceptions to the rule "raises doubts about its status as a 'rule' at all." See Note, *supra* note 15, at 972. For a discussion of the exceptions to the rule, see *infra* notes 124-51 and accompanying text.

both its definition and application.²³

The rule's appellation, the variance among the numerous formulations of the rule, and the observed tendency of the courts to manipulate the rule to avoid the harsh results that otherwise may flow from its strict application²⁴ all undoubtedly have contributed to the fog of uncertainty surrounding the rule. The primary source of confusion, however, lies with misapprehensions concerning both the nature of the policies that the rule intends to serve and those policies' relative merit.²⁵ An inquiry into the rule's underlying policy bases necessarily must preface any explication of the rule in its various forms. This inquiry is particularly important in assessing the degree to which using promissory estoppel to circumvent the rule may hinder attainment of the rule's underlying policy objectives.

A. *The Policy Bases of the Rule*

Scholars have traced the historical origins of the parol evidence rule to "a primitive formalism which attached mystical and ceremonial effectiveness to the *carta* and the seal."²⁶ The rule may have survived the disappearance from our law of other vestiges of formalism because a written contract furnishes more reliable evidence of the terms of an agreement than does the parties' oral testimony, which may be the product of faulty memory, wishful thinking, or outright prevarication.²⁷ A major function of the rule,

23. "Those who question whether the parol evidence rule is really a rule at all seem to feel that certainty of definition and application are required before something can be labeled a 'rule,' and the parol evidence rule tends to lack both." Wallach, *supra* note 17, at 651. For a discussion of the rule's various formulations, see *infra* notes 61-69 and accompanying text. For a discussion of the inconsistencies in the application of the rule, see *infra* text accompanying notes 120-23, 130, 149 & 153-61.

24. See *infra* notes 153-61 and accompanying text.

25. Professors Calamari and Perillo attribute the mystery surrounding the rule to a "basic disagreement as to the application of the . . . rule and as to the best method of ascertaining the intention of the parties—the process of contractual interpretation." J. CALAMARI & J. PERILLO, *supra* note 20, § 3-1. They further observe that: "The cases and treatises of the contract giants tend to conceal this conflict. While frequently masking disagreement by using the same terminology, Professors Williston and Corbin are often poles apart in the meaning they attach to the same terms." *Id.* Calamari and Perillo attribute this polarity of opinion to "conflicting value judgments as to policy issues that are as old as our legal system and that are likely to continue as long as courts of law exist." *Id.*

26. C. McCORMICK, *HANDBOOK OF THE LAW OF EVIDENCE* § 211, at 430 n.4 (1954); see also 9 J. WIGMORE, *supra* note 8, § 2426.

27. See, e.g., Patterson, *The Interpretation and Construction of Contracts*, 64 COLUM. L. REV. 833, 846 (1964) ("[T]he written instrument is a more reliable expression of the meaning of their contract than one party's not disinterested memory of his or the other

therefore, is the prevention of fraud or perjury.²⁸ This rationale, however, does not explain the application of the rule to written evidence of alleged terms that are not contained in the parties' final written contract.²⁹ This rationale also fails to explain the special role that the judge plays in parol evidence cases. The judge initially decides whether the rule excludes certain evidence. Why should jurors, whom courts rely upon to determine truth in a wide variety of legal contexts, lack the ability to determine truth reliably in matters of parol evidence?

Thus, while special devotion to the written word may have provided the origin of the parol evidence rule, the pervasive attitude that judges provide the best protection against perjured testimony³⁰ probably has been the reason for its continued viability.³¹ A distrust of the jury as a reliable mechanism for divining truth³² underlies the parol evidence rule. Left to their own devices, jurors may favor underdogs by relying upon alleged oral terms,³³ thereby deciding the case in a manner calculated to avoid a perceived injustice.³⁴ Jurors also may lack the sophistication needed to deal ef-

party's prior oral utterances."); see also *Kupka v. Morey*, 541 P.2d 740, 747-48 n.9 (Alaska 1975); C. McCORMICK, *supra* note 26, § 210, at 428; J. WHITE & R. SUMMERS, *HANDBOOK OF THE LAW UNDER THE UNIFORM COMMERCIAL CODE* § 2-9 (2d ed. 1980); Sweet, *supra* note 21, at 1036; Note, *supra* note 15, at 982.

28. See, e.g., *Kupka v. Morey*, 541 P.2d 740, 747-48 n.9 (Alaska 1975); *Masterson v. Sine*, 68 Cal. 2d 222, 227, 65 Cal. Rptr. 545, 548, 436 P.2d 561, 564 (1968); *Patterson*, *supra* note 27, at 846; Note, *supra* note 15, at 982.

29. The rule "would seem to lose much of its force if the prior expression is written. The prior writing is not unchanged [sic] at the time of trial any more than the subsequent writing." Murray, *supra* note 17, at 342; see also Farnsworth, "Meaning" in the Law of Contracts, 76 YALE L.J. 939, 958 (1967). One scholar, however, notes that the vast majority of decided cases involve alleged oral terms. Wallach, *supra* note 17, at 651.

30. "If the parol evidence doctrine is to be administered chiefly by juries and not by judges, its special protection for written transactions fades." C. McCORMICK, *supra* note 26, § 216, at 441-42; see also Note, *supra* note 15, at 982-83.

31. Professor McCormick argued that the formalist attributes of the rule had "a special survival value—the escape from the jury—which led the judges to retain for writings the conception that they had a sort of magical effect of erasing all prior agreements." C. McCORMICK, *supra* note 26, § 211, at 430 n.4.

32. "[L]awyers have a strong distrust of the judicial process, especially the jury, as a means of ascertaining the truth." Sweet, *supra* note 21, at 1049; see also Wallach, *supra* note 17, at 653 n.9.

33. Professor Sweet notes that juries may "favor the underdog who rarely has the writing on his side." Sweet, *supra* note 21, at 1050. See also C. McCORMICK, *supra* note 26, at § 210 ("[W]hen an issue arises involving choice between a writing and an alleged oral agreement, usually the one who sets up the spoken against the written word is economically the underdog.").

34. "If the jury is directed to bring in a general verdict (i.e. for the plaintiff or for the defendant), it may in doing so exercise its views of jury equity and thus impair the reliability of written instruments." *Patterson*, *supra* note 27, at 837. Professor McCormick similarly

fectively with complex commercial transactions involving numerous alleged oral and written contract terms.³⁵ Although under modern rules of procedure a judge may either overrule a jury's decision or decide the case himself when the proffered evidence reasonably supports only one conclusion, a judge may be unwilling to do so.³⁶ According to this view, the major service that the rule performs³⁷ is to allow judges to "strike down at the outset claims of oral variants on the writing, variants which the judge believes never were entered into but are fabrications, designed or unconscious."³⁸

Fears concerning the unreliability of oral testimony and the inadequacies of the jury as a device for divining the truth, however, do not fully explain the sanctity that the parol evidence rule affords written agreements. Legal scholars note that the parol evidence rule is a rule of form,³⁹ and, therefore, performs the "channeling" function common to rules of form⁴⁰ by encouraging contracting parties to embody the terms of their agreement in a proper writing.⁴¹ Under this view, courts deny enforcement of parol terms merely "because the parties have not used the proper form."⁴² Rules of form embody certain assumptions about the rela-

observed: "The average jury will, other things being equal, lean strongly in favor of the side which is threatened with possible injustice and certain hardship by the enforcement of the writing." C. McCORMICK, *supra* note 26, § 210, at 428.

35. "At an early date it was felt (and the feeling strongly remains) that writings require the special protection that is afforded by removing this issue from the province of unsophisticated jurors." J. CALAMARI & J. PERILLO, *supra* note 20, § 3-2, at 100. *See also* 9 J. WIGMORE, *supra* note 8, § 2426, at 86.

36. C. McCORMICK, *supra* note 26, § 216, at 440.

37. Other services of the rule include allowing trial judges to avoid issues concerning the agents' authority, the vagueness of oral agreements, and the relevance of oral conditions. Sweet, *supra* note 21, at 1050.

38. C. McCORMICK, *supra* note 26, § 216, at 440. As Professor Sweet noted, "[I]f a trial judge doubts that an asserted agreement ever took place, ruling on the basis of the parol evidence rule avoids commenting on the evidence, instructing on the weight and burden of proof, or granting a new trial. It also avoids branding the witness a liar." Sweet, *supra* note 21, at 1051. *See also* Wallach, *supra* note 17, at 653 (Trial judges who exclude evidence under the aegis of the rule do so either to protect the writing or to prevent less than credible evidence from reaching the jury.).

39. *See* J. CALAMARI & J. PERILLO, *supra* note 20, § 3-3, at 108 ("Since (and even before) the common law had its genesis, there has been a deeply-felt belief that transactions will be more secure, litigation will be reduced, and the temptation to perjury will be removed, if everyone will only use proper forms for his transactions.").

40. *See infra* notes 402-03 & 407-11 and accompanying text for a discussion of the channeling function in the Statute of Frauds context.

41. *See* Sweet, *supra* note 21, at 1036; Wallach, *supra* note 17, at 653; Note, *supra* note 15, at 982.

42. Sweet, *supra* note 21, at 1050.

tive reliability of formal versus informal transactions;⁴³ perhaps more importantly, however, the law perceives rules of form as adding stability to commercial transactions.⁴⁴ Thus, by protecting the integrity of written contracts the parol evidence rule may provide the predictability and certainty necessary for the efficient conduct of business.⁴⁵ The special role that the judge plays in parol evidence cases arguably is consistent with this view because jurors may not appreciate sufficiently the need for stability and certainty in commercial dealings.⁴⁶ A related subsidiary argument that commentators often advance in conjunction with this business efficiency rationale for the rule notes that the rule allows principals to control unauthorized commitments by their agents through the use of integration clauses in their contracts.⁴⁷

Finally, another body of thought exists that attaches significantly less importance to writings *qua* writings. Under this view, the parol evidence rule is nothing more than a particularized version of the basic contractual interpretation rule which stipulates that later final expressions of intent prevail over earlier tentative expressions of intent.⁴⁸ Two major differences exist between the parol evidence rule and the more general rule of interpretation.

43. On this point Sweet makes the following observation:

A rule of form assumes that parties generally follow formal rules. The absence of proper form indicates it is unlikely that the agreement was made. Also, a rule of form assumes that witnesses will lie or forget facts if it is to their advantage to do so. Such a rule assumes people have poor memories and that litigation is an inefficient method of ascertaining facts.

Id. at 1053.

44. See *supra* note 39.

45. The Minnesota Supreme Court in 1924 observed that

[w]ithout that rule there would be no assurance of the enforceability of a written contract. If such assurance were removed today from our law, generally disaster would result, because of the consequent destruction of confidence, for the tremendous but closely adjusted machinery of modern business cannot function at all without confidence in the enforceability of contracts. They must not be reduced to the innocuous character of a mere "scrap of paper."

Cargill Comm'n Co. v. Swartwood, 159 Minn. 1, 7, 198 N.W. 536, 538 (1924). See also *Holton v. Bivens*, 9 U.C.C. REP. SERV. (Callaghan) 836, 838 (Okla. Ct. App. 1971) ("Every person should have the right to rely upon written contracts. Allowing testimony of oral agreements in contradiction thereto can only result in confusion concerning reliability of the writing.").

46. Sweet, *supra* note 21, at 1036.

47. *Id.* at 1050, 1057. For a discussion of merger clauses and their effects, see *infra* note 189 and accompanying text.

48. See, e.g., J. CALAMARI & J. PERILLO, *supra* note 20, § 3-2; 3 A. CORBIN, CORBIN ON CONTRACTS § 574 (rev. ed. 1960); Farnsworth, *supra* note 29, at 958; Murray, *supra* note 17, at 338; Comment, *The Parol Evidence Rule: A Conservative View*, 19 U. CHI. L. REV. 348, 348 (1952).

The parol evidence rule only applies when the parties' final expression of intent is in written form, while the general rule of interpretation applies regardless of the final expression's form.⁴⁹ Furthermore, under the general rule of interpretation the jury determines whether the parties intended their last expression to supersede prior expressions, while under the parol evidence rule the trial judge makes this determination.⁵⁰

Under this view the primary purpose of the rule is to prevent courts from interpreting earlier, tentative agreements or negotiations as part of an integrated writing that the parties actually intended as the final expression of their agreement.⁵¹ Thus, according to this view the rule's justification is based upon the finality of the parties' written agreement.⁵² Courts exclude oral or written⁵³ terms extraneous to such a writing not because doubt exists concerning the terms' reliability, but rather because the terms are irrelevant, since the parties superseded them in the final integrated writing.⁵⁴

This last view of the rule—the rule as insurer that the final expression of intent governs—seems to be currently in vogue.⁵⁵

49. J. CALAMARI & J. PERILLO, *supra* note 20, § 3-2, at 99-100 (emphasis in original).

50. *Id.* at 100.

51. C. McCORMICK, *supra* note 26, § 210, at 428.

52. "[A]fter the parties have discussed and negotiated the terms of the transaction and have agreed on the terms by signing an instrument embodying them, all prior utterances have become discarded by the final act . . . which is the sole expression of their final promises and expectations." Patterson, *supra* note 27, at 846.

53. The policy of "*finality* applies to prior written as well as oral utterances." Patterson, *supra* note 27, at 846 (emphasis in original). Professor Murray, however, expresses some reservations on this issue:

The rule exists to afford special protection to subsequent written expressions of agreement when the parties intend such expressions to be complete and final. Whether the parties intended this result in a particular case is a question of fact which is decided by the trial judge . . . because juries cannot be trusted to accord proper weight to the writing over prior oral expressions. Stating the purpose and operation of the rule in this fashion, it is difficult to clearly apply it when the prior expression is written. If juries would not accord proper weight to a final writing over a prior writing, perhaps the reason for the rule remains even where both expressions are written. Otherwise, the question of fact involved should be decided as are other questions of fact.

Murray, *supra* note 17, at 342-43.

54. Professor Wigmore observed that when a judge decides to apply the parol evidence rule to the parties' agreement "he does not decide that the excluded negotiations did not take place, but merely that *if* they did take place they are nevertheless legally immaterial." 9 J. WIGMORE, *supra* note 8, § 2430, at 98 (emphasis in original). When the judge decides that the rule does not apply "he does not decide that the negotiations did take place, but merely that *if* they did, they are legally effective, and he leaves to the jury the determination of fact whether they did take place." *Id.* (emphasis in original).

55. *See, e.g., id.* § 2425, at 75-76; Sweet, *supra* note 21, at 1036.

This view has influenced both the Uniform Commercial Code's version of the rule⁵⁶ and the position of the *Restatement (Second) of Contracts*.⁵⁷ Nonetheless, the other rationales, especially the greater relative reliability of written versus oral expressions of intent, remain important justifications for several reasons. First, the major dispute in most parol evidence cases does not concern whether the contracting parties intended the writing to supersede alleged prior expressions of their intent, but whether the parties ever actually agreed to the prior expressions.⁵⁸ Second, although many courts discuss the ascertainment of intent, they often use language which suggests that they are barring parol evidence because of its unreliability.⁵⁹ Last, the influence of ideas relating to the special significance of writings is probably the best explanation for those formulations of the rule that take a more restrictive view of the critical process by which courts may ascertain best the parties' intent to integrate their agreement into a final writing.⁶⁰

B. The Parol Evidence Rule: Formulations, Exceptions, and Applications

While the essence of the parol evidence rule is clear, significant disagreement exists about the rule's various intricacies. Professor Murray explains the gist of the rule as follows: "When the parties to a contract embody the terms of their agreement in a writing, intending that writing to be the final expression of their agreement, the terms of the writing may not be contradicted by

56. See *infra* notes 99-106 and accompanying text.

57. See *infra* notes 107-15 & 131 and accompanying text.

58. As Professor McCormick noted:

Seldom from the case does one gain the impression that the dispute is really over whether an admitted oral agreement was intended to be superseded by the writing. Where the adversary's position is to be gleaned from the report, which is surprisingly seldom, it appears most often that he denies that any such oral counter-agreement was ever entered into.

C. McCORMICK, *supra* note 26, § 216, at 440. If Professor McCormick is correct, and intuitively he seems to be, then much of the recent concern with the rule as a device for effectuating the intent of the parties, however laudable that objective may be in terms of avoiding some of the obvious unfairness that results when courts treat the rule as a rule of form, may be largely misplaced. If the true function of the rule remains one of helping determine whether the parties ever entered into the alleged parol agreements, then conclusions about the rule's utility as a device in aiding the truth-seeking process should influence opinions about the desirability of the rule's continued viability.

59. Sweet, *supra* note 21, at 1048.

60. For a discussion of the polarity between Williston's and Corbin's views on this subject, see *infra* notes 78-98 and accompanying text.

evidence of any prior agreement.”⁶¹ Whether the rule applies to all prior agreements, regardless of whether they are in written or oral form is a topic of debate. Professors Corbin⁶² and Williston,⁶³ taking the majority position, answer in the affirmative. Statements of the rule, however, exist that limit its application to prior oral agreements.⁶⁴ Furthermore, while scholars unanimously agree that the rule applies to alleged agreements made prior to the writing, but not to alleged agreements made subsequent thereto,⁶⁵ disagreement exists about the proper treatment of agreements allegedly made contemporaneously with the writing.⁶⁶ Professor Williston⁶⁷ and the Uniform Commercial Code⁶⁸ treat contemporaneous oral terms like prior terms, but treat contemporaneous written terms as part of the integrated writing. Professor Corbin, on the other hand, argues that terms are either prior or subsequent, and deems the use of the word “contemporaneous” to be an obfuscation of the integration issue.⁶⁹

Further and more serious disagreement arises during actual application of the rule to particular fact situations, and concerns

61. Murray, *supra* note 17, at 337.

62. See 3 A. CORBIN, *supra* note 48, § 573, at 357, for the following statement of the rule:

When two parties have made a contract and have expressed it in a writing to which they have both assented as the complete and accurate integration of that contract, evidence, *whether parol or otherwise*, of antecedent understandings and negotiations will not be admitted for the purpose of varying or contradicting the writing.

Id. (emphasis added).

63. “Briefly stated, this rule requires, in the absence of fraud, duress, mutual mistake, or something of the kind, the exclusion of extrinsic evidence, *oral or written*, where the parties have reduced their agreement to an integrated writing.” 4 S. WILLISTON, A TREATISE ON THE LAW OF CONTRACTS § 631, at 948-49 (3d ed. W. Jaeger 1961) (emphasis added).

64. See, e.g., Sweet, *supra* note 21, at 1036 (“The parol evidence rule determines the provability of a prior or contemporaneous oral agreement when the parties have assented to a written agreement.”). Professor McCormick premises the entire discussion of the rule on its applicability only to prior oral agreements. See C. McCORMICK, *supra* note 26, §§ 210-22.

65. See, e.g., J. CALAMARI & J. PERILLO, *supra* note 20, § 3-6, at 113-14; J. WHITE & R. SUMMERS, *supra* note 27, § 2-9, at 78; Sweet, *supra* note 21, at 1042.

66. J. CALAMARI & J. PERILLO, *supra* note 20, § 3-2, at 100.

67. 4 S. WILLISTON, *supra* note 63, § 628, at 913-15, § 631, at 952-53; see also J. CALAMARI & J. PERILLO, *supra* note 20, § 3-2, at 100-01.

68. U.C.C. § 2-202 (1977) provides in part:

Terms with respect to which the confirmatory memoranda of the parties agree or which are otherwise set forth in a writing intended by the parties as a final expression of their agreement with respect to such terms as are included therein may not be contradicted by evidence of *any prior agreement or of a contemporaneous oral agreement . . .*

Id. (emphasis added). See also J. WHITE & R. SUMMERS, *supra* note 27, § 2-10, at 78-79; Wallach, *supra* note 17, at 665.

69. 3 A. CORBIN, *supra* note 48, § 577, at 400-01; see also J. CALAMARI & J. PERILLO, *supra* note 20, § 3-2, at 101; Murray, *supra* note 17, at 337 n.3.

the issues of the "finality" and "completeness" of a contract. Frequently, the "overriding issue in parol evidence disputes"⁷⁰ is whether the parties intended the written document to be a final and complete statement of their agreement, or a final statement of a part of their agreement, or perhaps nothing more than a memorandum not intended to be a final expression of their agreement at all.⁷¹ The issue of "finality" is crucial because the rule does not apply to tentative or preliminary writings.⁷² Scholars generally agree that parties may introduce any relevant evidence on the question of whether the parties intended the writing to be final.⁷³ While the question of finality often is cast as a question of law since the trial judge makes the determination, the question actually concerns an interpretation of intention, and thus is a question of fact.⁷⁴

The question of completeness is of paramount importance because courts generally draw a distinction between writings that are a partial integration of the parties' agreement and writings that are a total integration of the agreement. Courts characterize as total integrations those writings that they determine to be both final and complete. Thereafter, evidence of any prior agreements may not contradict or supplement the totally integrated writing.⁷⁵ Courts characterize as partial integrations those writings that are final but incomplete. Evidence of prior agreements may not contradict such writings, although evidence of consistent additional terms may supplement such incomplete writings.⁷⁶ The issue of completeness, like the issue of finality, is a question of law in the sense that the trial judge decides the issue.⁷⁷

The method that courts employ in determining whether a writing is complete is the subject of an intense conflict, with Professors Williston and Corbin heading the two sides of debate. Both scholars agree that the "intent" of the parties should be the

70. Wallach, *supra* note 17, at 655.

71. *Id.*

72. J. CALAMARI & J. PERILLO, *supra* note 20, § 3-2, at 101.

73. See, e.g., *id.*; 3 A. CORBIN, *supra* note 48, § 588, at 528-29.

74. J. CALAMARI & J. PERILLO, *supra* note 20, § 3-2, at 101-02.

75. *Id.* § 3-2, at 101.

76. *Id.*; see also 4 S. WILLISTON, *supra* note 63, § 636; Sweet, *supra* note 21, at 1038. Corbin agrees in substance but would abandon the term "partial integration" on the ground that parties rarely intend incomplete writings to be final. 3 A. CORBIN, *supra* note 48, § 581, at 441. See *infra* notes 119-23 and accompanying text for a discussion of the confusion surrounding the terms "contradict" and "consistent" in partial integration cases.

77. J. CALAMARI & J. PERILLO, *supra* note 20, § 3-2, at 103.

determinative element,⁷⁸ but the two use "intent" in very dissimilar ways.⁷⁹ Williston stresses the appearance of the writing as determinative in discerning the parties' intent;⁸⁰ he argues that the consideration of extrinsic evidence in determining the parties' intent emasculates the parol evidence rule.⁸¹ According to Williston, if the writing includes a merger clause (also called an "integration" or "entire contract" clause),⁸² then the writing is per se a total integration unless (1) the writing is obviously incomplete, (2) the merger clause is the product of fraud or mistake, or (3) other reasons exist that justify setting the contract aside.⁸³ If no merger clause exists and the writing is obviously incomplete on its face or only addresses the obligations of one party, the parties may introduce consistent additional terms.⁸⁴ If, however, the writing is apparently complete on its face, despite the absence of a merger clause, a court will treat it as a total integration unless the proffered additional terms are such as might "naturally and normally"⁸⁵ be made but not included in a written contract by similarly situated parties.⁸⁶ Parties always may prove the existence of "collateral" contracts, which are separate agreements supported by separate consideration, despite the apparent completeness of a particular writing.⁸⁷ Scholars have criticized Williston's approach

78. See 3 A. CORBIN, *supra* note 48, §§ 573-96; 4 S. WILLISTON, *supra* note 63, § 633.

79. J. CALAMARI & J. PERILLO, *supra* note 20, § 3-3.

80. Wallach, *supra* note 17, at 659. In this respect the Williston approach resembles the now-defunct "four corners" doctrine. *Id.* For a general discussion of the "four corners" doctrine, see *id.* at 656-57.

81. See 4 S. WILLISTON, *supra* note 63, § 633, at 1014. Determining the parties' intent through consideration of extrinsic evidence emasculates the rule because the existence of a collateral oral agreement conclusively indicates a partial integration, leaving as the only issue whether the parties made the alleged agreement. This issue presents a question of fact for the jury and, thus, eliminates the special protection that the trial judge should give the writing. J. CALAMARI & J. PERILLO, *supra* note 20, § 3-3, at 104.

82. For a discussion of merger clauses and their general operation in parol evidence cases, see generally J. WHITE & R. SUMMERS, *supra* note 27, § 2-12; Sweet, *supra* note 21, at 1037, 1045; Wallach, *supra* note 17, at 677-78; Comment, *The "Merger Clause" and the Parol Evidence Rule*, 27 TEX. L. REV. 361 (1949).

83. 4 S. WILLISTON, *supra* note 63, § 633, at 1014, § 634, at 1017-20; see also J. CALAMARI & J. PERILLO, *supra* note 20, § 3-3, at 104.

84. 4 S. WILLISTON, *supra* note 63, § 633, at 1014-15, §§ 636, 645; see also J. CALAMARI & J. PERILLO, *supra* note 20, § 3-3, at 105.

85. See Wallach, *supra* note 17, at 658. This test also is known as the "reasonable man" test. *Id.*; see also C. McCORMICK, *supra* note 26, § 216, at 441.

86. See 4 S. WILLISTON, *supra* note 63, §§ 638-39; see also J. CALAMARI & J. PERILLO, *supra* note 20, § 3-3, at 105.

87. J. CALAMARI & J. PERILLO, *supra* note 20, § 3-3, at 105; C. McCORMICK, *supra* note 26, § 211, at 431; Sweet, *supra* note 21, at 1038.

and its emphasis on the writing because that approach treats the parol evidence rule as a rule of form;⁸⁸ such treatment subordinates the true intent of the parties to their legally presumed intent.

Corbin argues that a court can never determine the true intent of the parties by confining its attention to the parties' writing.⁸⁹ He clearly expressed his scorn for the Willistonian approach:

Since the very existence of an "integration" . . . is dependent on what the parties thereto say and do, (necessarily extrinsic to the paper instrument) at the time they draw that instrument "in usual form," are we to continue like a flock of sheep to beg the question at issue, even when its result is to "make a contract for the parties," one that is vitally different from the one they made for themselves?⁹⁰

Corbin thus plainly is concerned with ascertaining the true intent of the parties to the writing:⁹¹ did the parties subjectively agree that the writing was a complete statement or merely a partial statement of their contract?⁹² A trial court adopting the Corbin approach would look at all available extrinsic evidence to divine the parties' true intent on the integration issue;⁹³ it would test the admissibility of a proffered parol term "solely by the credibility of the party seeking to introduce the term."⁹⁴ Corbin's emphasis on the true intent of the parties, however, "emasculates the traditional parol evidence rule"⁹⁵ by depriving the trial judge of the "judicial trump card"⁹⁶ afforded by the more traditional approaches that allow a judge to exclude suspect evidence from the jury without appearing to rule on its credibility.⁹⁷ Consequently, the Corbin approach largely eliminates the "rule of form" aspects of the rule by denying writings the special treatment that the Willistonian approach affords them.⁹⁸

88. See J. CALAMARI & J. PERILLO, *supra* note 20, § 3-3, at 108. Courts following the Willistonian approach see the essential purpose of the parol evidence rule as protecting the integrity of written contract. *Id.* § 3-3, at 107.

89. "[I]t can never be determined by mere interpretation of the words of a writing whether it is an 'integration' of anything, whether it is 'the final and complete expression of the agreement' or is a mere partial expression of 'the agreement.'" 3 A. CORBIN, *supra* note 48, § 581, at 442.

90. *Id.* § 582, at 463.

91. See J. CALAMARI & J. PERILLO, *supra* note 20, § 3-3, at 106.

92. See Wallach, *supra* note 17, at 663.

93. See *id.* at 644.

94. *Id.*

95. J. CALAMARI & J. PERILLO, *supra* note 20, § 3-3, at 106 n.45.

96. See Wallach, *supra* note 17, at 665.

97. See *id.*; see also *supra* notes 36-38 and accompanying text.

98. Wallach, *supra* note 17, at 664.

Despite the criticism that the Corbin approach emasculates the parol evidence rule, Corbin's arguments clearly have influenced recent formulations of the rule and have resulted in a general shift favoring attempts to ascertain the parties' true intent on the integration issue. For example, the Uniform Commercial Code's version of the rule significantly increases the chances that the trier of fact will consider evidence of extrinsic terms.⁹⁹ The Code suggests that partial integration is the norm¹⁰⁰ by initially focusing on whether the writing was "*intended* by the parties as a final expression of their agreement *with respect to such terms as are included therein . . .*"¹⁰¹ A court will bar evidence of consistent additional terms only when it "finds the writing to have been intended also as a complete and exclusive statement of the terms of the agreement."¹⁰² Furthermore, a court will deem the writing "complete and exclusive" and bar evidence of additional terms only if the court finds that the purported additional terms "if agreed upon, . . . would *certainly* have been included in the document."¹⁰³ This standard for a complete integration is much narrower than Williston's "naturally and normally" test.¹⁰⁴ The Code also allows evidence of course of dealing, course of performance, and usage of trade to explain or supplement the terms included in the writing,¹⁰⁵ even when this evidence appears to contradict apparently unambiguous terms in the writing.¹⁰⁶

The *Restatement (Second) of Contracts* likewise reflects the influence of Corbin's ideas.¹⁰⁷ It provides that "a writing which in view of its completeness and specificity reasonably appears to be a complete agreement . . . is taken to be an integrated agreement unless it is established by other evidence that the writing did not

99. See *id.* at 666-67.

100. *Id.* at 666; see also J. CALAMARI & J. PERILLO, *supra* note 20, § 3-7.

101. U.C.C. § 2-202 (1977) (emphasis added); see also Wallach, *supra* note 17, at 666. Note also the Code's emphasis on the parties' intent. U.C.C. § 2-202 comment 2 further exemplifies this emphasis by stating that the object of the parol evidence process is to find "the true understanding of the parties as to the agreement . . ." *Id.* (emphasis added).

102. U.C.C. § 2-202(b) (1977).

103. U.C.C. § 2-202 comment 3 (1977) (emphasis added); see also J. CALAMARI & J. PERILLO, *supra* note 20, § 3-7, at 115; J. WHITE & R. SUMMERS, *supra* note 27, § 2-10, at 79-80; Wallach, *supra* note 17, at 667-68.

104. Wallach, *supra* note 17, at 668.

105. U.C.C. § 2-202(a) (1977).

106. See J. CALAMARI & J. PERILLO, *supra* note 20, § 3-7, at 115; J. WHITE & R. SUMMERS, *supra* note 27, § 2-10, at 85-87; Wallach, *supra* note 17, at 665-66.

107. J. CALAMARI & J. PERILLO, *supra* note 20, § 3-3, at 109.

constitute a final expression."¹⁰⁸ The *Restatement* rejects the notion that the writing itself can "prove its own completeness"¹⁰⁹ and notes that "wide latitude must be allowed for inquiry into circumstances bearing on the intention of the parties."¹¹⁰ A court must determine the parties' intent to integrate by looking at "all relevant evidence,"¹¹¹ including usages of trade, course of dealing, and course of performance;¹¹² even the presence of a merger clause is not conclusive on the integration issue.¹¹³ If the alleged omitted term is one that "in the circumstances might naturally be omitted from the writing," a court will deem the writing a partial integration.¹¹⁴ That an omission "does not seem natural," however, will not prevent its admission "unless the court finds that the writing was intended as a complete and exclusive statement of the terms of the agreement."¹¹⁵

The clear thrust, then, of both the *Restatement (Second)*¹¹⁶ and the Code¹¹⁷ is to add momentum to what Williston himself perceived as the modern trend "toward increasing liberality in the admission of parol agreements."¹¹⁸ By increasing the likelihood that courts will hold writings to be partial integrations rather than complete integrations, these more liberal formulations of the rule not only provide a formal rationale for the admission of terms inadmissible under form-oriented versions of the rule but also provide an increased opportunity for *sub rosa* judicial behavior aimed at producing "just" results at the expense of consistent application of the rule. Judges confronted with alleged terms that they believe are genuine may admit such terms by finding that the parties intended their writing to be only a partial integration of their agreement, regardless of the writing's facial completeness.¹¹⁹ Of course,

108. RESTATEMENT (SECOND) OF CONTRACTS § 209(3) (1981).

109. *Id.* § 210 comment b.

110. *Id.*

111. *Id.* § 209 comment c; *see id.* § 210 comment b.

112. *Id.* § 209 comment a.

113. *Id.* § 209 comment b; *see id.* § 216 comment e.

114. *Id.* § 216(2)(b). A comment to this section observes that if "the omission seems natural in the circumstances, it is not necessary to consider further the questions whether the agreement is completely integrated and whether the omitted term is within its scope, although factual questions may remain." *Id.* at comment d.

115. *Id.* at comment d ("[T]here is no rule or policy penalizing a party merely because his mode of agreement does not seem natural to others.").

116. *See* J. CALAMARI & J. PERILLO, *supra* note 20, § 3-3, at 111.

117. *See* Wallach, *supra* note 17, at 666-67.

118. 4 S. WILLISTON, *supra* note 63, § 638, at 1045.

119. C. McCORMICK, *supra* note 26, § 212, at 432-33.

only such terms as are "consistent" with the writing are admissible, and terms which "contradict" terms in a partial integration are never admissible.¹²⁰ Whether a proffered extraneous term contradicts a term contained within the writing, however, is a difficult¹²¹ and perhaps misleading¹²² question, which itself provides fertile ground for judicial legerdemain. Courts in numerous recent cases have admitted evidence of extraneous terms on the grounds of "consistency" when any meaningful application of the consistency standard would dictate a different result.¹²³

A brief examination of the numerous judicially-created exceptions to the rule's operation gives further evidence of the modern trend toward liberality in applying the parol evidence rule and of the confusion and inconsistency surrounding its application.¹²⁴ For example, courts generally agree¹²⁵ that the parol evidence rule does not apply to proof that the contract which a writing represents is subject to a condition precedent.¹²⁶ The courts, however, should distinguish between conditions precedent to the formation of a contract and conditions precedent to the performance of a contract.¹²⁷ Courts should admit on its own merit, rather than as an exception, evidence that the failure of a condition precedent prevented the formation of a contract since the rule only applies when a court determines that a complete or partial contract exists.¹²⁸ On

120. See *supra* note 76 and accompanying text.

121. See J. WHITE & R. SUMMERS, *supra* note 27, § 2-10, at 81-84; Sweet, *supra* note 21, at 1038.

122. Professor Wigmore suggested that the only sensible inquiry concerning this aspect of the integration question is whether the proffered term is "additional" to the writing in the sense that it addresses issues not touched upon in the writing. Any additional term arguably contradicts previous terms. Parties hardly would seek to introduce extraneous terms that would have no effect on their obligations under the writing. 9 J. WIGMORE, *supra* note 8, § 2430.

123. See, e.g., *Bunge Corp. v. Recker*, 519 F.2d 449 (8th Cir. 1975); *Hunt Foods & Indus. v. Doliner*, 26 A.D. 2d 41, 270 N.Y.S.2d 937 (1966); *Recreatives, Inc. v. Travel-On Motorcycles Co.*, 29 N.C. App. 727, 225 S.E.2d 637 (1976). See Wallach, *supra* note 17, at 669-70, 674-76, for a discussion of these cases.

124. See J. CALAMARI & J. PERILLO, *supra* note 20, § 3-3, at 110, for the allegation that the genesis of at least some of these exceptions is "[d]issatisfaction with rigid application" of the rule.

125. Some courts, however, refuse to recognize the condition precedent exception on the ground that such conditions "are a type of contract term and arguably should be subject to some type of restriction on admissibility if not found in the written agreement." Wallach, *supra* note 17, at 654.

126. See J. CALAMARI & J. PERILLO, *supra* note 20, § 3-4, at 111-12; C. McCORMICK, *supra* note 26, § 221, at 449-50.

127. See J. CALAMARI & J. PERILLO, *supra* note 20, § 3-4, at 112 n.4.

128. See 3 A. CORBIN, *supra* note 48, § 589, at 530-32; Broude, *The Consumer and the*

the other hand, the parol evidence rule should apply to conditions precedent to the performance of a contract; a party should be allowed to prove such conditions only if the writing is a partial integration and the alleged conditions merely supplement, rather than contradict, the writing.¹²⁹ Of course, any condition precedent is arguably inherently inconsistent with a facially unconditional writing. Nevertheless, cases exist in which courts have admitted as consistent conditions that plainly were "contradictory" under any reasonable construction of the term.¹³⁰ In addition, the *Restatement (Second)* plainly rejects the "consistency" standard as it applies to conditions.¹³¹

Parties also may avoid the strictures of the parol evidence rule by raising a variety of defenses attacking the validity of the underlying agreement that the writing represents. These defenses include most traditional contract defenses and constitute "a complex set of subrules and exceptions that equals the parol evidence rule itself in unevenness of application and confusion."¹³² Courts may admit evidence which shows that the contract lacks consideration,¹³³ is illegal,¹³⁴ or is voidable due to fraud, duress, undue influence, or mistake.¹³⁵ Of these exceptions, the most important—not only because significant disagreement exists concerning the scope of the exception, but also because the exception serves as a convenient entree for judicial manipulation of the rule¹³⁶—is the exception for fraud.¹³⁷ Indeed, the fraud exception could become "an

Parol Evidence Rule: Section 2-202 of the Uniform Commercial Code, 1970 DUKE L.J. 881, 891.

129. See J. CALAMARI & J. PERILLO, *supra* note 20, § 3-4, at 112; Sweet, *supra* note 21, at 1040.

130. See, e.g., *Williams v. Johnson*, 229 A.2d 163, 166 (D.C. 1967) (allowing admission of a condition precedent in the face of a merger clause). For a discussion of *Williams* and other similar cases, see Broude, *supra* note 128, at 891-99.

131. See RESTATEMENT (SECOND) OF CONTRACTS § 217 comment h (1981).

132. Sweet, *supra* note 21, at 1039.

133. J. CALAMARI & J. PERILLO, *supra* note 20, § 3-4, at 112; Sweet, *supra* note 21, at 1040.

134. J. CALAMARI & J. PERILLO, *supra* note 20, § 3-4, at 112; C. McCORMICK, *supra* note 26, § 221, at 451.

135. J. CALAMARI & J. PERILLO, *supra* note 20, § 3-4, at 112; C. McCORMICK, *supra* note 26, § 221, at 450-51; Sweet, *supra* note 21, at 1039.

136. Some scholars argue that "to circumvent the rule fraud has been found and reformation has been granted in situations where these concepts are not ordinarily deemed applicable." J. CALAMARI & J. PERILLO, *supra* note 20, § 3-3, at 110. See also Childres & Spitz, *Status in the Law of Contract*, 47 N.Y.U. L. REV. 1, 26 n.118 (1972) (collecting cases in which courts have expanded the fraud and mistake defenses to admit terms under the exception to the rule).

137. J. WHITE & R. SUMMERS, *supra* note 27, § 2-11. But cf. Sweet, *Promissory Fraud*

exception that swallows up the entire parol evidence rule."¹³⁸ Although some authorities only allow proof of fraud in the execution¹³⁹ of a contract to contradict the writing,¹⁴⁰ a party generally may show proof of fraud in the inducement¹⁴¹ even though such proof contradicts an integration.¹⁴² Authorities disagree about whether to allow proof of promissory fraud¹⁴³ to contradict an integration.¹⁴⁴

Finally, another means of access for parol evidence and source of disagreement about the scope of the rule is the generally accepted principle that the rule does not apply to evidence aimed at interpretation of a writing, as long as such evidence does not vary, add to, or contradict the writing.¹⁴⁵ This principle in effect stipulates that such parol evidence is inadmissible only if the written terms are plain, clear, and unambiguous.¹⁴⁶ Authorities have criticized this corollary to the rule on the grounds that language is inherently ambiguous¹⁴⁷ and that courts which profess to find the language of a writing clear and unambiguous, of necessity, are en-

and the Parol Evidence Rule, 49 CALIF. L. REV. 877 (1961) (To term the admission of evidence of fraud as an "exception" to the rule is erroneous because the rule only applies to valid contracts.).

138. J. WHITE & R. SUMMERS, *supra* note 27, § 2-11.

139. Fraud in the execution of a contract relates to deception concerning the contents of a writing. Sweet, *supra* note 137, at 888.

140. *Id.* at 887-88.

141. Fraud in the inducement exists when one party, by false representations, induces the other party to enter the contract. *Id.* at 888.

142. J. CALAMARI & J. PERILLO, *supra* note 20, § 3-4, at 112 n.77; 3 A. CORBIN, *supra* note 48, § 580.

143. Promissory fraud consists of making a promise without intending to perform. J. CALAMARI & J. PERILLO, *supra* note 20, § 3-4, at 112 n.77.

144. Some jurisdictions have held that a party may not prove the existence of a promise which contradicts an integration on a theory of promissory fraud. *Id.* A majority of jurisdictions, however, now treat the making of promises without intent to perform as the equivalent of a misrepresentation of fact and, thus, treat such misrepresentations like fraud in the inducement. *Id.* § 9-19, at 286-87; Sweet, *supra* note 137, at 888-89.

145. See C. McCORMICK, *supra* note 26, §§ 217-20; J. WHITE & R. SUMMERS, *supra* note 27, § 2-11, at 89; Sweet, *supra* note 21, at 1041.

146. See Farnsworth, *supra* note 29, at 959.

147. "The very concept of plain meaning finds scant support in semantics, where one of the cardinal teachings is the fallibility of language as a means of communication." *Id.* at 952. See Young, *Equivocation in the Making of Agreements*, 64 COLUM. L. REV. 619, 632 n.57 (1964), for this discussion:

Ambiguity in the sense of vagueness is the usual *carte d'entrée* for parol evidence, but ambiguity in this sense is a matter of degree: "as all language will bear some different meanings, some evidence is always admissible; the line of exclusion depends on how far the words will stretch, and how alien is the intent they are asked to include."

Id. (quoting *Eustis Mining Co. v. Beer, Sondheimer & Co.*, 239 F. 976, 985 (S.D.N.Y. 1917)).

gaging in an exercise in interpretation.¹⁴⁸ Numerous examples of judicial confusion about, and manipulation of, this corollary exist.¹⁴⁹ Nonetheless, Corbin has argued that the parol evidence rule should not apply to evidence aimed at issues of interpretation since a court cannot know whether proffered evidence supplements or contradicts a writing and, hence, is admissible under the rule until it ascertains the meaning of the terms contained in the writing.¹⁵⁰ Both the *Restatement (Second)*¹⁵¹ and the Uniform Commercial Code¹⁵² embrace Corbin's position and thereby significantly broaden the admissibility of parol evidence offered under the pretext of aiding the interpretive process.

Given the foregoing catalog of confusion, disagreement, and judicial manipulation, some observers of the parol evidence rule believe that a significant dichotomy often exists between the realities of the rule's operation and its various official formulations.

148. When a court makes the often repeated statement that the written words are so plain and clear and unambiguous that they need no interpretation . . . it is making an interpretation on the sole basis of the extrinsic evidence of its own linguistic experience and education, of which it merely takes judicial notice.

Corbin, *supra* note 20, at 189.

149. According to Corbin,

[t]here are many court decisions, made by highly respected courts, that are inconsistent with the repeated rule. If extrinsic evidence is admitted without objection, the trial court is never reversed for considering it in the process of interpretation. There are many cases, practically never subjected to criticism, in which the court has considered extrinsic evidence as a basis for finding that the written words are ambiguous; instead of ambiguity admitting the evidence, the evidence establishes the ambiguity. Learned judges have often differed as to whether the written words are ambiguous, each one sometimes asserting that *his* meaning is plain and clear. All that any court has to do in order to admit relevant extrinsic evidence is to assert that the written words are ambiguous; this has been done in many cases in which the ordinary reader can perceive no ambiguity until he sees the extrinsic evidence.

Id. at 161-62.

150. *Id.* at 188-90.

151. RESTATEMENT (SECOND) OF CONTRACTS § 212 comment b (1981) provides:

It is sometimes said that extrinsic evidence cannot change the plain meaning of a writing, but meaning can almost never be plain except in a context. Accordingly, the rule stated in Subsection (1) [on interpretation of integrated agreements] is not limited to cases where it is determined that the language used is ambiguous. Any determination of meaning or ambiguity should only be made in the light of the relevant evidence of the situation and relations of the parties, the subject matter of the transaction, preliminary negotiations and the statements made therein, usages of trade, and the course of dealing between the parties.

152. The actual Code provisions expressly consider only evidence relating to course of performance, usage of trade, or course of dealing. *See supra* notes 105-06 and accompanying text. U.C.C. § 2-202 comment 1(c) (1972), however, says that "[t]his section definitely rejects" any requirement that "a condition precedent to the admissibility" of such evidence "is an original determination by the court that the language used is ambiguous."

Professors Calamari and Perillo observe that: "It would . . . be a mistake to suppose that the courts follow any of these rules blindly, literally or consistently. As often as not they choose the standard or rule that they think will give rise to a just result in the particular case."¹⁵³ A study of the application of the parol evidence rule, then, must attempt to pinpoint the factors that influence judicial perception about the "justice" of a particular result. One commentator avers that "the proponent of the oral agreement will be permitted to prove it if the trial judge thinks it likely that the agreement was made and if there are no cogent reasons why it should not be enforced."¹⁵⁴ Unfortunately, since judicial manipulation of the rule varies from judge to judge, the outcome of any particular parol evidence case remains difficult to predict:

Although the outcome of a case is often correct because courts, as a rule, have a good sense of fairness, there are cases that simply come out wrong. There are non-result-oriented judges who mechanically follow cases phrasing the Rule in its traditional form. Other judges, believing the Rule expresses a sound judicial policy, may refuse to admit the testimony of the oral agreement even if they believe the agreement took place and was intended to stand.¹⁵⁵

Perhaps the most interesting suggestion about the realities of the rule's operation—a suggestion that emerges from a study of 149 parol evidence cases¹⁵⁶—is that an examination of the status of the parties to the contract and the nature of their contract better explains decisions in parol evidence cases than does the rule itself.¹⁵⁷ The authors who conducted the study found that the courts' application of the rule in cases concerning "informal" contracts or contracts involving an abuse of the bargaining process varied greatly from the application of the rule in cases involving "formal" contracts.¹⁵⁸ The authors classified as "formal" any contract in which sophisticated parties "negotiated fairly and in detail"¹⁵⁹ and deemed "informal" any contract between unsophisticated par-

153. J. CALAMARI & J. PERILLO, *supra* note 20, § 3-14.

154. Sweet, *supra* note 21, at 1045.

155. *Id.* at 1046. For a discussion of a case in which a court excluded evidence of extrinsic terms despite convincing proof of the terms' existence, see *infra* notes 437-39 and accompanying text.

156. Professors Childres and Spitz conducted the study. See Childres & Spitz, *supra* note 136.

157. See generally *id.*

158. "Precisely those factors regarded as decisive in excluding alleged oral understandings in the formal contracts category were ignored by the courts when deciding cases involving informal contracts." *Id.* at 17.

159. *Id.* at 4.

ties.¹⁶⁰ The authors concluded that

the rule functions effectively, indeed one can almost say it only functions, in cases assigned to the "formal contracts" category. In cases involving informal contracts, the parol evidence rule was almost always avoided, and where the court chose not to circumvent the rule, its decision was based on the credibility of proffered evidence, not its admissibility. As for cases in our third category—those involving an abuse of the bargaining process—the cases indicated that the parol evidence rule had no application at all with regard to the aggrieved party.¹⁶¹

C. *The Parol Evidence Rule: Criticisms and Suggestions for Improvement.*

Commentators frequently have assailed the inconsistency typifying the parol evidence rule's application and the confusion and disagreement about its underlying policy foundations. The parol evidence rule, scholars note, is difficult to apply consistently because it is riddled with exceptions¹⁶² and is not sufficiently self-executing.¹⁶³ The result is a rule that encourages litigation,¹⁶⁴ "adversely affect[s] both the counseling of clients and the litigation process,"¹⁶⁵ and hurts the administration of justice.¹⁶⁶

Questions concerning the validity of the basic assumptions upon which the rule ostensibly is premised have fueled the critical attack against the rule. Many feel that supporters of the rule have exaggerated the extent to which courts would fall prey to per-

160. *Id.*

161. *Id.* at 7.

162. See Note, *supra* note 15, at 974.

163. The rule is not sufficiently self-executing because it does not contain any internal test for determining the finality and completeness of a writing. Consequently, courts and commentators are free to devise conflicting tests. *Id.* at 973.

164. "When any rule of law is riddled through with exceptions and applications difficult to reconcile, it is believed that litigation is stimulated rather than reduced." J. CALAMARI & J. PERILLO, *supra* note 20, § 3-3, at 110-11.

165. Sweet, *supra* note 21, at 1036. Professor Sweet argues that "[t]he different policies behind the rule have varying degrees of persuasiveness in different fact situations. The ceaseless flow of parol evidence opinions and the refusal of courts to give the real reasons for their decisions contribute to litigation prediction difficulties." *Id.* at 1044.

166. As Sweet notes:

The administration of justice also suffers because of the parol evidence rule. Almost every parol evidence case involves lengthy and often fruitless bickering on the part of the attorneys. Much time is spent trying to unravel the intricacies of the rule. In addition, because the rule is phrased in admissibility terms, there is a substantial chance of a reversal of trial court decisions because the rule is often as misunderstood by appellate courts as by trial courts.

Id. at 1047; see also Hale, *The Parol Evidence Rule*, 4 OR. L. REV. 91, 91 (1925) (describing the rule as a "positive menace to the due administration of justice").

jured testimony absent the rule.¹⁶⁷ At any rate, fear of perjury does not justify refusing to enforce the parties' entire agreement.¹⁶⁸ Detractors of the rule further point out that modern juries are better educated than their predecessors,¹⁶⁹ are subject to more significant controls,¹⁷⁰ play a lesser role in dispute settlement,¹⁷¹ and decide issues of credibility in many important contexts.¹⁷² Thus, a primary justification of the rule—the rule as a jury control device¹⁷³—may be unsupportable.

Indeed, the merits of each of the myriad justifications offered in support of the rule are debatable. Critics answer the argument that the rule allows judges the convenience of excluding untrustworthy evidence without branding witnesses as liars,¹⁷⁴ by arguing that the price paid for this convenience is too high.¹⁷⁵ They counter the assertion that the rule affords a degree of certainty to written contracts which is necessary to the efficient conduct of business,¹⁷⁶ by observing that the assumption that businessmen need such certainty is untested.¹⁷⁷ Besides, the critics note, the numerous ways of avoiding the rule and the various ways in which courts apply the rule operate to decrease the reliability of written contracts.¹⁷⁸ In any event, detractors feel that courts should

167. See *supra* notes 27-38 and accompanying text.

168. J. CALAMARI & J. PERILLO, *supra* note 20, § 3-3, at 109-10.

169. See C. McCORMICK, *supra* note 26, § 216, at 441-42; J. WHITE & R. SUMMERS, *supra* note 27, § 2-9, at 77-78.

170. See J. WHITE & R. SUMMERS, *supra* note 27, § 2-9, at 77-78; Sweet, *supra* note 21, at 1056; Note, *supra* note 15, at 987.

171. See Sweet, *supra* note 21, at 1054 & n.79.

172. Professor Sweet questioned whether parol evidence cases would prove any more difficult for the jury to decide than cases concerning construction accidents, consumer injuries, or gift tax cases—all areas in which the jury exercises great control. See Sweet, *supra* note 21, at 1055; see also J. WHITE & R. SUMMERS, *supra* note 27, § 2-9, at 75-76; Note, *supra* note 15, at 985.

173. See *supra* notes 30-35 and accompanying text.

174. See *supra* notes 36-38 and accompanying text.

175. As Professor Sweet pointed out, "[p]roper issues may be missed or ignored, and the real reasons for the decision may not be given" when judges exclude evidence from the outset. Sweet, *supra* note 21, at 1057.

176. See *supra* notes 44-47 and accompanying text.

177. "Like most of the law's basic assumptions this one has never been tested by any survey of the actual effects in business of the presence or absence of such assurance." See C. McCORMICK, *supra* note 26, § 210. Indeed, as Professor McCormick notes, "[p]erhaps the special protection [afforded by the rule] is no longer needed. The telephone, and the urgent call for high speed in certain types of important transactions, such as security trading, have accustomed businessmen to rely upon word of mouth, and dispense with the safeguard of writing." *Id.* § 216, at 441-42.

178. "Because it is impossible to forecast whether or not the facts of a given transaction will come within one of the exceptions or various tests of the rule, the assumption that

subordinate the goal of certainty to the goal of effectuating the parties' actual intentions in entering into their agreement.¹⁷⁹ Additionally, critics argue that the rule's emphasis on the reliability of written evidence is misplaced because writings may be forgeries or otherwise inaccurate.¹⁸⁰ They further argue that viewing the rule as a rule of form¹⁸¹ is an unsatisfactory justification for its existence, since the rule really has not performed the "channeling" function that rules of form supposedly accomplish.¹⁸² Finally, critics note that treating the parol evidence rule as a rule of form, in some cases, can extend unjustifiably the scope of the Statute of Frauds.¹⁸³

For the purposes of determining whether the parol evidence rule is a likely candidate for the ministrations of promissory estoppel, the most significant criticisms of the rule attack the rule's potential for working injustice. People do enter parol agreements that they do not include in their written agreements. To the extent that the rule operates to bar proof of such agreements, it frustrates the true intentions of the parties¹⁸⁴ and results in courts enforcing a contract to which neither party agreed.¹⁸⁵ In many instances, a

the rule is indispensable to business stability is specious." Note, *supra* note 15, at 983.

179. For example, Professor Sweet questions the priority of the rule's objectives:

How clear is the need to protect writings from gullible or soft hearted juries or judges? In an era dominated by adhesion contracts, inequality of bargaining power and the pervasive use of liability limitations and exculpations, such commercial certainty should be subordinate to the protection of reasonable expectations. The law should be more concerned with protecting the *actual* agreement of the parties than with protecting [a] written agreement that *appears* to constitute the entire agreement.

Sweet, *supra* note 21, at 1056.

180. See, e.g., J. WHITE & R. SUMMERS, *supra* note 27, § 2-9, at 77-78.

181. See *supra* notes 39-43 and accompanying text.

182. Indeed, "[d]espite the long existence of the parol evidence rule, contracts which are partially written and partially oral are not uncommon." Note, *supra* note 15, at 983. The same author also observes that "informal business transactions between friends or long-time business associates are likely to involve 'understandings' between the parties that are not reduced to writing." *Id.*; see also Sweet, *supra* note 21, at 1036, 1047.

183. Since the Statute of Frauds never has required the entire agreement to be expressed in the memorandum, the treatment of the parol evidence rule as a rule of form effectively extends the scope of the Statute. See Sweet, *supra* note 21, at 1054. On the "channeling" function of rules of form, see *infra* notes 402-03 and accompanying text.

184. See *supra* notes 88-90 and accompanying text; see also Sweet, *supra* note 21, at 1058; Wallach, *supra* note 17, at 653; Note, *supra* note 15, at 974.

185. One observer notes that

[s]ince parties *do* make oral agreements outside their written contract and *do* use words in other than the usual sense, the exclusion of this evidence by the parol evidence rule may force upon the parties a contract that they never intended to make. Thus, because the parol evidence rule may exclude as much truthful testimony as it does perjurious testimony, the rule constitutes a major source of injustice in contract

party may rely on oral assurances that the writing's failure to reflect accurately all the terms of the agreement is "no problem."¹⁸⁶ The tendency of the rule to favor unduly the party relying on the writing¹⁸⁷ may operate particularly to the detriment of consumers, who may be unable to understand the terms of the contract even if they actually read them.¹⁸⁸ In any event, consumers are unlikely to understand the significance of technical devices such as merger clauses that they are likely to encounter in standardized form contracts.¹⁸⁹

A plethora of proposed alternatives or modifications to the parol evidence rule has arisen from the stinging criticism surrounding the rule. Some commentators conclude that the most expeditious way of overcoming the numerous difficulties that the rule poses is simply to discard the rule as a relic of the past that has outlived its original purposes and causes more problems than it solves.¹⁹⁰ Given the rule's long existence, such a resolution is unlikely,¹⁹¹ although as Justice Hohnes once observed:

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 limitation of the past.¹⁹²

Others propose less drastic means than complete abrogation for reducing the dilemmas associated with the rule's operation. They suggest that the party standing on the writing bear the burden of proving its completeness and finality,¹⁹³ or that the rule operate as

law.

Note, *supra* note 15, at 974-75 (emphasis added). See also *supra* note 90 and accompanying text.

186. See C. McCORMICK, *supra* note 26, § 210, at 428 n.6.

187. See J. WHITE & R. SUMMERS, *supra* note 27, § 2-9, at 770-78.

188. "In most [consumer] situations, it is unlikely that the buyer will understand the terms of the writing even if he could and did read them." Broude, *supra* note 128, at 906.

189. Professor McCormick notes that

[i]n certain types of relatively standardized transactions, particularly in the sale to consumers of appliances such as [sic] automobiles, air-conditioners, tractors, and television sets, it is customary for the standard preprinted form of contract or order blank to include a clause, frequently in fine print, providing that there are no promises, warranties, conditions or representations not appearing in the writing. Seldom are these provisions actually read by the purchaser and even less often would he understand their effect upon the statements and promises made by the salesman.

C. McCORMICK, *supra* note 26, § 222.

190. See Note, *supra* note 15, at 976, 987-88.

191. "The rule . . . has existed so long that its total abandonment is not likely even if it could be shown that it is not needed." Sweet, *supra* note 21, at 1059.

192. Holmes, *The Path of the Law*, 10 HARV. L. REV. 457, 469 (1897).

193. See J. WHITE & R. SUMMERS, *supra* note 27, § 2-9, at 78.

a rebuttable presumption that the writing is inclusive and final—a presumption that a party may challenge by clear and convincing evidence to the contrary.¹⁹⁴ Another alternative proposal is that the rule only apply to truly integrated agreements¹⁹⁵ and that the courts recognize that the number of such agreements is relatively small.¹⁹⁶ Some commentators have added further qualifications to this proposal: that courts use the rule only to strike down prior terms that directly contradict the writing,¹⁹⁷ that they never use the rule to strike down contemporaneous terms,¹⁹⁸ and that when the Statute of Frauds requires written evidence of the parties' agreement courts discard the rule entirely in favor of an inquiry into whether a memorandum sufficient to satisfy the Statute exists. The last qualification would leave to the jury the issue of whether the parties indeed made the proffered prior oral agreement.¹⁹⁹ Finally, some authorities suggest that the manner of the rule's application should vary depending upon whether the contract in question is "formal,"²⁰⁰ "informal,"²⁰¹ or involves an abuse of the bargaining process.²⁰²

The likelihood of the parol evidence rule's continued existence in some form or other forces consideration of whether the doctrine of promissory estoppel properly can play a part in ameliorating

194. See Hale, *supra* note 166, at 122; Sweet, *supra* note 21, at 1061; Note, *supra* note 15, at 986.

195. See Sweet, *supra* note 21, at 1059-60, 1063. Professor Sweet says that the hallmark of such a contract is that the parties wrote it carefully and methodically. Sweet offers numerous criteria for determining the existence of such a contract. *Id.* at 1063-67.

196. See *id.* at 1064.

197. See *id.* at 1061.

198. See *id.*

199. *Id.* at 1054.

200. See *supra* notes 158-59 and accompanying text. Professors Childres and Spitz suggest that the rule should find its fullest application to "formal" contracts, but even in such cases they feel that its application should vary depending upon the nature of the proffered agreement. Childres & Spitz, *supra* note 136, at 7-12.

201. See *supra* note 160 and accompanying text. Regarding "informal" contracts, the Childres and Spitz study concludes that

[t]he reasonable expectations of the parties at the time of entering these transactions can be reached only by shaping the parol evidence rule to an informal model, not by forcing the informal contract to abide by a rule designed for keenly negotiated, formal transactions. And shaping the rule for informal transactions seems to us to require ignoring it.

Childres & Spitz, *supra* note 136, at 17.

202. See *supra* note 158 and accompanying text. In situations in which one party has abused the bargaining process, Childres and Spitz conclude that "[a]ny person who alleges inferior bargaining position or an abuse of bargaining power should and usually does get his evidence to the trier of fact." Childres & Spitz, *supra* note 136, at 24.

both the confusion and the potential for injustice associated with the rule. The resolution of this issue, however, first requires an exploration of the promissory estoppel doctrine's evolution in general, and, in particular, its evolution as a device for circumventing the Statute of Frauds.

III. PROMISSORY ESTOPPEL—THE EVOLUTION OF THE DOCTRINE

The "bargain" theory of consideration,²⁰³ which developed during the nineteenth century²⁰⁴ served to limit a promisor's liability by denying promisees recovery for detrimental, but unbar-gained-for, reliance on a promise.²⁰⁵ Courts applying this doctrine often reached seemingly unfair results.²⁰⁶ Early twentieth century jurists, for a multiplicity of reasons,²⁰⁷ afforded more protection to a promisee's reliance interests than their predecessors were willing to grant.²⁰⁸ These jurists countered the classic response that reli-

203. According to bargain theory, the consideration given on one side of a contractual obligation must be the "price" of the consideration given on the other side of the agreement. G. GILMORE, *THE DEATH OF CONTRACT* 19-21 (1974). Justice Holmes formulated the classic statement of the bargain idea:

[I]t is the essence of a consideration, that, by the terms of the agreement, it is given and accepted as the motive or inducement of the promise. Conversely, the promise must be made and accepted as the conventional motive or inducement for furnishing the consideration. The root of the whole matter is the relation of reciprocal inducement, each for the other, between consideration and promise.

O. HOLMES, *THE COMMON LAW* 230 (1963).

204. The "bargain theory" arguably did not even exist prior to the nineteenth century. See G. GILMORE, *supra* note 203, at 19-21; see also Farnsworth, *The Past of Promise: An Historical Introduction to Contract*, 69 COLUM. L. REV. 576, 577 (1969). On the development of "pure" or "classical" contract theory in the nineteenth century, see generally L. FRIEDMAN, *A HISTORY OF AMERICAN LAW* 244 (1973) and M. HORWITZ, *THE TRANSFORMATION OF AMERICAN LAW* 160 (1977).

205. G. GILMORE, *supra* note 203, at 19-21.

206. See, e.g., *Thorne v. Deas*, 4 Johns. 84 (N.Y. Sup. Ct. 1809) (denying the part-owner of a sailing ship any recovery on the "gratuitous" promise of another part-owner to secure insurance to cover the vessel on its ill-fated voyage).

207. For a detailed exposition of the socio-historical considerations relating to this shift in judicial stance and their relevance to the evolution of "modern" contract law in general and promissory estoppel in particular, see Metzger & Phillips, *Promissory Estoppel and the Evolution of Contract Law*, 18 AM. BUS. L.J. 139, 141-58 (1980).

208. See Summers, *The Doctrine of Estoppel Applied to the Statute of Frauds*, 79 U. PA. L. REV. 440, 448 (1931). Professor Summers argued that "the analytical and historical schools of the past century have given way to the philosophical and sociological schools of the present and morality has re-asserted itself in the law." *Id.*; see also Note, *C & K Engineering Contractors v. Amber Steel Co.: Promissory Estoppel and the Right to Trial by Jury in California*, 31 HASTINGS L.J. 697, 711-12 (1980) (Promissory estoppel's extension of liability for promises that induce reliance is consistent with other changes in modern contract law that also recognize fairness as a substantive basis of contractual liability.); *infra* note 213.

ance upon gratuitous promises is unreasonable by simply noting the common occurrence of such reliance.²⁰⁹ They argued, furthermore, that the bargain theory ignored the promisor's role in inducing the reliance.²¹⁰ Allowing individuals who had made promises to avoid all responsibility even though the promisors knew the promises likely would induce reliance by others would countenance a manifest injustice. Their search for a flexible doctrinal device to "do justice" in reliance cases led these early twentieth century jurists to principles of equity in general,²¹¹ and equitable estoppel²¹² in particular.²¹³

209. See Seavey, *Reliance Upon Gratuitous Promises or Other Conduct*, 64 HARV. L. REV. 913, 924-25 (1951).

210. See Henderson, *Promissory Estoppel and Traditional Contract Doctrine*, 78 YALE L.J. 343, 373 (1969); Seavey, *supra* note 209, at 925 ("[O]ne who makes a promise intending not to keep it misrepresents his intent and, as in other cases of deceitful misrepresentation, it should not be a defense that the defendant succeeded in taking advantage of the plaintiff's credulity.").

211. Chancellors of courts of equity traditionally had acted as the "keeper of the king's conscience," and equity had long served to mitigate the sometimes harsh results that otherwise would flow from the strict application of statutes or common law rules. L. FRIEDMAN, *supra* note 204, at 22.

212. G. GILMORE, *supra* note 203, at 63-64. Equitable estoppel, or estoppel *in pais*, was a well-established outgrowth of the familiar equitable maxim that "he who has committed inequity shall not have equity." Annot., 56 A.L.R.3d 1037, 1040-41 (1974). Professor Pomeroy outlined the classic elements of equitable estoppel:

1. There must be conduct—acts, language, or silence—amounting to a representation or a concealment of material facts.
2. These facts must be known to the party estopped at the time of his said conduct, or at least the circumstances must be such that knowledge of them is necessarily imputed to him.
3. The truth concerning these facts must be unknown to the other party claiming the benefit of the estoppel, at the time when such conduct was done, *and at the time when it was acted upon by him*.
4. The conduct must be done with the intention, or at least with the *expectation*, that it will be acted upon by the other party, or under such circumstances that it is both natural and probable that it will be so acted upon
5. The conduct must be relied upon by the other party, and, thus relying, he must be led to act upon it.
6. He must in fact act upon it in such a manner as to change his position for the worse

3 J. POMEROY, A TREATISE ON EQUITY JURISPRUDENCE § 805, at 191-92 (5th ed. 1941) (emphasis in original). Scholars also have expressed the equitable estoppel test in terms of whether the promisor's conduct inducing reliance was "unconscionable." See L. VOLD, HANDBOOK OF THE LAW OF SALES § 15, at 93 & n.62 (2d ed. 1959).

213. As Professor Gilmore noted,

[m]any judges . . . were not prepared to look with stony-eyed indifference on the plight of a plaintiff who had, to his detriment, relied on a defendant's assurances without the protection of a formal contract. However, the new doctrine [the bargain theory of consideration] precluded the judges of the 1900 crop from saying, as their predecessors would have said a half-century earlier, that the "detriment" itself was "consideration." They had to find a new solution, or at least, a new terminology. In such a situation the word that comes instinctively to the mind of any judge is, of course, "estoppel"

. . . .

G. GILMORE, *supra* note 203, at 63-64.

By preventing a person "from denying or asserting anything to the contrary of that which by the person's own deeds, acts, or representations, has been set forward as the truth," equitable estoppel protected individuals relying on that party's representations.²¹⁴ Equitable estoppel in its original form, however, insufficiently protected reliance interests because of certain evidentiary requirements. Early statements of the doctrine required the plaintiff to prove that the defendant had actual fraudulent intent.²¹⁵ Although courts had reduced this requirement by the end of the nineteenth century and forced plaintiffs to prove only a misrepresentation that would "work a fraud" on them,²¹⁶ a more significant limitation remained. Traditionally, equitable estoppel applied only to misrepresentations of present or past facts.²¹⁷ Equitable estoppel, therefore, was theoretically inapplicable to cases in which the promisor's representations consisted solely of a promise of some future performance.²¹⁸ This "fact/promise" distinction was patently artificial²¹⁹ and undoubtedly produced unjust results.²²⁰ Therefore, that "a rule of promissory estoppel would develop in recognition of the applicability of the estoppel principle to promises" was "inevitable."²²¹

While Williston apparently was the first to use the designation "promissory estoppel,"²²² the historical origins of the doctrine lie

214. Annot., 56 A.L.R.3d 1037, 1041 (1974).

215. See Note, *Statute of Frauds—The Doctrine of Equitable Estoppel and the Statute of Frauds*, 66 MICH. L. REV. 170, 174 n.21 (1967).

216. See Note, *Part Performance, Estoppel, and the California Statute of Frauds*, 3 STAN. L. REV. 281, 290 (1951).

217. See, e.g., Henderson, *supra* note 210, at 376; Seavey, *supra* note 209, at 914.

218. See, e.g., *Bank of America v. Pacific Ready-Cut Homes, Inc.*, 122 Cal. App. 554, 562-63, 10 P.2d 478, 482 (Cal. Dist. Ct. App. 1932); *Fiers v. Jacobson*, 123 Mont. 242, 251, 211 P.2d 968, 972-73 (1949); *In re Estate of Watson*, 177 Misc. 308, 317, 30 N.Y.S.2d 577, 586-87 (Sup. Ct. 1941).

219. Promises concerning the future are difficult to distinguish from statements of present or past fact, since "every statement of the future includes some statement of present facts." See Seavey, *supra* note 209, at 922-23. The difficulty arises because "statements concerning the future and . . . promises [both] involve representations as to the present state of mind of the speaker." *Id.* at 914. As Lord Bowen noted: "[T]he state of a man's mind is as much of a fact as the state of his digestion." *Edgington v. Fitzmaurise*, 29 Ch. D. 459, 483 (1885).

220. Affording protection to promisees who rely on misrepresentations of fact while denying protection to promisees who rely on promises seems inherently unjust. See Seavey, *supra* note 209, at 914.

221. Henderson, *supra* note 210, at 376.

222. Boyer, *Promissory Estoppel: Requirements and Limitations of the Doctrine*, 98 U. PA. L. REV. 459, 459 n.1 (1950) (citing 1 S. WILLISTON, *THE LAW OF CONTRACTS* § 139, at 308 (1st ed. 1920)).

in a much earlier series of cases. Dean Boyer's masterful article on the doctrine's origins²²³ pointed to several classes of cases in which turn-of-the-century courts, by a variety of judicial devices—for instance, by finding non-existent "bargains"—protected the promisees' unbargained-for reliance, which they could not have protected by strict application of classical contract principles. Dean Boyer convincingly demonstrated that promissory estoppel principles afforded the clearest explanation for such judicial machinations.²²⁴ Dean Boyer also found historical support for promissory estoppel principles in cases concerning gratuitous bailments,²²⁵ gratuitous agency,²²⁶ waivers,²²⁷ rent reductions,²²⁸ and promises of bonuses and pensions.²²⁹ Finally, he found numerous cases purporting to apply the doctrine of equitable estoppel in instances in which the doctrine's traditional elements simply were not present.²³⁰ The

223. Boyer, *Promissory Estoppel: Principle From Precedents: Parts I & II*, 50 MICH. L. REV. 639, 873 (1952).

224. For example, Boyer noted that in charitable subscriptions cases many courts enforced promisors' pledges when charitable organizations acted in reliance upon the pledges. *Id.* at 644-46. While courts did this under a variety of guises, most often by trying to transform the parties' relationship into a "bargain" of some sort, no bargain actually existed in these cases under traditional contract principles. *Id.* at 646-49. In cases concerning oral promises to make gifts of land, numerous courts enforced such promises—despite the absence of bargained-for consideration and the promisor's failure to satisfy the Statute of Frauds's writing requirement—when the promisee (who was often the promisor's child) had taken possession and had made substantial improvements to the property. *See, e.g.*, Greiner v. Greiner, 131 Kan. 760, 293 P. 759 (1930); Evenson v. Aamodt, 153 Minn. 14, 189 N.W. 584 (1922); Seavey v. Drake, 62 N.H. 393, 394 (1882); Reid v. Reid, 115 Okla. 58, 241 P. 797 (1925). In such cases courts surmounted the absence of bargained-for consideration problem either by treating the promise as an executed gift or by finding that the promisee's expenditures in reliance amounted to "consideration" for the promisor's promise. *See Boyer, supra* note 223, at 657-62. In reliance situations courts circumvented the Statute of Frauds in the same way they did situations involving the equitable doctrine of part performance, which technically was applicable only to oral contracts to sell land and not to gratuitous oral promises to deliver title. *See id.* at 656.

225. *See, e.g.*, Coggs v. Bernard, 92 Eng. Rep. 107 (1703); Wheatley v. Low, 79 Eng. Rep. 578 (1623); *see generally* Boyer, *supra* note 223, at 665-74.

226. *See, e.g.*, Maddock v. Riggs, 106 Kan. 808, 190 P. 12 (1920); Barile v. Wright, 256 N.Y. 1, 175 N.E. 351, 245 N.Y.S. 899 (1931); Elam v. Smithdeal Realty & Ins. Co., 182 N.C. 599, 109 S.E. 632 (1921); *see generally* Boyer, *supra* note 223, at 873-83.

227. *See, e.g.*, Zarthar v. Saliba, 282 Mass. 558, 185 N.E. 367 (1933); Parish Mfg. Corp. v. Martin-Parry Corp., 293 Pa. 422, 143 A. 103 (1928); *see generally* Boyer, *supra* note 223, at 888-92.

228. *See, e.g.*, William Lindeke Land Co. v. Kalman, 190 Minn. 601, 252 N.W. 650 (1934); *see generally* Boyer, *supra* note 223, at 892-98.

229. *See, e.g.*, Langer v. Superior Steel Corp., 105 Pa. Super. 579, 161 A. 571 (1932), *rev'd on other grounds*, 318 Pa. 490, 178 A. 490 (1935); *see generally* Boyer, *supra* note 223, at 883-88.

230. *See, e.g.*, Ricketts v. Scothorn, 57 Neb. 51, 77 N.W. 365 (1898).

California Supreme Court significantly advanced this tendency to broaden the application of equitable estoppel when in *Seymour v. Oelrichs*²³¹ the court barred an employer-promisor's Statute of Frauds defense to a ten-year oral employment contract because of the employer's oral promise to reduce the contract to written form.

Additional support for the promissory estoppel principle came in a series of cases that the New York Court of Appeals decided during Justice Cardozo's tenure. In these cases the court took an "expansive" view of contract;²³² the court found "consideration" when none really existed under "bargain" principles,²³³ supplied missing contractual terms,²³⁴ implied promises in the absence of express promises,²³⁵ and enforced charitable pledges in part on promissory estoppel grounds.²³⁶ In each of these cases the court protected the reliance interests of promisees in situations in which strict application of classical contract principles would operate to deny recovery.

The widespread *de facto* and *de jure* recognition of promissory estoppel motivated the authors of the *Restatement of Contracts* to grant the doctrine formal recognition in section 90:²³⁷ "A promise which the promisor should reasonably expect to induce action or forbearance of a definite and substantial character on the part of the promisee and which does induce such action or forbearance is binding if injustice can be avoided only by enforcement of the promise."²³⁸ The language of section 90 represents a clear departure from the traditional elements of equitable estoppel.²³⁹ A promisee's reliance on a "promise" triggers the operation of section 90 without the promisee showing any misrepresentation of fact.²⁴⁰ Nonetheless, the section covers only promises likely to induce reli-

231. 155 Cal. 782, 106 P. 88 (1909). For a discussion of the case, see *infra* notes 356-61 and accompanying text.

232. See G. GILMORE, *supra* note 203, at 62, for a discussion of these cases.

233. See, e.g., *De Cicco v. Schweizer*, 221 N.Y. 431, 117 N.E. 807 (1917).

234. See, e.g., *Cohen & Sons, Inc. v. Lurie Woolen Co.*, 232 N.Y. 112, 133 N.E. 370 (1921).

235. See, e.g., *Wood v. Lucy, Lady Duff-Gordon*, 222 N.Y. 88, 118 N.E. 214 (1917).

236. See, e.g., *Allegheny College v. National Chautauqua Bank*, 246 N.Y. 369, 159 N.E. 173 (1927).

237. See G. GILMORE, *supra* note 203, at 63-64, for an account of the incorporation of § 90 into the *Restatement*. Gilmore contends that Professor Corbin essentially forced the other authors to include § 90 by citing the numerous cases in which the courts imposed promissory liability in the absence of bargained-for consideration. *Id.*

238. RESTATEMENT OF CONTRACTS § 90 (1932).

239. See *supra* note 212.

240. See *supra* notes 217-20 and accompanying text.

ance of a "definite and substantial character." Courts determine the definiteness and substantiality of reliance from the viewpoint of the promisor's reasonable expectations at the time the promise was made.²⁴¹ Whether, given the circumstances then known to him, the promisor reasonably should have expected the promise to induce significant reliance is the issue for the court's determination. While this objective test focuses on the promisor, section 90 analysis also inherently involves an inquiry into the reasonableness of the promisee's reliance.²⁴² Even given the existence of such a promise, section 90 only applies when the promisor's statements result in significant reliance. Furthermore, the reasonableness inquiry necessarily includes a judgment about the fairness of the result the promisee seeks.²⁴³ The section specifically articulates this latter consideration in its insistence that enforcement of a promise is proper only when "injustice" would result from nonenforcement. While this element expressly focuses only on the effect that nonenforcement will have on the promisee,²⁴⁴ the court impliedly considers the promisor's position in determining whether his promise was likely to induce substantial reliance by the promisee.²⁴⁵

Express recognition of the promissory estoppel principle in section 90, however, was not sufficient to effect an immediate incorporation of the newly articulated doctrine into the general body of law. For example, a few relatively recent cases exist in which the court required proof of elements properly associated with equitable estoppel before applying promissory estoppel principles.²⁴⁶ Some of these cases may result from judicial confusion²⁴⁷ or conserva-

241. See Boyer, *supra* note 222, at 461.

242. Plainly, the courts rarely, if ever, would hold a promisor to "reasonably expect" unreasonable reliance by his promisee. A comment to the *Restatement (Second)* version of § 90 notes that "the reasonableness of the promisee's reliance" is a factor for courts to consider in determining whether enforcement of the promisor's promise is necessary "to avoid injustice." *RESTATEMENT (SECOND) OF CONTRACTS* § 90 comment b (1981).

243. See Boyer, *supra* note 222, at 475. A court is unlikely to believe that enforcement of a promise is necessary to avoid injustice if the reliance upon it has been insubstantial. *Id.*

244. See *id.* at 485.

245. See *supra* note 241 and accompanying text. Arguably, a court could justify any hardship that it imposed on the promisor by enforcing his promise since the promisor should have foreseen the promisee's reliance. See Boyer, *supra* note 222, at 461-62, 485.

246. See, e.g., *Sacred Heart Farmers Coop. Elevator v. Johnson*, 305 Minn. 324, 232 N.W.2d 921 (1975); *Skillman v. Lynch*, 74 S.D. 212, 50 N.W.2d 641 (1951); see also Henderson, *supra* note 210, at 377 n.192 (collecting cases).

247. See Note, *Estoppel and the Statute of Frauds*, 26 KAN. L. REV. 327, 329-30 (1978) ("The two doctrines [equitable and promissory estoppel] are . . . so similar that many courts fail to distinguish adequately between them, and some use the terms interchangeably.").

tism,²⁴⁸ but many undoubtedly reflect legitimate judicial concerns about the potential effects of the doctrine's application. After all, in classical contract doctrine, gratuitous promises were among those least deserving of judicial protection.²⁴⁹ Some of the judicial resistance that promissory estoppel has encountered may stem from judicial doubts concerning the reasonableness of unbargained-for reliance in bargain situations.²⁵⁰

This latter concern for the unbridled use of the reliance doctrine in bargain contexts played at least a *sub rosa* role in the first major restraint that the courts imposed on the application of section 90—the inapplicability of the section to cases involving “commercial” (bargained for) promises. Judge Learned Hand’s 1933 opinion in *James Baird Co. v. Gimbel Brothers, Inc.*²⁵¹ firmly established this restraint for the next quarter-century. Judge Hand concluded that promissory estoppel did not apply to promises or offers contemplating a bargained-for exchange; such promises would become enforceable only when the promisee gave the bargained-for consideration. Promissory estoppel, he reasoned, would not apply in such situations because it applied only to purely “donative” promises for which the promisor never had anticipated receiving any exchange of consideration.²⁵²

While this restricted view of promissory estoppel found historical support in the gratuitous promise cases in which courts first applied the doctrine’s principles²⁵³ and in the knowledge that such cases were apparently the sole concern of the authors of the *Restatement* when they drafted section 90,²⁵⁴ several considerations conspired to extend promissory estoppel beyond its role as a “con-

248. “The failure of the courts to adequately distinguish promissory estoppel from either equitable estoppel or part performance is due to the fact that equitable estoppel and part performance are much more established in the law, and courts are loath to establish new exceptions.” Note, *Promissory Estoppel as a Means of Defeating the Statute of Frauds*, 44 *FORDHAM L. REV.* 114, 120 n.53 (1975).

249. See Henderson, *supra* note 210, at 346.

250. See *id.* at 360.

251. 64 F.2d 344 (2d Cir. 1933). The *James Baird* case concerned a subcontractor-defendant seeking to revoke its bid to supply materials for a construction project on the grounds of an alleged mistake after the plaintiff general contractor relied upon the subcontractor’s bid by incorporating it into the general contractor’s own bid on the prime contract. For a more complete discussion of *James Baird*, see Boyer, *supra* note 222, at 491-93.

252. 64 F.2d at 346.

253. See *supra* notes 224-29 and accompanying text.

254. See Henderson, *supra* note 210, at 345 (“The proceedings leading to the drafting of Section 90 . . . evidence concern solely with justifiable detrimental reliance on promises for which no agreed equivalent has been asked or tendered.”).

sideration substitute" in donative promise cases.²⁵⁵ Section 90's language²⁵⁶ did not—either expressly²⁵⁷ or implicitly²⁵⁸—bar the principle's application in the bargain context. Furthermore, and perhaps most importantly, injustice was as predictable a consequence of reliance in the bargain context as it was in situations concerning donative promises.²⁵⁹ These factors, and the inherently expansionist nature of the reliance principle,²⁶⁰ combined to make the eventual extension of promissory estoppel into the bargain context a foregone conclusion. Promissory estoppel finally entered into the bargain context in another subcontractor bid case, *Drennan v. Star Paving Co.*²⁶¹ In *Drennan*, Chief Justice Traynor bridged the gap between the donative promise and bargain realms by first implying a promise on a subcontractor's part not to revoke its bid due to the general contractor's foreseeable and unavoidable reliance on it, and then by invoking section 90 to prevent the revocation of this implied subsidiary promise.²⁶² The result in *Drennan* since has gained general acceptance in the courts²⁶³ and received the official sanction of the *Restatement (Second)*.²⁶⁴

Another major source of judicial resistance to promissory es-

255. See J. CALAMARI & J. PERILLO, *supra* note 20, § 6-7; see also Note, *Promissory Estoppel and the Statute of Frauds in California*, 66 CALIF. L. REV. 1219, 1222-23 (1978).

256. See *supra* note 238 and accompanying text.

257. See Boyer, *supra* note 222, at 492.

258. See *id.*; see also Note, *Promissory Estoppel, Equitable Estoppel and the Farmer as a Merchant: The 1973 Grain Cases and the U.C.C. Statute of Frauds*, 1977 UTAH L. REV. 59, 82 ("To say that reliance can serve as a basis for enforcing gratuitous promises does not exclude the use of reliance to enforce promises in other contexts.").

259. "Injustice can result where a gratuitous promise is given in connection with a commercial transaction as easily as it can in the instance of a charitable subscription." Boyer, *supra* note 222, at 492-93.

260. Observers have argued that the reliance principle and the bargain theory of consideration are not only inherently contradictory, but also ultimately may be mutually exclusive. See G. GILMORE, *supra* note 203, at 61 ("The one thing that is clear is that these two contradictory propositions cannot live comfortably together: in the end one must swallow the other up.").

261. 51 Cal. 2d 409, 333 P.2d 757 (1958). A few cases prior to *Drennan* departed from the *James Baird* rule, but without the influential effect of Justice Traynor's decision in *Drennan*. See, e.g., *Robert Gordon, Inc. v. Ingersoll-Rand Co.*, 117 F.2d 654 (7th Cir. 1941); *Northwestern Eng'g Co. v. Ellerman*, 69 S.D. 397, 10 N.W.2d 879 (1943).

262. 51 Cal. 2d at 414, 333 P.2d at 760. For a fuller analysis of *Drennan*, see Henderson, *supra* note 210, at 355-56.

263. See, e.g., *Lyon Metal Prods., Inc. v. Hagerman Constr. Corp.*, 391 N.E.2d 1152, 1154-55 (Ind. Ct. App. 1979); *E.A. Coronis Assocs. v. M. Gordon Constr. Co.*, 90 N.J. Super. 69, 216 A.2d 246 (N.J. Super. Ct. App. Div. 1966); see also Note, *supra* note 208, at 713 (Promissory estoppel "has grown from a rule which originally was thought to cover only purely gratuitous promises to a rule potentially applicable in commercial dealings.").

264. See RESTATEMENT (SECOND) OF CONTRACTS § 87(2) & illustration 6 (1981).

toppel stemmed from the fear that, in some cases, the doctrine might work an unfair hardship on promisors.²⁶⁵ This fear largely is due to the uncertainty in determining the proper measure of damages in promissory estoppel cases. Courts generally agree that promissory estoppel will not support recovery when *quantum meruit* recovery of a promisee's restitution interest alone²⁶⁶ would prevent injustice. The justification for this stance is attributable to the traditional rule that equitable relief is available only when no adequate remedy at law exists²⁶⁷ and to the notion that no further enforcement of a promisor's promise is necessary to avoid injustice if the court fairly can place the promisee in *status quo*.²⁶⁸ Often, however, a restitutionary recovery will be inadequate because the value to the promisor of the promisee's reliance will be significantly less than the promisee's losses resulting from his reliance.²⁶⁹ In such situations courts generally fully enforce the promisor's promise and award expectation damages,²⁷⁰ even though the *raison d'être* of promissory estoppel is to protect promisees' reliance interests.²⁷¹ Section 90, which makes promises that meet its criteria "binding" and tries to avoid injustice through the "enforcement of the promise,"²⁷² predictably supports the award of expectation damages. Fully enforcing the promisor's promise also is consonant with the traditional view that promissory estoppel is essentially contractual in nature, serving to provide the missing elements needed to create a binding contract.²⁷³ Traditionally, courts view

265. See Boyer, *supra* note 222, at 489; Note, *supra* note 255, at 1245; *infra* notes 278-81 and accompanying text.

266. A promisee's restitution interest is equal to the value his reliance has conferred upon the promisor. See Fuller & Perdue, *The Reliance Interest in Contract Damages: 1*, 46 *YALE L.J.* 52, 53-54 (1936).

267. See Note, *Equitable Estoppel and the Statute of Frauds in California*, 53 *CALIF. L. REV.* 590, 601 (1965).

268. See Boyer, *supra* note 222, at 485.

269. See *id.* at 486.

270. Expectation damages attempt to put plaintiffs in the position that they would be in if the defaulting party had performed the contract. See Fuller & Perdue, *supra* note 266, at 54.

271. A promisee's reliance interest equals the value of his actual change in position in anticipation of the promisor's performance. The objective that granting a reliance-based recovery serves is to return the promisee to the position he was in before the promisor made his promise. *Id.*

272. See *supra* note 238 and accompanying text.

273. See, e.g., Note, *supra* note 248, at 126 (Promissory estoppel's role as a "consideration substitute" theoretically justifies an expectation award.). Williston apparently believed that an expectation award was the only proper measure of recovery in promissory estoppel cases. See Fuller & Perdue, *supra* note 266, at 64 & n.14.

reliance damages as a creature of tort law²⁷⁴ while they view restitutionary or expectation awards as contractual remedies.²⁷⁵ Finally, in some cases a reliance-based award would not compensate adequately the promisee for his injuries,²⁷⁶ while full enforcement of the promisor's promise "satisfies the expectations which have been aroused justifiably."²⁷⁷

The "all or nothing" approach to damages in promissory estoppel cases—granting full expectation recovery to promisees at one end of the spectrum while restricting promisees at the other end to a restitutionary recovery—undoubtedly can produce unjust results.²⁷⁸ For example, if the value of the promisor's full performance greatly exceeds the promisee's reliance losses, then injustice will result since the promisor experiences more harm from full enforcement than the promisee encounters from non-enforcement.²⁷⁹ This potential for injustice suggests that a reliance-based recovery may be more appropriate in some promissory estoppel cases;²⁸⁰ several courts that have disregarded the traditional damage theory employ such a situation-sensitive approach.²⁸¹

The conceptual difficulties associated with awarding full expectation damages in promissory estoppel cases are especially apparent in the recent and highly controversial extension of promissory estoppel into the context of contract negotiations. Traditionally, courts would enforce only genuine promises by using promissory estoppel principles.²⁸² Indefinite²⁸³ or illusory²⁸⁴ promises, scholars felt, were improper subjects for the doctrine's

274. See Fuller & Perdue, *supra* note 266, at 90 n.61. Fuller and Perdue conclude, however, that reliance always has been an element of contract damages. *Id.* at 89-96.

275. See *id.*; see also Note, *supra* note 255, at 1245.

276. See Note, *supra* note 248, at 128-29.

277. Boyer, *supra* note 223, at 664.

278. See Boyer, *supra* note 222, at 487.

279. See *id.* at 489.

280. See *id.* at 497.

281. For example, the Wisconsin Supreme Court has concluded that "[w]here damages are awarded in promissory estoppel instead of specifically enforcing the promisor's promise, they should only be such as in the opinion of the court are necessary to prevent injustice. Mechanical or rule-of-thumb approaches to the damage problem should be avoided." Hoffman v. Red Owl Stores, Inc., 26 Wis. 2d 683, 701, 133 N.W.2d 267, 276 (1965); see also Goodman v. Dicker, 169 F.2d 684 (D.C. Cir. 1948); Wheeler v. White, 398 S.W.2d 93, 97 (Tex. 1965).

282. See Spooner v. Reserve Life Ins. Co., 47 Wash. 2d 454, 287 P.2d 735 (1955); see also Henderson, *supra* note 210, at 361.

283. See, e.g., Perlin v. Board of Educ., 86 Ill. App. 3d 108, 114, 407 N.E.2d 792, 798 (1980); Malaker Corp. Stockholders Protective Comm. v. First Jersey Nat'l Bank, 163 N.J. Super. 463, 479, 395 A.2d 222, 230 (N.J. Super. Ct. App. Div. 1978); Boddy v. Gray, 497 S.W.2d 600 (Tex. Civ. App. 1973).

284. See, e.g., 1A A. CORBIN, CORBIN ON CONTRACTS § 201 (1963) (Section 90 does not

application because a promise capable of enforcement did not exist in either situation. Clearly, application of promissory estoppel principles to such cases would depart from promissory estoppel's more traditional "consideration substitute" role. Furthermore, such an extension of the doctrine would be inconsistent with the traditional principles of offer and acceptance. The primary thrust of these principles is "to guarantee parties seeking an exchange extensive freedom to express, or to refuse to express, a willingness to be bound,"²⁸⁵ and "to insure that, in most instances, obligation attaches only when it has been deliberately undertaken."²⁸⁶ Nonetheless, given the tendency of twentieth century jurists to seek "just" results,²⁸⁷ the general trend "toward a fuller and wider securing of interests and hence toward a wider and fuller enforcement of promises,"²⁸⁸ and the ease of application of the broadly formulated reliance principle of section 90,²⁸⁹ that courts ultimately applied promissory estoppel principles in the negotiation context is not surprising.

The most famous example²⁹⁰ of promissory estoppel's extension into the area of contract negotiations is the Wisconsin Su-

apply to illusory promises since reliance can not transform such a promise into a true promise.).

285. Henderson, *supra* note 210, at 357.

286. *Id.* at 358.

287. *See supra* notes 207-13 and accompanying text.

288. *See* Pound, *Consideration in Equity*, 13 ILL. L. REV. 667, 679 (1918).

289. *See* Henderson, *supra* note 210, at 353-54. Professor Henderson notes that [b]ecause of the flexibility which results from the generalized phrasing of the doctrine, many courts have apparently concluded that cases can be decided more easily by the use of promissory estoppel than by consideration rules. . . . The broad language of Section 90 also enables courts to avoid struggling with the more unintelligible aspects of consideration doctrine, such as the so-called requirement of 'mutuality of obligation.'

Id.

290. For another example of the doctrine's extension into the negotiation context, see *Wheeler v. White*, 385 S.W.2d 619 (Tex. Civ. App. 1964), *rev'd*, 398 S.W.2d 93 (Tex. 1965). In *Wheeler* the court allowed a plaintiff, who tore down an existing rental property in reliance upon the defendant's promise either to secure financing for improvements upon the plaintiff's real property or to provide the financing himself if none was available from third parties, to recover damages based on his reliance losses. The court based its decision on promissory estoppel despite the defendant's argument that the plaintiff's promise was too indefinite to be enforceable—an argument that both the trial court and the intermediate appellate court had accepted. *See also* *Associated Tabulating Servs., Inc. v. Olympic Life Ins. Co.*, 414 F.2d 1306, 1310-11 (5th Cir. 1969); *Hunter v. Hayes*, 533 P.2d 952, 953 (Colo. Ct. App. 1975); *Mooney v. Craddock*, 35 Colo. App. 20, 25-26, 530 P.2d 1302, 1305 (Colo. Ct. App. 1974).

preme Court's *Hoffman v. Red Owl Stores, Inc.*²⁹¹ decision. The plaintiff in *Hoffman* sought a Red Owl franchise store and, in reliance on the defendant's promise that it ultimately would grant a franchise, sold his bakery at a loss, bought a small grocery to gain experience in the grocery business, moved his family, and bought an option on a site for the franchised store. The parties never consummated the proposed deal. When Hoffman filed suit Red Owl argued that no contract existed between the parties because the parties never reached agreement on the essential terms governing the proposed franchise relationship. The Wisconsin Supreme Court concurred with Red Owl's argument that the parties had not created a contract in the traditional sense of the term, but allowed Hoffman to recover a reliance-based award²⁹² under the theory of promissory estoppel. In the process, the court observed that "it would be a mistake to regard an action grounded on promissory estoppel as the equivalent of a breach-of-contract action,"²⁹³ and noted that nothing in the language of section 90 required a promise serving as the basis of promissory estoppel to be "so comprehensive in scope as to meet the requirements of an offer."²⁹⁴

The above-quoted language from *Hoffman* certainly suggests that promissory estoppel may be evolving into a theory of recovery independent of contract²⁹⁵—a proposition with profound implications for rules of form such as the parol evidence rule.²⁹⁶ Numerous observers have argued that the reliance principle properly resides outside the framework of contract; these observers have characterized it as a creature of tort law,²⁹⁷ as a separate theory of recovery distinct from tort or contract,²⁹⁸ as a hybrid between tort and con-

291. 26 Wis. 2d 683, 133 N.W.2d 267 (1965). For more extensive discussions of *Hoffman*, see Henderson, *supra* note 210, at 358-60, and Note, *Contracts: Reliance Losses: Promissory Estoppel as a Basis of Recovery of Breach for Agreement to Agree*, *Hoffman v. Red Owl Stores, Inc.*, 51 CORNELL L.Q. 351 (1966).

292. See *supra* note 281.

293. 26 Wis. 2d at 698, 133 N.W.2d at 275.

294. *Id.*

295. See Henderson, *supra* note 210, at 359. Henderson, however, later concludes that promissory estoppel is contractual in nature. *Id.* at 378.

296. See *infra* note 434 and accompanying text.

297. Professor Seavey notes that

[e]stoppel is basically a tort doctrine and the rationale of . . . [§ 90] is that justice requires the defendant to pay for the harm caused by foreseeable reliance upon the performance of his promise. The wrong is not primarily in depriving the plaintiff of his promised reward but in causing the plaintiff to change position to his detriment.

Seavey, *supra* note 209, at 926.

298. See Beale, *Gratuitous Undertakings*, 5 HARV. L. REV. 222, 225 (1891); Note, *Promissory Estoppel—The Basis of a Cause of Action Which is Neither Contract, Tort or*

tract,²⁹⁹ and as an "anti-contract"³⁰⁰ doctrine that "may ultimately provide the doctrinal justification for the fusing of contract and tort in a unified theory of civil obligation."³⁰¹ Other *Hoffman* explanations, however, continue to place the promissory estoppel doctrine within the traditional contract framework. Arguably, *Hoffman* merely allows the reliance that promissory estoppel protects to "substitute for" the offer and acceptance which traditional contract theory requires, in the same way that reliance substituted for consideration in the early donative promise cases.³⁰² Some commentators have asserted that *Hoffman* represents an extension of the concept of "good faith"³⁰³ beyond the performance and enforcement areas and into the context of contract negotiations.³⁰⁴ Allowing promissory estoppel to impose liability in the absence of the definite expressions of mutual intent to contract, which have been the *sine qua non* of classical contract liability, however, weakens the argument that the situation involves contract law in the traditional sense of that term.³⁰⁵ Furthermore, support for the proposition that promissory estoppel is evolving into a non-contractual theory of recovery appears in several post-*Hoffman* cases that seem to treat estoppel as a basis of liability independent of

Quasi-Contract, 40 Mo. L. Rev. 163 (1975).

299. See Shattuck, *Gratuitous Promises—A New Writ?*, 35 MICH. L. REV. 908, 942 (1937).

300. See G. GILMORE, *supra* note 203, at 61.

301. *Id.* at 90.

302. See Metzger & Phillips, *supra* note 207, at 173. Of course, since the application of promissory estoppel to donative promise cases dispenses with the element of bargain, even this limited application of promissory estoppel is a departure from traditional contract doctrine. See Annot., 115 A.L.R. 152, 154 (1938).

303. Authorities have recognized expressly the "good faith" requirement in § 1-203 of the Uniform Commercial Code and in § 205 of the *Restatement (Second) of Contracts*. Both of these provisions, however, concern good faith in the performance or enforcement of contracts; the provisions do not consider expressly good faith in contract formation. See RESTATEMENT (SECOND) OF CONTRACTS § 205 comment c (1981).

304. See Summers, "Good Faith" in *General Contract Law and the Sales Provisions of the Uniform Commercial Code*, 54 VA. L. REV. 195, 223 (1968); see also Fuller, *Consideration and Form*, 41 COLUM. L. REV. 799, 811 (1941).

305. See Metzger & Phillips, *supra* note 207, at 186. For a case in which promissory estoppel serves as a substitute for a proper acceptance, see *N. Litterio & Co. v. Glassman Constr. Co.*, 319 F.2d 736, 738-40 (D.C. Cir. 1963).

contract law³⁰⁶ and in the nontraditional approach to damages that some promissory estoppel cases³⁰⁷ and the *Restatement (Second) of Contracts*³⁰⁸ take.

Regardless of the outcome on the independent theory debate, two observations about the promissory estoppel doctrine seem undeniable. First, promissory estoppel has developed well beyond its origins and, in doing so, has gone well beyond classical contract theory. Second, the expansionist nature of the reliance principle and the general tendency of modern courts to seek "just" results suggest that further expansions of the promissory estoppel principle are likely. The *Restatement (Second) of Contracts* certainly has encouraged wider acceptance of the various applications of promissory estoppel. The new version of section 90,³⁰⁹ for example, deletes the requirement that the reliance be of a "definite and substantial character."³¹⁰ Instead, a comment to the new section 90 expressly directs courts to consider, among other things, the reasonableness of the promisee's reliance,³¹¹ its substantiality in relation to the remedy he seeks, the formality surrounding the promise's formation, and the extant evidence of the promise's existence

306. See, e.g., *Debron Corp. v. National Homes Constr. Corp.*, 493 F.2d 352 (8th Cir. 1974). In *Debron* a contractor's acceptance varied the terms of a subcontractor's bid. The contractor voluntarily dismissed the contract portion of its claim before trial and the court allowed it to proceed solely on its estoppel claim. The court held that promissory estoppel could serve as the basis for a separate cause of action in which proof of all of the elements required in a breach of contract action would be unnecessary. See *id.* at 357; see also *Allen v. A.G. Edwards & Sons, Inc.*, 606 F.2d 84, 87 (5th Cir. 1979); *Northwestern Bank of Commerce v. Employers' Life Ins. Co. of Am.*, 281 N.W.2d 164 (Minn. 1979).

307. See *supra* note 281 and accompanying text.

308. See *infra* notes 315-17 and accompanying text. Nonetheless, the drafters of the *Restatement (Second)* apparently still view promissory estoppel as a contract law doctrine. The drafters include § 90 under the heading "Contracts Without Consideration," and a comment to the section notes that "[a] promise hindering under this section is a contract . . ." RESTATEMENT (SECOND) OF CONTRACTS § 90 comment d (1981).

309. RESTATEMENT (SECOND) OF CONTRACTS § 90 (1981) provides:

(1) A promise which the promisor should reasonably expect to induce action to forbearance on the part of the promisee or a third person and which does induce such action or forbearance is binding if injustice can be avoided only by enforcement of the promise. The remedy granted for breach may be limited as justice requires.

Id. For a complete discussion of this new version of § 90, see generally Knapp, *supra* note 5.

310. See *supra* note 238 and accompanying text. One authority suggests that the drafters deleted this language to accommodate partial enforcement of promises in promissory estoppel cases. See Knapp, *supra* note 5, at 58.

311. This concern about the reasonableness of the promisee's reliance was already an inherent part of the original section's requirement that the promisor foresee the likelihood of definite and substantial reliance upon his promise. See *supra* note 242 and accompanying text.

in deciding whether to grant enforcement.³¹² In addition, the new section 90 adds no requirement that the promise inducing reliance be so definite as to be enforceable as a contract, despite the stir that *Hoffman v. Red Owl Stores* created.³¹³ Indeed, an illustration accompanying the new section expressly sanctions the *Hoffman* result.³¹⁴ More importantly, by expressly stating that the remedy under the section "may be limited as justice requires,"³¹⁵ the drafters specifically have sanctioned an award of reliance-based damages in promissory estoppel cases when such an award is necessary to avoid injustice to promisors.³¹⁶ The drafters thus have removed a significant stumbling block to broader judicial acceptance of the reliance principle.³¹⁷

IV. PROMISSORY ESTOPPEL AND THE STATUTE OF FRAUDS

A. *The Statute of Frauds*

The British Parliament passed the original Statute of Frauds³¹⁸ when the law of contract was in its infancy³¹⁹ and when a primitive law of evidence excluded the testimony of persons with an interest in the outcome of the case.³²⁰ Since the evidentiary rule prevented even parties to the case from testifying, a significant danger existed that the courts would impose contract liability on defendants for contractual obligations to which they never had assented.³²¹ That courts of the era had minimal control over jury verdicts and jurors were free to disregard the evidence and rule in accordance with their own knowledge of the facts exacerbated this danger.³²² Thus, Parliament sought to achieve two goals when it passed the original Statute in 1677:³²³ the protection of defendants

312. See RESTATEMENT (SECOND) OF CONTRACTS § 90 comment b (1981).

313. See *supra* notes 291-304 and accompanying text.

314. See RESTATEMENT (SECOND) OF CONTRACTS § 90 comment d, illustration 10 (1981).

315. See *supra* note 309.

316. See *supra* notes 278-79 and accompanying text. A comment to the section, however, indicates that while "relief may sometimes be limited to restitution or to damages or specific relief measured by the extent of the promisee's reliance," "full-scale enforcement by normal remedies is often appropriate." RESTATEMENT (SECOND) OF CONTRACTS § 90 comment d (1981).

317. See *supra* note 265 and accompanying text.

318. An Act for Prevention of Frauds and Perjuries, 1677, 29 Car. 2, ch.3.

319. Summers, *supra* note 208, at 441.

320. 6 W. HOLDSWORTH, A HISTORY OF ENGLISH LAW 387-89 (1924).

321. Summers, *supra* note 208, at 441; see also Willis, *The Statute of Frauds—A Legal Anachronism*, 3 IND. L.J. 427, 430 (1928).

322. 6 W. HOLDSWORTH, *supra* note 320, at 388.

323. See generally Costigan, *The Date and Authorship of the Statute of Frauds*, 26

from perjured testimony³²⁴ and the placement of a "curb [on] the power of the juries."³²⁵ The means that Parliament chose to attain these ends was a writing requirement for certain classes of contractual obligations.³²⁶

Almost from its inception, scholars have subjected the Statute to virulent criticism. These detractors have argued that the Statute generates as many "frauds and perjuries" as it prevents.³²⁷ They also have castigated the Statute as the source of endless litigation.³²⁸ Commentators also frequently suggest that modern developments in the law of contracts³²⁹ and evidence,³³⁰ when combined with the substantial control that modern courts can exercise over jury verdicts,³³¹ have eclipsed the major reasons for the Statute's enactment and have rendered it a proper subject for repeal.³³²

The most telling complaints about the Statute's operation, however, have focused on its potential for working injustice. The parties to an oral agreement within the ambit of the Statute may be ignorant of the Statute's existence or of its relevance to their transaction.³³³ Their consequent failure to reduce their agreement

HARV. L. REV. 329 (1913) (noting the uncertainty surrounding the authorship and the date of enactment of the Statute). For an amusing account of one of the early cases that reputedly moved Parliament to take action, see J. WHITE & R. SUMMERS, *supra* note 27, § 2-1.

324. See Summers, *supra* note 208, at 441.

325. See *id.* at 458; see also Costigan, *supra* note 323, at 343.

326. The most important types of obligations requiring written evidence were contracts concerning an interest in land (except leases not exceeding three years in length), promises by an executor or administrator to answer for a decedent's debts out of his personal estate, promises by persons to pay for the debts of another, contracts made on consideration of marriage, bilateral contracts not capable of execution within one year of their making, and contracts for the sale of goods of a value exceeding ten pounds. See generally 6 W. HOLDSWORTH, *supra* note 320, at 384-87.

327. See, e.g., *Child v. Godolphin*, 21 Eng. Rep. 181 (1723); Stephen & Pollock, *Section Seventeen of the Statute of Frauds*, 1 L.Q. REV. 1 (1885); Sunderland, *A Statute for Promoting Fraud*, 16 COLUM. L. REV. 273 (1916).

328. "There is no other statute that has been the source of so much litigation." J. SMITH, *THE LAW OF FRAUDS* 327 (1907).

329. See Willis, *supra* note 321, at 431.

330. Some authorities have characterized § 17 of the original Statute [the sale of goods provision] as "a relic of times when the best evidence on such subjects was excluded on a principle now exploded." Stephen & Pollock, *supra* note 327, at 6.

331. Professor Willis notes that "with the control exercised by the courts over juries in modern times the danger that juries will hold people liable on promises they never made is better protected by such court control than by a Statute of Frauds." Willis, *supra* note 321, at 430. Cf. *supra* notes 169-73 and accompanying text (parol evidence context).

332. Justice Stephen, speaking of the sale of goods provision of the original Statute, observed that "it should be thrown out of the window . . . the 17th section should be repealed, and the cases upon it be consigned to oblivion." Stephen & Pollock, *supra* note 327, at 5.

333. Justice Stephen noted that, despite the Statute's long tenure, "[i]n the great

to written form, therefore, may operate to preclude proof of the bona fide agreement that they actually made.³³⁴ Indeed, defendants in modern times use the Statute to defend against agreements that they voluntarily assumed much more frequently than they use it to disprove fictitious agreements that perjurious plaintiffs fabricate.³³⁵ These negative effects of the Statute have generated suggestions that the Statute's writing requirement is ill-suited to its avowed purpose of preventing fraud³³⁶ and that courts should utilize other, more satisfactory, devices in grappling with the evidentiary problems inherent in oral agreements.³³⁷

The Statute, while it retains considerable vitality, has not emerged unscathed in the face of such continuing criticism. Recent versions of the Statute, in particular section 2-201 of the Uniform Commercial Code, have made significant attempts at reducing its potential for fraud,³³⁸ and courts have reacted to the Statute's potential for injustice by carving out numerous "exceptions" to its operation³³⁹ under the guise of statutory "interpretation." Such ju-

mass of cases the contracting party is as unconscious of the existence of the Statute of Frauds as of the pressure of the atmosphere." *Id.* at 3.

334. See Willis, *supra* note 321, at 539; see also Stephen & Pollock, *supra* note 327, at 3.

335. Professors White and Summers note that

[t]he possibility that plaintiffs [might] conjure up forged writings and perjured oral contracts out of whole cloth is unreal. These plaintiffs would be most unlikely to survive cross-examination, motions for directed verdict, and a jury's own scrutiny. But the possibility that defendants might get out of actual contracts simply for lack of a signed writing is not unreal at all.

J. WHITE & R. SUMMERS, *supra* note 27, § 2-8, at 74; see also Note, *supra* note 267, at 591. For a classic judicial response to such a use of the Statute, see Chief Judge Campbell's statement in *Sivewright v. Archibald*, 20 L.J.Q.B. (n.s.) 529, 536 (1851):

I regret to say that the view which I take of the law in this case compels me to come to the conclusion that the defendant is entitled to our judgement, although the merits are entirely against him; although, believing that he had broke his contract, he could only have defended the action in the hope of mitigating the damages, and although he was not aware of the objection on which he now relies till within a few days before the trial.

336. "A true 'means to an end' surely should not serve commonly as a means to disserve that end Yet, centuries of experience and tons of case law testify that a statute of frauds can be an instrument of perjury and fraud." J. WHITE & R. SUMMERS, *supra* note 27, § 2-8, at 74.

337. See, e.g., Summers, *supra* note 208, at 464 (hold plaintiffs seeking to prove the existence of oral contracts to the "beyond a reasonable doubt" standard of proof); Note, *supra* note 267, at 609 (require defendants to deny the existence of an oral contract under oath as a precondition of raising the Statute as a defense).

338. Corbin, *The Uniform Commercial Code—Sales; Should It Be Enacted?*, 59 YALE L.J. 821, 829 (1950). For a detailed exposition of this aspect of § 2-201, see Metzger & Phillips, *Promissory Estoppel and Section 2-201 of the Uniform Commercial Code*, 26 VILL. L. REV. 63, 69-74 (1980-1981).

339. These exceptions include the part performance doctrine, quasi-contract, fraud,

dicial incursions on the Statute's province have been so substantial that one legal historian claimed that "[i]n one sense, the statute of frauds [is] hardly a statute at all. It [is] so heavily warped by 'interpretation' that it [has] become little more than a set of common-law rules, worked out in great detail by the common-law courts."³⁴⁰ The courts also have engaged in some rather questionable feats of judicial legerdemain when strict construction of the Statute would produce injustice.³⁴¹ Given this history of judicial circumvention of the strictures of the Statute and the inexorable advance of the doctrine of promissory estoppel,³⁴² courts' incorporation of promissory estoppel into the judicial arsenal of devices designed to circumvent the Statute was inevitable.

B. *Promissory Estoppel and the Statute of Frauds*

The equitable doctrine of part performance was the first judicially-created exception that courts employed to circumvent the Statute of Frauds' operation.³⁴³ This initial equitable intrusion into the Statute's domain was consistent with equity's traditional role of preventing the harsh operation of statutes.³⁴⁴ Courts of equity also commonly enforced oral contracts that were within the scope of the Statute in cases involving outright fraud on the theory that the supreme duty of equity was to prevent fraud.³⁴⁵ When cases arose that did not involve fraud, but in which denial of an oral

the imposition of constructive trusts, the "joint obligor" rule, the "main purpose" rule, equitable estoppel, and promissory estoppel. See Note, *supra* note 248, at 115-16.

340. L. FRIEDMAN, *supra* note 204, at 246-47.

341. See, e.g., Summers, *supra* note 208, at 442. Summers notes that [h]ow many unjust suits have been prevented as a result of the statute cannot be estimated, but the reports are filled with cases where just claims have been defeated by its operation. This has resulted in a distorting of the statute, in order to prevent injustice, into the most inconceivable meanings, so that cases might be ruled to fall without its provisions.

Id.

342. See *supra* notes 203-317 and accompanying text.

343. See Costigan, *supra* note 323, at 344. According to Professor Costigan, [t]he statute's framers were thoroughly familiar with the part-performance problem, and the decisions which shortly after the passage of the Statute of Frauds settled the law that part-performance would make the oral contract for the sale of land enforceable in chancery, notwithstanding the statute, are conclusive evidence that its framers never intended the statute to prevent the giving of equitable relief in the part-performance cases.

344. *Id.* See Summers, *supra* note 208, at 447.

345. See, e.g., *Wakeman v. Dodd*, 27 N.J. Eq. 564 (1876). In such cases courts justified the circumvention of the Statute by arguing: "relief is afforded in equity because of the fraud, and not by virtue of the contract." *Id.* at 565 (dictum).

promise's enforceability because of its failure to comply with the Statute's writing requirement would impose severe hardships on the promisee, equity courts began to employ the doctrine of equitable estoppel³⁴⁶ to defeat the Statute. Courts justified equitable estoppel's use in Statute of Frauds cases by noting that "the invocation of the statute would allow the perpetration of a moral fraud,"³⁴⁷ and that the prevention of such an inequitable use of the Statute would "be an aid in the ultimate function of the statute in preventing fraud."³⁴⁸

Equitable estoppel's utility as a device for protecting a promisee's reliance interests stemming from promises within the scope of the Statute suffered because of the doctrine's application only to misrepresentations of fact.³⁴⁹ This limitation confined early applications of estoppel in the Statute of Frauds context to situations in which the promisor either had made misrepresentations concerning facts that would obviate the need for a writing,³⁵⁰ or had represented that the promisor had executed a writing sufficient to satisfy the Statute.³⁵¹ The artificiality of the fact/promise distinction³⁵² and the manifest injustice that often flowed from its operation³⁵³ eventually would cause courts to apply the emerging doctrine of promissory estoppel to Statute of Frauds cases.

At the turn of the century Statute of Frauds cases began to appear in which courts that purported to apply equitable estoppel³⁵⁴ deviated significantly from that doctrine's traditional elements.³⁵⁵ Perhaps the most important of these cases was the California Supreme Court's decision in *Seymour v. Oelrichs*.³⁵⁶

346. See *supra* notes 212-14 and accompanying text for a discussion of equitable estoppel.

347. See Summers, *supra* note 208, at 446.

348. *Id.* at 447. Summers and others have argued that the part-performance doctrine, in reality, was nothing more than a limited variant of equitable estoppel. *Id.* The two doctrines clearly overlap and numerous examples exist of judicial confusion concerning their proper application. See Note, *supra* note 216, at 282-83 & nn.11-15.

349. See *supra* notes 217-18 and accompanying text.

350. See Note, *supra* note 267, at 595.

351. See Edwards, *The Statute of Frauds of the Uniform Commercial Code and the Doctrine of Estoppel*, 62 MARQ. L. REV. 205, 215 (1978).

352. See *supra* note 219 and accompanying text.

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

354. See, e.g., *Kingston v. Walters*, 16 N.M. 59, 113 P. 594 (1911) (seller orally promised to extend future payment date of debt); *Perkiomen Brick Co. v. Dyer*, 187 Pa. 470, 41 A. 326 (1898) (purchaser orally agreed to subscribe to stock).

355. See *supra* note 212 for the traditional elements of equitable estoppel.

356. 156 Cal. 782, 106 P. 88 (1909). The plaintiff in *Seymour*, a police captain, gave up secure employment, including pension rights and the right to a "good cause" discharge, in

Although purporting to base its decision on equitable estoppel,³⁵⁷ the *Seymour* court actually estopped the surviving defendants from asserting their Statute of Frauds defense because of their ancillary promise to reduce the parties' agreement to writing.³⁵⁸ Promissory, rather than equitable estoppel, more appropriately explains the *Seymour* result, based as it is upon plaintiff's reliance on the defendant's promise and not upon any misrepresentation of fact.³⁵⁹ Nonetheless, the court in *Seymour* understandably attributed its holding to equitable estoppel since the legal community neither formally nor informally had recognized the doctrine of promissory estoppel.³⁶⁰

Courts subsequently broadened the *Seymour* "ancillary promise" exception to equitable estoppel's misrepresentation of fact requirement to include, among other things, promises by the defendant not to use the Statute as a defense.³⁶¹ Courts justified the circumvention of the Statute of Frauds in "ancillary" promise situations on two theoretical grounds. First, the courts wanted to prevent "frauds" upon promisees who relied upon oral promises.³⁶² Second, the courts defeated the Statute only indirectly since they used estoppel to enforce the ancillary promise, which was not within the Statute's scope, rather than the underlying promise, which was within the Statute's ambit.³⁶³ However circuitous their logic, the courts ultimately were able to enforce promises that the Statute otherwise rendered unenforceable. Appropriate attribution of the *Seymour* result to promissory estoppel principles would not occur until the holding of the case gained official sanction of the *Restatement of Contracts* in 1932.³⁶⁴

reliance upon a ten-year oral contract to serve as a property manager for the defendants. He alleged that one of the defendants, who later died while traveling in Europe, orally had promised to reduce the agreement to writing upon his return. *Id.* at 792, 106 P. at 93. Since the parties could not perform the contract within one year of its making, it fell within the ambit of the California Statute of Frauds. *Id.* at 786, 106 P. at 90.

357. *See id.* at 800, 106 P. at 96.

358. *Id.* at 799-800, 106 P. at 96.

359. *See Summers, supra* note 208, at 454.

360. *See supra* note 222 and accompanying text.

361. *See, e.g., Zellner v. Wassman*, 184 Cal. 80, 86-87, 193 P. 84, 87 (1920).

362. *See, e.g., Seymour v. Oelrichs*, 156 Cal. at 797-800, 106 P. at 94-96.

363. *See Note, supra* note 248, at 117; *Note, supra* note 258, at 74.

364. *See* RESTATEMENT OF CONTRACTS § 178 comment f (1932). Comment f provides: Though there has been no satisfaction of the Statute, an estoppel may preclude objection on that ground in the same way that objection to the non-existence of other facts essential for the establishment of a right or a defense may be precluded. A misrepresentation that there has been such satisfaction if substantial action is taken in reliance on the representation, precludes proof by the party who made the representation that

Nothing in the language of section 90 of the original *Restatement*³⁶⁵—the general section sanctioning promissory estoppel—suggested that promissory estoppel should enjoy any broader application in Statute of Frauds cases. Some modern courts accordingly continue to limit the use of promissory estoppel in Statute of Frauds cases to factual situations concerning ancillary promises.³⁶⁶ Commentators have suggested that the reluctance of such courts to extend the application of promissory estoppel beyond cases involving ancillary promises may be attributable to fears that such an extension could result in the complete abrogation of the Statute.³⁶⁷ Exactly how much protection the Statute affords promisors even in jurisdictions recognizing only the ancillary promise exception remains in doubt. A perjurer attempting to prove the existence of an oral agreement otherwise within the ambit of the Statute seemingly could assert the existence of a fictitious promise to reduce the alleged agreement to written form and thus vitiate any residual protection the Statute affords.

Nevertheless, limiting promissory estoppel to the ancillary promise context potentially works great injustice upon the beleaguered promisee. For one thing, such a limitation forgets that a promisee who relies on an underlying promise to perform experiences as great an injury if a court refuses to enforce the promise as a promisee who relies on an ancillary promise to put the agreement in writing.³⁶⁸ More importantly, in most ancillary promise cases the promisee more likely is relying on performance of the underlying agreement than upon the ancillary promise. Not until the Califor-

it was false; and a promise to make a memorandum, if similarly relied on, may give rise to an effective promissory estoppel if the Statute would otherwise operate to defraud.

Id. (emphasis added).

365. See *supra* note 238 and the accompanying text.

366. See, e.g., *21 Turtle Creek Square, Ltd. v. New York State Teachers' Retirement Sys.*, 432 F.2d 64, 65-66 (5th Cir. 1970), *cert. denied*, 401 U.S. 955 (1971) (applying Texas law); *Tiffany, Inc. v. W.M.K. Transit Mix, Inc.*, 16 Ariz. App. 415, 420-21, 493 P.2d 1220, 1225-26 (1972).

367. See Note, *supra* note 255, at 1229.

368. Commentators have argued that the ancillary promise rubric often has served to mask the true policy basis for circumventing the Statute's operation in such cases:

It is appropriate for modern courts to cast aside the raiments of conceptualism which cloak the true policies underlying the reasoning behind the many decisions enforcing contracts that violate the Statute of Frauds. There is certainly no need to resort to legal rubrics or meticulous legal formulas when better explanations are available. The policy behind enforcing an oral agreement which violated the Statute of Frauds [is] a policy of avoiding unconscionable injury . . .

McIntosh v. Murphy, 52 Hawaii 29, 35, 469 P.2d 177, 180 (1970).

nia Supreme Court's decision in *Monarco v. Lo Greco*,³⁶⁹ however, did a court explicitly recognize this.³⁷⁰ Although the *Monarco* court adhered to the precedent set in *Seymour* by purporting to base its decision on equitable estoppel,³⁷¹ the doctrine they actually applied is promissory estoppel.³⁷² Several post-*Monarco* decisions recognizing the general proposition that reliance on oral promises can operate to circumvent the Statute regardless of the existence of an ancillary promise correctly have characterized the animating principle underlying their decisions as promissory estoppel.³⁷³ Courts also have used promissory estoppel to circumvent the Uniform Commercial Code's Statute of Frauds requirement,³⁷⁴ although the extant cases on point manifest considerable judicial confusion on the subject.³⁷⁵

Although the judicial tendency to circumvent the Statute of Frauds via promissory estoppel is apparently well-established and, indeed, seems to be gaining momentum, this Article would deceive if it failed to acknowledge the considerable resistance that the doctrine has encountered in the Statute of Frauds context. One obsta-

369. 35 Cal. 2d 621, 220 P.2d 737 (1950). In *Monarco* the court enforced an oral promise two parents made to their son that they would leave him the hulk of their property at their death. In reliance on this promise the son remained at home and worked on the family farm for twenty years. Although defendants made no ancillary promise to reduce the agreement to written form or to refrain from raising the Statute as a defense, the court concluded that estoppel appropriately could exist when a promisor induces the promisee's reliance by promising to perform his part of their oral agreement. The court observed that "[i]n reality it is not the representation that the contract will be put in writing or that the statute will not be invoked, but the promise that the contract will be performed that a party relies upon when he changes his position because of it." *Id.* at 626, 220 P.2d at 741. The court analyzed prior decisions and concluded that whenever unconscionable injury or unjust enrichment would result from enforcement of the Statute, "the doctrine of estoppel . . . applied whether or not plaintiff relied upon representations going to the requirements of the statute itself." *Id.* at 625, 220 P.2d at 741. The *Monarco* court characterized earlier cases refusing to apply estoppel as doing so either because a restitutionary remedy would compensate adequately the promisee or because no unconscionable injury would result from nonenforcement of the agreement. *Id.* at 623-24, 220 P.2d at 740.

370. A few cases decided in the interim between *Seymour* and *Monarco*, however, evidenced some considerable judicial stretching of the ancillary promise requirement in the quest for fair results. See Note, *supra* note 216, at 293-94.

371. 35 Cal. 2d at 625, 220 P.2d at 740.

372. See *supra* notes 359-60 and accompanying text. But see Note, *supra* note 255, at 1221-22 (describing *Monarco* as allowing proof of detrimental reliance sufficient to establish promissory estoppel to serve as the basis for an equitable estoppel). The Note's contrary position results from the author's view that promissory estoppel traditionally has functioned only as a consideration substitute. See *id.* at 1222-23.

373. See, e.g., *McIntosh v. Murphy*, 52 Hawaii 29, 36-37, 469 P.2d 177, 181 (1970); *Alpark Distrib., Inc. v. Poole*, 95 Nev. 605, 607-08, 600 P.2d 229, 230-31 (1979).

374. U.C.C. § 2-201 (1977).

375. See Metzger & Phillips, *supra* note 338, at 91-96.

cle to the recognition of promissory estoppel as a legitimate device to circumvent the Statute has been judicial fears that observers would charge courts with usurping legislative power.³⁷⁶ This concern, however, ignores several important facts. First, equity always has had the power to mitigate the harsh operation of statutes.³⁷⁷ Second, the Statute has from its inception been the object of equitable intervention;³⁷⁸ indeed, one prominent historian of the Statute has argued that its authors never intended the Statute to apply to the courts of equity.³⁷⁹ Last, the usurpation argument fails to consider the judiciary's primary role in shaping the Statute.³⁸⁰

The usurpation argument against recognition of promissory estoppel as a legitimate means around the Statute recently has resurfaced with particular vigor in the discussion of section 2-201 of the Uniform Commercial Code. In the discussion, opponents of the widespread recognition of promissory estoppel in Statute of Frauds cases have coupled the usurpation argument with arguments premised on the familiar, though often discounted,³⁸¹ maxim of statutory construction: *Expressio unius est exclusio alterius*.³⁸² The major thrust of this combined attack is that, since the drafters of the Code specifically enumerated several methods³⁸³ of avoiding

376. See, e.g., *Tanenbaum v. Biscayne Osteopathic Hosp., Inc.*, 190 So. 2d 777, 779 (Fla. 1966).

377. See *supra* note 344 and accompanying text.

378. See *supra* notes 343 & 345-48 and accompanying text.

379. The judges who framed the Statute of Frauds were so anxious to tie the hands of juries and so possessed by the idea that the statute would not apply *ex proprio vigore* to chancery cases that they neglected to be as explicit in the wording of the statute as they should have been.

Costigan, *supra* note 323, at 344-45.

380. See *supra* notes 339-40 and accompanying text.

381. See Radin, *Statutory Interpretation*, 43 HARV. L. REV. 863, 873-74 (1930), for the following observation:

The rule that the expression of one thing is the exclusion of another is in direct contradiction to the habits of speech of most persons. To say that all men are mortal does not mean that all women are not, or that all other animals are not. There is no such implication, either in usage or in logic, unless there is a very particular emphasis on the word *men*. It is neither customary nor convenient to indicate such emphasis in statutes, and without this indication, the first comment on the rule is that it is not true.

Id. (emphasis in original). See also *National Petroleum Refiners Ass'n v. FTC*, 482 F.2d 672, 676 (D.C. Cir. 1973).

382. Enumeration of specific items impliedly excludes all others. See *United States v. Robinson*, 359 F. Supp. 52, 58-59 (S.D. Fla. 1973) (interpreting wiretap authorization statute).

383. See U.C.C. § 2-201(2)-(3) (1977). These methods include the use of merchants' written confirmations, the manufacture of goods not suitable for sale to others in the regular course of business, an admission by the party against whom enforcement is sought that a contract for sale was made, and the prior acceptance of, coupled with either payment for or

the section's writing requirement,³⁸⁴ the drafters impliedly precluded avoidance by any other manner, including estoppel.³⁸⁵ This view is incorrect for various reasons.³⁸⁶ The most important of these is that section 1-103 of the Code expressly provides that "[u]nless displaced by the particular provisions of this Act, the principles of law and equity, including . . . the law relative to . . . estoppel . . . shall supplement its provisions."³⁸⁷ This section applies prospectively as well as retrospectively,³⁸⁸ and, therefore, can serve as a vehicle for incorporating into Code decisions legal developments that occur after the Code's enactment.³⁸⁹

Another, more important obstacle to the recognition of promissory estoppel as a legitimate means of circumventing the Statute is the fear of numerous courts that such recognition would result in the Statute's complete abrogation.³⁹⁰ This fear, however, is unfounded. While promissory estoppel would permit proof of some oral agreements that the Statute otherwise would preclude the Statute would remain applicable both to wholly executory promises within its scope and to oral promises that have induced reliance insufficient to trigger application of estoppel.³⁹¹ Furthermore, although allowing estoppel to intrude into the Statute of Frauds context certainly will diminish some protection that the Statute otherwise affords contracting parties, substantial controversy exists over whether the Statute truly offers *any* significant protection. Given the Statute's numerous judicially created exceptions,³⁹² a dishonest plaintiff conceivably could "frame" a contract

receipt of, the goods in issue.

384. See U.C.C. § 2-201(1) (1977).

385. See *C.R. Fedrick, Inc. v. Borg-Warner Corp.*, 552 F.2d 852, 858 (9th Cir. 1977) (Duniway, J., concurring); *Cox v. Cox*, 292 Ala. 106, 111, 289 So. 2d 609, 613 (1974); *C.G. Campbell & Son, Inc. v. Comdeq Corp.*, 586 S.W.2d 40, 41 (Ky. Ct. App. 1979); see also *Edwards*, *supra* note 351, at 218.

386. See *Metzger & Phillips*, *supra* note 338, at 97-99 (fully discussing why the Code's enumeration does not preclude other methods of avoidance).

387. U.C.C. § 1-103 (1977).

388. "What the section [1-103] invites is not limited to law which exists as of the date of particular enactments of the Code." Summers, *supra* note 304, at 197 n.9.

389. Specifically, courts, under § 1-103, appropriately may incorporate the express recognition of promissory estoppel as a means of circumventing the Statute in *Restatement (Second) of Contracts* § 139.

390. This fear has been a strong consideration in those cases in which courts refused to countenance the circumvention of the Statute under any circumstances. See, e.g., *Kahn v. Cecelia Co.*, 40 F. Supp. 878, 880 (S.D.N.Y. 1941); *Tanenbaum v. Biscayne Osteopathic Hosp., Inc.*, 173 So. 2d 492, 495 (Fla. Dist. Ct. App. 1965), *aff'd*, 190 So. 2d 777 (Fla. 1966).

391. See *Metzger & Phillips*, *supra* note 338, at 82.

392. See *supra* notes 339-40 and accompanying text.

that was beyond the Statute's scope.³⁹³ Moreover, that a writing satisfying the Statute exists does not indicate necessarily the true existence of a contract between the parties. The writing might be a forgery,³⁹⁴ or might be a draft that parties optimistically signed in advance, looking forward to an agreement that failed to materialize.³⁹⁵ Finally, memoranda satisfying the Statute are often incomplete;³⁹⁶ thus, a perjurious promisee conceivably could convince a judge or jury of the existence of contract terms to which the promisor never actually assented.³⁹⁷ All of these possibilities have led one commentator to observe that "there is a very serious doubt as to what protection the statute really does afford. Despite the lofty words about its glory that are so often pronounced, its protection, in most cases, is more illusory than real."³⁹⁸

Proponents of the expansion of promissory estoppel into the Statute of Frauds context, in sum, have convincing responses to all judicial fears concerning the abrogation of the Statute. Perhaps, the chief fear concerns the loss of the evidentiary function that the writing requirement serves.³⁹⁹ The reliance that the promissory estoppel doctrine requires, however, likewise can serve this evidentiary function.⁴⁰⁰ As one author notes "it is almost inconceivable that anyone should materially change his position, so as to satisfy all the elements of an estoppel, on the expectation of recouping himself on a 'framed' contract."⁴⁰¹ Another fear concerns the loss of the channeling and cautionary functions that the Statute's writing requirement serves.⁴⁰² The channeling function of the Statute encourages knowledgeable parties to reduce their agreements to

393. See Summers, *supra* note 208, at 460.

394. J. WHITE & R. SUMMERS, *supra* note 27, § 2-8, at 73.

395. *Id.*

396. *Id.* § 2-3, at 57 (sales contract context). The various incarnations of the Statute never have required that a memorandum sufficient to satisfy the Statute's writing requirement be a complete integration of the parties' agreement. J. CALAMARI & J. PERILLO, *supra* note 20, § 19-27 & n.70.

397. J. WHITE & R. SUMMERS, *supra* note 27, § 2-8, at 72-73 (sales contract context).

398. Summers, *supra* note 208, at 460. See also J. WHITE & R. SUMMERS, *supra* note 27, § 2-8, at 73 (The Uniform Commercial Code's writing requirement is "so far from any kind of guarantee against successful perjury that it is inappropriate even to call it a means to fraud prevention at all.").

399. See *supra* notes 319-24 and accompanying text.

400. See *infra* notes 423-24 and accompanying text for a discussion of the *Restatement (Second) of Contracts'* recognition of promissory estoppel's evidentiary function. See also Note, *supra* note 208, at 710-11.

401. Summers, *supra* note 208, at 459-60.

402. See Note, *supra* note 215, at 182; see also J. CALAMARI & J. PERILLO, *supra* note 20, § 19-1, at 673.

written form, and thus provides courts with a handy basis for distinguishing between enforceable and unenforceable agreements.⁴⁰³ The cautionary function of the writing requirement discourages parties from entering into ill-considered agreements by forcing them to reduce their agreement to writing and impressing upon them the seriousness of their actions.⁴⁰⁴ While the promissory estoppel reliance requirement cannot serve either of these functions,⁴⁰⁵ significant grounds exist for concluding that the writing requirement also is inadequate as a channeling and cautionary device.⁴⁰⁶

Before the Statute can serve the channeling and cautionary functions, the parties must be aware of its existence and of its applicability to their agreement—a prerequisite often absent in everyday transactions.⁴⁰⁷ An empirical study of actual business practice relating to the reduction of oral agreements to written form that the *Yale Law Journal* presented in 1957 indicated that the Statute plays a relatively insignificant role in inducing businessmen to put their agreements in written form.⁴⁰⁸ Firms generally put their agreements in writing because they consider such devotion to the written word sound business policy.⁴⁰⁹ Parties often put their agreements in writing simply to avoid problems of interpretation or to prevent each other from denying the existence of the agreement.⁴¹⁰ Even if the parties are aware of the Statute's writing requirement, they may refrain from creating a writing because one

403. See Note, *supra* note 215, at 170-71; see also Fuller, *supra* note 304, at 800-01 (A formality like the writing requirement can provide a simple, external test for enforceability.).

404. See Fuller, *supra* note 304, at 800; Note, *supra* note 215, at 170.

405. But see Note, *supra* note 208, at 710. Since promissory estoppel only imposes liability for promises on which the promisor should foresee injurious reliance it, therefore, arguably fulfills a cautionary function because "[l]iability thus is not imposed for breach of impulsive promises made under circumstances in which legal consequences are, or should be, unexpected." *Id.*

406. The drafters of the Uniform Commercial Code apparently were not concerned with the channeling and cautionary functions of the writing requirement, since the various alternative methods they provided for satisfying the Statute in Code cases all concern the evidentiary function of the Statute. See Edwards, *supra* note 351, at 218.

407. See *supra* note 333. For more recent evidence of the truth of this assertion, see *The Statute of Frauds and the Business Community: A Re-Appraisal in Light of Prevailing Practices*, 66 *YALE L.J.* 1038, 1057-58 (1957) [hereinafter cited as *Yale Study*], which indicates that the Sales Act's requirement of written evidence of sales contracts for \$500 or more had very little to do with whether businessmen reduced their agreements to written form.

408. See *id.*; see also Note, *supra* note 267, at 592.

409. *Yale Study*, *supra* note 407, at 1064.

410. See Note, *supra* note 267, at 593 n.16.

of the parties relies on the other's promise either to put their agreement in writing or to refrain from raising the Statute as a defense, or because both mistakenly believe that they have satisfied the writing requirement.⁴¹¹

The Yale Study also raised significant questions concerning the identity of the individuals or groups most likely to benefit from the Statute's writing requirement. According to the study, large manufacturers are more likely to demand written evidence of oral agreements than their smaller competitors⁴¹² because of the difference in the nature of large- and small-scale business operations⁴¹³ and the comparatively lesser bargaining power that smaller companies enjoy.⁴¹⁴ Another reason that those operating smaller concerns are less writing-conscious may be that the small firms are less knowledgeable than their larger competitors about the Statute's existence and its applicability. The Statute, therefore, may be most likely to affect adversely those parties least likely to know of its requirements, least able to secure compliance with the requirements of which they are aware, and least capable of absorbing any losses associated with their reliance if others successfully use the Statute to bar their claims.⁴¹⁵

The foregoing discussion suggests that the channeling and cautionary functions attributed to the Statute's writing requirement may be largely illusory since they are at odds with the realities of prevailing business practice; therefore, additional impetus exists for protecting a party's reliance on an oral promise within the ambit of the Statute. As Justice Stephen observed long ago, "[l]aws ought to be adjusted to the habits of society, and not to aim at remoulding them."⁴¹⁶ This observation seems particularly true if actual experience with a rule indicates that it has had little real effect in shaping human behavior despite over three centuries of existence. The Statute's failure to fulfill its channeling and cautionary functions, the possible elimination of the original justifications for the Statute by subsequent legal developments,⁴¹⁷ and courts' wide recognition of the Statute's potential for working in-

411. *Id.* at 597.

412. *See Yale Study, supra* note 407, at 1047.

413. *Id.* at 1051.

414. *Id.* at 1051-55.

415. Metzger & Phillips, *supra* note 338, at 103.

416. Stephen & Pollock, *supra* note 327, at 6.

417. *See supra* notes 329-32 and accompanying text.

justice⁴¹⁸ combine to present a compelling case for circumvention of the Statute via promissory estoppel in appropriate cases, if not its outright repeal.⁴¹⁹

Two recent developments should serve to accelerate the growing judicial trend⁴²⁰ toward defeating the Statute's writing requirement via promissory estoppel. The first, and least controversial, is the *Restatement (Second) of Contracts'* express recognition in section 139⁴²¹ that courts appropriately may utilize promissory estoppel to circumvent the Statute. The second development is the emergence of the notion of promissory estoppel as a theory of recovery independent of contract.

In addition to the obvious "respectability" conferred by the quasi-official sanction of the *Restatement (Second)*, section 139 includes evidentiary and remedial provisions⁴²² that should allay judicial fears concerning this use of promissory estoppel and should enhance the possibility that the future development of the doctrine will proceed in a more systematized, rational fashion. Subsection 139(2)(c)⁴²³ of the *Restatement (Second)* expressly considers the evidentiary value of reliance by directing courts to consider the extent to which the promisee's reliance or other evidence introduced at trial corroborates the existence of the alleged oral promise. This

418. See *supra* notes 333-36 and accompanying text.

419. Parliament repealed the bulk of the original Statute in 1954, retaining the writing requirement only for contracts for the sale of land and promises to answer for the debt, default, or miscarriage of another. Law Reform (Enforcement of Contracts) Act, 1954, 2 & 3 Eliz. 2, ch. 34. See also *supra* note 192 and accompanying text.

420. See *supra* notes 354-64 & 369-75 and accompanying text.

421. *Restatement (Second) of Contracts* § 139, reads as follows:

Enforcement by Virtue of Action in Reliance

(1) A promise which the promisor should reasonably expect to induce action or forbearance on the part of the promisee or a third person and which does induce the action or forbearance is enforceable notwithstanding the Statute of Frauds if injustice can be avoided only by enforcement of the promise. The remedy granted for breach is to be limited as justice requires.

(2) In determining whether injustice can be avoided only by enforcement of the promise, the following circumstances are significant:

(a) the availability and adequacy of other remedies, particularly cancellation and restitution;

(b) the definite and substantial character of the action or forbearance in relation to the remedy sought;

(c) the extent to which the action or forbearance corroborates evidence of the making and terms of the promise, or the making and terms are otherwise established by clear and convincing evidence;

(d) the reasonableness of the action or forbearance;

(e) the extent to which the action or forbearance was foreseeable by the promisor.

422. See *infra* notes 425-30 and accompanying text.

423. See *supra* note 421.

provision should ensure wider acceptance of promissory estoppel in the Statute of Frauds context by satisfying the evidentiary purposes that the writing requirement ordinarily served and by obviating the possibility that courts will refuse to recognize an estoppel because of fears that the alleged oral agreement may be fictitious.⁴²⁴ Section 139 also expressly treats the issue of the proper measure of damages in promissory estoppel cases—a subject which has generated considerable confusion and inconsistency since the initial formal recognition of the doctrine.⁴²⁵ By stating that the “remedy granted for breach is to be limited as justice requires,”⁴²⁶ section 139(1) expressly authorizes the use of reliance-based damages in promissory estoppel cases.⁴²⁷ This not only provides for the judicial discretion necessary to deal effectively with promissory estoppel’s remedial dilemmas,⁴²⁸ but also should enable courts to retain whatever measure of the channeling function the writing requirement provides⁴²⁹ by limiting recovery under oral promises to reliance damages. The judicial position espoused in section 139(1) should encourage knowledgeable parties to continue reducing their agreements to written form and should represent a proper balance between the abrogation of the Statute that complete enforcement of oral promises would cause, and the injustice that would result from rote denial of recovery for failure to comply with the strictures of the Statute.⁴³⁰

The second and more controversial support for using promissory estoppel to circumvent the Statute comes from the emergence of authorities suggesting the doctrine’s possible status as an independent theory of recovery.⁴³¹ Some recent cases suggest that the Statute does not bar claims premised upon promissory estoppel since such claims are not contractually based and, therefore, are beyond the Statute’s scope.⁴³² Such a suggestion represents a sig-

424. See *supra* notes 400-01 and accompanying text.

425. See *supra* notes 265-81 and accompanying text.

426. RESTATEMENT (SECOND) OF CONTRACTS § 139(1) (1981).

427. But see RESTATEMENT (SECOND) OF CONTRACTS § 139 comment d (1981). This comment provides that “when specific enforcement is available under the rule stated in § 129 [providing specific enforcement of land sales contracts when courts can avoid injustice only by specific performance], an ordinary action for damages is commonly less satisfactory” *Id.*

428. See *supra* notes 265-81 and accompanying text.

429. See *supra* notes 402-15 and accompanying text.

430. See Note, *supra* note 255, at 1242.

431. See *supra* notes 295-308 and accompanying text.

432. See, e.g., *R.S. Bennett & Co. v. Economy Mechanical Indus.*, 606 F.2d 182, 188 (7th Cir. 1979) (permitting plaintiff to assert promissory estoppel claim notwithstanding

nificant theoretical departure from the courts' traditional rationale in using estoppel to circumvent the Statute's operation—the prevention of injustice that otherwise would result from the Statute's operation.⁴³³ While little practical difference arguably exists between allowing promissory estoppel to “substitute for” the Statute and treating the doctrine as an independent theory of recovery, the rationale of the latter theory is pregnant with significant implications concerning the fate of other traditional “contract” defenses such as the parol evidence rule. As Grant Gilmore observed,

[S]ome of the recent cases are beginning to suggest that liability under § 90 or the doctrine of promissory estoppel or however it is described is somehow different from liability in contract. Thus, it may be, defenses based on the *statute of frauds* or the contract statute of limitation or the *parol evidence rule*—all these being looked on as contract-based defenses—are no longer available if the underlying theory of liability—§ 90 or an analogue—is not contract theory at all.⁴³⁴

Treating promissory estoppel as an independent theory of recovery, furthermore, could provide judges with an obvious, albeit technical, device for circumventing the Statute in situations in which mechanical application of the Statute otherwise would produce harsh results. Finally, by giving promissory estoppel independent theory status, judges would avoid charges of judicial usurpation of legislative power and abrogation of the Statute's writing requirement.⁴³⁵

V. PROMISSORY ESTOPPEL AND THE PAROL EVIDENCE RULE

A. *The Cases*

While, as the foregoing discussion plainly indicates, promissory estoppel has gained widespread legitimacy as a device for circumventing the Statute of Frauds, the same is not true in the parol evidence context. Upon survey of extant judicial authority on the subject, several things become abundantly clear. First, very few cases exist in which the issue of promissory estoppel's applicability to the parol evidence rule ever has arisen. Second, the fundamental attribute that all of the cases in which the issue has arisen share is a paucity of significant judicial analysis on the estoppel issue. Last,

Statute's bar to contract recovery); *N. Litterio & Co. v. Glassman Constr. Co.*, 319 F.2d 736, 738-40 & n.9 (D.C. Cir. 1963); *Janke Constr. Co. v. Vulcan Materials Co.*, 386 F. Supp. 687, 696-97 (W.D. Wis. 1974), *aff'd*, 527 F.2d 772 (7th Cir. 1976).

433. See *supra* notes 347-48 and accompanying text.

434. G. GILMORE, *supra* note 203, at 66 (emphasis added).

435. See *supra* notes 376-89 and accompanying text.

and perhaps most significantly,⁴³⁶ very few cases exist in which a court has refused to consider convincing evidence of an extrinsic promise in situations in which such refusal, in the court's judgment, would work a serious injustice on the party relying on the extrinsic promise. In one of the earliest cases to consider the applicability of promissory estoppel in the parol evidence context, however, the court did refuse to consider such convincing evidence.

In *Berverdor, Inc. v. Salyer Farms*⁴³⁷ the California Court of Appeals held the trial court in error for admitting parol evidence of an omitted lease condition since the written lease was "on its face, full and complete and no fraud, mistake, imperfection, illegality or ambiguity was proved."⁴³⁸ The appellate court rejected any application of equitable estoppel because the defendants had failed to prove the misrepresentation of fact that the doctrine required. The court then rejected application of promissory estoppel, although it acknowledged that promissory estoppel properly could be founded on a promise of future performance. The court observed that "we have been cited no case and we have been unable to find any case which in any way aids in the application of that doctrine to the case at bar."⁴³⁹ Given the court's form-oriented approach to the parol evidence rule, its views on the unenforceability of the extrinsic agreement, and the realization that the court decided *Berverdor* nearly three months before the California Supreme Court's landmark *Monarco* decision on promissory estoppel's applicability to the Statute of Frauds,⁴⁴⁰ that the court gave

436. See *infra* notes 580-83 and accompanying text.

437. 97 Cal. App. 2d 459, 218 P.2d 138 (1950).

Plaintiff-lessor sued defendant-lessees for the last year of rent due under a three-year farm lease. Defendants acknowledged their failure to pay, but argued that the written lease was subject to a parol condition that the lessor would excuse rent payments if it became unprofitable for the lessee to pay. *Id.* at 460, 218 P.2d at 139-40. The trial court allowed defendants to introduce certain letters evidencing the alleged condition in support of this contention. Plaintiff's managing agent had signed the letters and had sent them to defendants prior to the execution of the lease. *Id.* at 461, 218 P.2d at 140. The trial court also allowed defendants to introduce evidence of certain conversations with plaintiff's agent who died shortly after execution of the lease; the most important of these conversations focused on defendants' concerns that the lease did not contain the agreed upon rent adjustment provision and on the agent's assurances that this omission was unimportant. *Id.* The trial court further allowed defendants to introduce evidence that in the year when they did not pay the rent insufficient water was available to support a profitable crop. *Id.* The trial court rendered judgment for plaintiff in a reduced amount, and both parties appealed.

438. *Id.* at 464, 218 P.2d at 142.

439. *Id.*

440. See *supra* notes 369-72 and accompanying text. The California Supreme Court decided *Monarco* on August 1, 1950. The California appeals court decided *Berverdor* on

defendants' estoppel arguments relatively short shrift is not surprising.

Nonetheless, the treatment that the defendants received in *Berverdor* stands in rather stark contrast to the treatment afforded parties relying on alleged oral agreements in much earlier California Statute of Frauds cases. In *Seymour v. Oelrichs*,⁴⁴¹ for example, the court allowed a party who had relied upon an alleged oral promise to reduce a wholly oral ten-year employment contract to writing to prove the existence of that contract. In *Berverdor*, on the other hand, the court denied defendants the opportunity to prove the existence of a single term of an otherwise written contract, even though defendants had credible written evidence supporting the existence of that term, were arguably no less reasonable in relying on the agent's promises and assurances than their counterpart in *Seymour*, and had suffered a significant injustice as a result of their reliance.

While the scant attention the *Berverdor* court paid to defendants' promissory estoppel arguments, perhaps, is excusable since the promissory estoppel doctrine had not departed yet from its traditional roles in any dramatic way when the court decided the case, later cases in which litigants sought to apply promissory estoppel in parol evidence cases unfortunately fail to provide any greater judicial analysis of the issue. For example, in *Mack v. Earle M. Jorgensen Co.*⁴⁴² the court found that the parties' written contract represented a complete integration of their agreement and that the parol evidence rule barred any proof of alleged oral promises. The court refused to allow proof of the promises under the umbrella of promissory estoppel, but did not reject outright the possibility that the reliance doctrine could, in some circumstances, operate to circumvent the parol evidence rule. Instead, the court concluded that "the evidence presented precludes the application of the doctrine of promissory estoppel,"⁴⁴³ observing that "We are not aware of any decision applying the doctrine of promissory estoppel where, as in the instant case, the alleged oral agree-

May 10, 1950.

441. See discussion *supra* notes 356-60 and accompanying text.

442. 467 F.2d 1177 (7th Cir. 1972). In *Mack* a terminated manufacturer's representative sought to introduce proof of oral assurances that his employer would not discharge him as long as he performed adequately. He alleged that various officers of the defendant company made these assurances after the company acquired his original employer and before he signed a written contract of employment containing a thirty-day termination clause. *Id.* at 1178-79.

443. *Id.* at 1179.

ment or promise was followed by a written contract, the terms of which are in direct conflict with the alleged oral agreement or promise."⁴⁴⁴ Thus, the perceived conflict between the alleged oral assurances and the written contract served, in the eyes of the court, to prevent any further consideration of the estoppel issue because "[i]n such a situation, we seriously doubt whether the promisee could successfully argue that his reliance on the promise was justifiable."⁴⁴⁵ Of course the alleged oral promises conceivably were not inconsistent with the written terms, but rather represented a condition upon the terms. Viewed in this light, the plaintiff's reliance seems more reasonable. The court's finding of inconsistency and its ultimate resolution of the case may reflect the court's doubts about whether the parties ever made the alleged promises.⁴⁴⁶ Furthermore, the facts as presented by the court evidence no particular reliance by the plaintiff—no foregone alternative employment opportunities or other such substantial indicia of reliance that would produce injustice if the court refused to enforce the defendant's promise. Perhaps, given facts more conducive to a promissory estoppel claim the court might have considered its application to a parol evidence case.⁴⁴⁷

While the Seventh Circuit in *Mack* hinted at a willingness to apply promissory estoppel in certain parol evidence cases, the Supreme Court of Alaska unceremoniously discarded the applicability of the doctrine in *Johnson v. Curran*.⁴⁴⁸ The court in *Johnson*

444. *Id.*

445. *Id.*

446. The court, when discussing plaintiff's allegation that defendant's vice president made the alleged promise on the day plaintiff signed the written contract, stated that plaintiff "claims that Dellinger [the vice president] told him that he wouldn't be terminated as long as he did a good job." *Id.* (emphasis added). The court followed this statement with a discussion noting that plaintiff signed the written contract without objection, even though he had insisted on scratching out another provision which related to credits for orders booked before termination and which defendant originally included in the same paragraph as the termination clause. *Id.*

447. The validity of this suspicion is borne out by the court's subsequent decision in *Ehret Co. v. Eaton, Yale & Towne, Inc.*, 523 F.2d 280 (7th Cir. 1975), *cert. denied*, 425 U.S. 943 (1976). See *infra* notes 484-96 and accompanying text.

448. 633 P.2d 994 (Alaska 1981). In *Johnson* plaintiffs brought suit against a nightclub owner who fired them despite the fact that their eight-week written employment contract (an American Federation of Musicians standard form contract) provided for two weeks more employment. Defendant, in an uncontroverted affidavit in opposition to the plaintiff's motion for summary judgment, asserted that a band member misrepresenting himself as the band's leader orally had assured her, before she signed the contract, that the agreement was subject to termination on two week's notice if the band did not draw well. *Id.* at 995. Defendant argued that the writing at best was a partial integration, that it was ambiguous, and that, in any event, evidence of the parol agreement should be admissible because of fraud in

affirmed the trial court's refusal to admit parol evidence of a termination clause not contained in the parties' standard form contract. While the court recognized that evidence of fraud is admissible even when a writing is a complete integration, the court held that the fraud exception to the rule did not apply because defendant failed to show that plaintiff's misrepresentation induced the defendant to enter the contract.⁴⁴⁹ More important for our purposes, the court summarily dismissed defendant's promissory estoppel argument, noting that "the rule has no applicability in the factual context presented in the instant case; promissory estoppel is a principle applicable to situations where a promise unsupported by consideration is sought to be enforced and thus has no direct relevance to the case at bar."⁴⁵⁰ The court apparently did not have any doubts about the credibility of the defendant's assertions since it "assumed" that the defendant's assertions were true.⁴⁵¹ The defendant's reliance losses, however, were not of the sort likely to present a particularly compelling case for relief. This consideration, the court's narrow conception of promissory estoppel, and its rather mechanical approach to the parol evidence rule make the scant attention afforded the promissory estoppel argument more understandable, if not more defensible.

In at least one more case in which the promissory estoppel issue surfaced, the court not only sidestepped the estoppel question, but also refused to grant relief on any other basis. In *Clark Oil & Refining Corp. v. Leistikow*⁴⁵² the court found that defendant's promissory estoppel argument presented no Statute of Frauds problem, but observed that "there is . . . difficulty with this defense as it relates to the parol evidence rule" since the leases in question were clear as to the duration of the agreement and contained an integration clause requiring modifications to be in writing and signed by the lessor's vice president.⁴⁵³ Rather than address the issue of whether a party can use promissory estoppel to circumvent the parol evidence rule, the court concluded that the

the inducement and promissory estoppel. *Id.* at 995-96.

449. *Id.* at 997.

450. *Id.* at 996.

451. *Id.* at 997-98.

452. 69 Wis. 2d 226, 230 N.W.2d 736 (1975). In *Clark Oil* two service station lessees raised promissory estoppel as a defense to an eviction action that followed the expiration of the parties' written lease agreement. The lessees argued that the lessor's agents orally promised that they could remain dealers as long as their performance was satisfactory, and that they relied upon plaintiff's promise by leaving long-standing former employment.

453. *Id.* at 237-38, 230 N.W.2d at 743.

defendant's allegations were insufficient to raise a promissory estoppel claim. The defendants, according to the court, had not shown that the court could avoid injustice only by enforcing the plaintiff's promises. The court also expressed doubts about whether foregoing prior employment constituted reliance of a "definite and substantial character."⁴⁵⁴ The quoted language plainly addresses the requirements of section 90 as manifested in the *Restatement*.⁴⁵⁵ While the *Restatement (Second)* since has dropped the definite and substantial reliance requirement,⁴⁵⁶ and other courts have found foregoing prior employment to be sufficient reliance for a promissory estoppel claim,⁴⁵⁷ the heart of the *Clark Oil* court's refusal to confront the promissory estoppel issue or to employ any other legal device to protect the defendants' reliance probably lies with its conclusion that defendants failed to demonstrate sufficient injustice to justify enforcing plaintiff's oral promise. Defendants, after all, could have entered into a new lease at a higher rate if they were willing to pay for it. Furthermore, defendants did not stand to lose a substantial investment in the franchise. Given the absence of a compelling claim for relief, the court, not surprisingly, eschewed defendants' invitation to apply promissory estoppel to the parol evidence rule.

While the courts in the foregoing cases denied application of promissory estoppel in the parol evidence context and refused to admit proffered extrinsic evidence on any other grounds, a number of cases exist in which courts declined formal use of promissory estoppel to circumvent the parol evidence rule, but found independent bases for enforcing parol promises that engendered significant reliance. For example, in *Wojciechowski v. Amoco Oil Co.*⁴⁵⁸ plaintiff, a gas station franchisee, sought to enjoin the termination of a written "trial franchise" agreement drafted in conformity with the Petroleum Marketing Practices Act [PMPA].⁴⁵⁹ Plaintiff argued that defendant's agents induced him to enter the trial agreement by making oral misrepresentations that the PMPA required that defendant treat him as a trial franchisee⁴⁶⁰ and by orally promising that defendant would renew plaintiff's franchise if plaintiff "per-

454. *Id.* at 239, 230 N.W.2d at 744.

455. *See supra* note 238 and accompanying text.

456. *See supra* notes 309-12.

457. *See, e.g.,* *Hunter v. Hayes*, 533 P.2d 952 (Colo. Ct. App. 1975).

458. 483 F. Supp. 109 (E.D. Wis. 1980).

459. 15 U.S.C. §§ 2801-2806 (Supp. IV 1980).

460. 483 F. Supp. at 111, 114.

formed adequately.”⁴⁶¹ Defendant’s witnesses disputed plaintiff’s version of these events, but the court concluded that plaintiff “was a more credible witness.”⁴⁶² The court, nonetheless, rejected plaintiff’s argument that defendant’s promise to continue his franchise if he performed adequately gave rise to a claim based upon promissory estoppel. Promissory estoppel, according to the court, “provides no help to plaintiff’s claim [because the oral promises] were made prior to the written contract.”⁴⁶³ Therefore, the court concluded, “the parol evidence rule nullifies the salutary [sic] effects of the doctrine of promissory estoppel.”⁴⁶⁴

Despite its holding on promissory estoppel, the *Wojciechowski* court did not want to deprive plaintiff of his significant investment in the franchise. Consequently, it held that defendant’s misrepresentations constituted a fraud that prevented the creation of a trial franchise under the PMPA.⁴⁶⁵ Thus, defendant could not terminate plaintiff’s franchise until defendant conformed with the more stringent PMPA requirements for terminating regular franchises.⁴⁶⁶ Regarding the misrepresentations that PMPA required defendant to treat plaintiff as a trial franchisee, the court acknowledged the general rule that fraud cannot arise from misrepresentations of law, but held that defendant’s misstatements concerning the PMPA’s requirements fell within a recognized exception to the general rule predicated upon defendant’s intentional misrepresentation, the relationship of trust between the parties, and defendant’s superior knowledge and skill.⁴⁶⁷ To defendant’s objection that its promise relating to renewal of the franchise could not constitute fraud because it represented acts Amoco was to do in the future, the court responded that “if [defendant] had a pre-

461. *Id.* at 114.

462. *Id.*

463. *Id.* at 115.

464. *Id.* *Wojciechowski* relied heavily on *McConnell v. L.C.L. Transit Co.*, 42 Wis. 2d 429, 167 N.W.2d 226 (1969), which concerned an alleged oral promise to a man employed as a general manager that the company would not discharge him as long as the business earned an annual profit. His written employment contract, however, clearly gave the employer the right to terminate him upon payment of an agreed sum as liquidated damages. The *McConnell* court did not reach the promissory estoppel issue because plaintiff asserted that defendant made many of the alleged promises *after* the parties signed the agreement. Thus, the parol evidence rule did not bar evidence of their existence and of their role in inducing the plaintiff to forego other employment opportunities. *Id.* at 437, 167 N.W.2d at 229-30. The *Wojciechowski* court, therefore, correctly concluded that *McConnell* was “of little help to plaintiff.” 483 F. Supp. at 115.

465. 483 F. Supp. at 114-15 (citing 15 U.S.C. § 2803(b)(1) (Supp. V 1981)).

466. *Id.* at 115 (citing 15 U.S.C. § 2802 (Supp. V 1981)).

467. *Id.* at 114-15.

sent intention not to continue the franchise even if the performance was adequate, then this would constitute actionable fraud."⁴⁶⁸ The court then concluded that "there is sufficient proof at this juncture in the litigation to show that defendant misrepresented its intention."⁴⁶⁹

*Walker v. KFC Corp.*⁴⁷⁰ is similar in thrust, if not in methodology, to *Wojciechowski*. The plaintiff in *Walker* had purchased a restaurant franchise from defendant franchisor. Plaintiff claimed that defendant made numerous fraudulent parol representations concerning the viability of the franchise and its intentions concerning the franchise and the plaintiff. Plaintiff brought suit on numerous theories, including breach of contract, promissory estoppel, and fraud. When the jury found for plaintiff on the promissory estoppel and fraud claims, defendant moved for a judgment n.o.v., arguing (1) that the existence of a contract between the parties precluded any recovery under promissory estoppel and (2) that, in any event, promissory estoppel was inapplicable to promises preceding a written contract because of the parol evidence rule.⁴⁷¹

On the first point, the court acknowledged the general rule that existence of a valid contract prevents any recovery based on estoppel,⁴⁷² but noted that "defendants provide no authority for the proposition that the mere existence of a contract between two parties precludes a recovery based on promises that were not part of the contract."⁴⁷³ The court then proceeded to hold the oral promises as separate from the contract and to award recovery to plaintiff. The court concluded that since the jury found for defendant on the breach of contract claim, it necessarily based its verdict on promises that "were not bargained for and that were not included in the contract."⁴⁷⁴

On defendant's second argument, the court noted with approval the instructions to the jury providing that evidence of extrinsic terms was inadmissible for purposes of plaintiff's contract and estoppel claims unless the jury found that the written contract

468. *Id.* at 115 (citing 37 AM. JUR. 2d *Fraud* § 59 (1968)).

469. *Id.*

470. 515 F. Supp. 612 (S.D. Cal. 1981).

471. *Id.* at 616.

472. *Id.*; see, e.g., *Guaranty Bank v. Lone Star Life Ins. Co.*, 568 S.W.2d 431, 434 (Tex. Civ. App. 1978); *Pasadena Assocs. v. Connor*, 460 S.W.2d 473, 481 (Tex. Civ. App. 1970). The emergence of promissory estoppel as an independent theory of recovery, however, would render such a rule obsolete. See Metzger & Phillips, *supra* note 3, at 549.

473. 515 F. Supp. at 616 (citations omitted).

474. *Id.*

either failed to reflect the parties' real intention due to fraud, mistake, or accident, or that the parties did not intend their writing to be an exclusive statement of their agreement.⁴⁷⁵ The preceding language seems to paint the court's refusal to set aside the jury's verdict as standing for the relatively unexceptionable proposition that parties may prove and courts may enforce on promissory estoppel grounds those parol promises unsupported by consideration and extrinsic to an incomplete, or fraudulently induced, written contract.⁴⁷⁶ The court, however, achieved this result in an interesting fashion. The jury, rather than the trial judge, ruled on the parol evidence question, an approach that emasculates the parol evidence rule and that virtually guarantees enforcement of parol promises that the jury believes the parties actually made.⁴⁷⁷

In two other cases courts nimbly maneuvered around the question of whether a court may use promissory estoppel to protect significant reliance on extrinsic promises that the parol evidence rule otherwise bars, but, nonetheless, found other legal justification for protecting such reliance. In *Prudential Insurance Co. of America v. Clark*,⁴⁷⁸ for example, the court allowed beneficiaries of a life insurance policy issued to a soldier killed in a helicopter crash in Vietnam to recover for his death despite war risk and aviation exclusion clauses in the policy. Prudential's agent induced the insured to drop a prior policy that contained neither exclusion by promising to obtain a similar policy. The insured died after filing an application and prepaying the premium, but before finding out that the policy contained the exclusionary clauses. Pruden-

475. *Id.* at 616-17.

476. For another recent case taking a similar approach, see *Kramer v. Alpine Valley Resort, Inc.*, 108 Wis. 2d 417, 321 N.W.2d 293 (1982). In *Kramer*, a school teacher/artisan leased space in a commercial complex, reduced his teaching load to half-time, and made a significant investment in labor and materials to construct a workshop in reliance upon parol promises that the complex would be open all year and would attract sizeable walk-through traffic. *Id.* at 420, 321 N.W.2d at 294-95. Although his written lease agreement contained neither of these promises, the Wisconsin Supreme Court rejected defendant's argument that the existence of an unambiguous contract between the parties precluded plaintiff's resort to promissory estoppel, holding that "where the contract fails to embody essential elements of the total business relationship of the parties . . . the existence of a contract does not bar recovery under promissory estoppel." *Id.* at 421-22, 321 N.W.2d at 295. The parol evidence rule, the court held, did not prevent proof of the alleged extrinsic promises because the written lease agreement "did not embody the entire business relationship" of the parties and the extrinsic promises "in no way" varied or contradicted "anything in the lease agreement." *Id.* at 426, 321 N.W.2d at 297-98.

477. On the traditional role of the judge in parol evidence cases, see *supra* notes 30-38, 46 & 50 and accompanying text.

478. 456 F.2d 932 (5th Cir. 1972).

tial initially paid the claim, then sued for return of the payment, alleging oversight and mistake.⁴⁷⁹ The court treated the agent's promise as a separate promise independent of the written policy and enforced it on the basis of promissory estoppel, observing that:

[T]his verdict recognized a duty of Prudential, dehors the writing, to act in an honorable and upright way in accordance with its agent's promise. Thus, *application of promissory estoppel in no way trammels upon the parol evidence rule*. Involved here is a separate enforceable promise and not a variance or modification of the terms of the policy.⁴⁸⁰

Similarly, in *Hohenstein v. S.M.H. Trading Corp.*,⁴⁸¹ a stevedore firm's officer made an undisputed oral promise that the cargo vessel which the firm selected to carry the plaintiff shipper's cargo had sufficient capacity to store the shipper's entire cargo. The parties, however, did not incorporate this promise into their written contract. Subsequently, the shipper had to make other arrangements when the stevedore firm could not load all the shipper's cargo in the provided vessel. The court observed that "were this not complicated by an integrated writing constituting 'the' contract, there would be no doubt that this inquiry [about the capacity of the selected vessel] and response and reliance thereon would constitute a classic case analogous to 'promissory estoppel.'" ⁴⁸² Rather than confront the estoppel issue, however, the court held that a fair reading of the written contract revealed a commitment by the stevedore that it could and would load all of the plaintiff's cargo on the specified vessel.⁴⁸³ *

While the equities of the cases and the courts' sense of justice may explain the decisions in the foregoing cases and the courts' failure to devote any significant attention to promissory estoppel's possible application to circumvent the parol evidence rule, three cases furnish more formal, albeit quite limited, support for the reliance doctrine's operation in the parol evidence context. The United States Court of Appeals for the Seventh Circuit's opinion

479. *Id.* at 934-35.

480. *Id.* at 937 (emphasis added). For a case in which the parties never raised promissory estoppel arguments, but in which the court employed similar reasoning to circumvent the parol evidence rule, see *Service Iron Foundry, Inc. v. M.A. Bell Co.*, 2 Kan. App. 2d 662, 669-70, 588 P.2d 463, 470-71 (1978) (sales agent who controlled insolvent principal held to have made independent express warranty to disappointed buyer of pollution control device; parol evidence rule inapplicable to sales agent who was not a party to written sales contract).

481. 382 F.2d 530 (5th Cir. 1967).

482. *Id.* at 535.

483. *Id.* at 536.

in *Ehret Co. v. Eaton, Yale & Towne, Inc.*⁴⁸⁴ contains a curious admixture of both equitable and promissory estoppel principles. In *Ehret*, the plaintiff sales representative for the defendant manufacturer signed a "Commission Sales Agreement" that contained a "duration of agency" clause which gave either party the right to terminate the contract upon thirty days' notice. The contract further provided that in the event of termination the principal would pay commissions only on orders that it accepted prior to termination and that were deliverable within three months from the termination date.⁴⁸⁵ Plaintiff, prior to signing, expressed concern about the termination provision because the nature of the defendant's products required the plaintiff to engage in substantial development work prior to the consummation of a sale. In response to such concern, defendant's General Sales Manager in a letter acknowledged the possible effect of the clause, but assured plaintiff that "in those few cases where the contract has been cancelled by us we have always been much more liberal than provided for in the contract."⁴⁸⁶ He further assured plaintiff that "this discussion is probably academic only." [I]n the very unlikely event of cancellation, you will have to rely on receiving extremely fair treatment."⁴⁸⁷ After receiving these assurances, plaintiff signed the agreement. Eleven months later defendant requested that plaintiff sign a new contract reflecting that defendant had changed its name. The new contract contained the same termination clause and an integration clause cancelling all prior agreements. Twenty-six months after plaintiff signed the new contract, defendant notified plaintiff that it was terminating the agreement and intended to construe the termination clause literally.⁴⁸⁸

Plaintiff filed suit for breach of contract, claiming that the General Sales Manager's letter entitled it to better treatment than defendant had given. Plaintiff succeeded in winning a jury verdict at trial. On appeal, defendant argued that the new contract contained the entire agreement between the parties, but the court upheld the trial judge's holding that defendant's representations "estopped" it from asserting its rights under the termination clause.⁴⁸⁹ The court's language initially pertains to equitable estoppel, as the

484. 523 F.2d 280 (7th Cir. 1975), cert. denied, 425 U.S. 943 (1976).

485. *Id.* at 282.

486. *Id.*

487. *Id.*

488. *Id.*

489. *Id.* at 283-84.

court refers to the "representations" in the General Sales Manager's letter and plaintiff's reliance thereon, and notes that allowing defendant "to disclaim its representations after receiving the benefits therefrom . . . would have the fraudulent effect that an estoppel was designed to prevent."⁴⁹⁰

While defendant's statements that it always had been "more liberal" in past termination cases and that the new contract was necessary to reflect its name change were plainly representations of fact that could be the proper subject for an equitable estoppel, defendant's assurance of "extremely fair treatment" plainly is a promise of future performance and, thus, the proper subject of a promissory estoppel action.⁴⁹¹ Promissory estoppel further enters into the court's reasoning in the following confused statement:

The question of damages was submitted to the jury on the theory of a possible breach of contract to give the plaintiff extremely fair treatment. Support for the treatment as an enforceable [sic] promise of the promise which, as a result of the plaintiff's reliance, creates an estoppel is supported by Restatement, Contracts, § 90 and Restatement 2d, Contracts, § 90 (Tent. Draft No. 2, 1965). Following this view, admission of the "extremely fair treatment" letter is not in conflict with the parol evidence rule as the defendant contends. An estoppel is an equitable remedy with its own independent force.⁴⁹²

Of course, section 90 deals exclusively with promissory estoppel and has nothing whatever to do with equitable estoppel, a point that may have eluded the court. Later discussion by the court on the damage issue, however, clearly indicates that the court is enforcing the "extremely fair treatment" promise as a *promise*.⁴⁹³

Following this tentative endorsement of promissory estoppel, the *Ehret* court blunted the impact of its decision by advancing a more conventional, if not more convincing, rationale for admitting

490. *Id.* at 283. The dissent concurred with the majority's "finding of equitable estoppel," but disagreed with its interpretation of the parties' agreement. *See id.* at 285-86 (Swygert, J., dissenting).

491. On the importance of the "fact/promise" distinction and its role in the evolution of promissory estoppel, see *supra* notes 217-21 and accompanying text.

492. 523 F.2d at 283-84.

493. *See id.* at 284 (quoting *Hoffman v. Red Owl Stores, Inc.*, 26 Wis. 2d 683, 701, 133 N.W.2d 267, 276 (1964), a famous § 90 case discussed *supra* notes 291-304 and accompanying text) for the following statement:

We recognize that there is authority in jurisdictions other than Illinois for the proposition that the promise which becomes the basis for the estoppel is not to be enforced as a promise, and that damages, if awarded, "should be only such as in the opinion of the court are necessary to prevent injustice."

The court, however, subsequently concluded that "in the circumstances of this case, this latter theory produces the same result as determining damages under the contract." 523 F.2d at 284.

the General Sales Manager's letter: the court claimed that the language in the duration of agency clause was ambiguous, and noted that admitting parol evidence is proper in determining the meaning of contract terms.⁴⁹⁴ Thus, the court in a classic instance of judicial legerdemain⁴⁹⁵ has "put the cart before the horse" by admitting an extrinsic term to show an ambiguity that then justifies the term's admission.⁴⁹⁶ While few conclusions unequivocally emerge from the welter of confusion that *Ehret* represents, one thing seems plain enough: despite its unclear reasoning, the *Ehret* court protected plaintiff's reliance on an extrinsic promise the existence of which defendant did not dispute and the terms of which contradicted an unambiguous clause in the parties' contract.

The *Philo Smith & Co. v. USLIFE Corp.*⁴⁹⁷ decision exhibits similar judicial confusion concerning the elements of equitable and promissory estoppel. In *Philo Smith* the court dismissed plaintiffs' contract and quantum meruit claims, holding that the Statute of Frauds and the parol evidence rule barred the claims. Plaintiffs' promissory estoppel claim, however, went to the jury, with the trial

494. 523 F.2d at 284.

495. See *supra* note 149 and accompanying text.

496. On the subject of the "interpretation" exception to the parol evidence rule, see *supra* notes 145-52 and accompanying text.

497. 420 F. Supp. 1266 (S.D.N.Y. 1976), *aff'd per curiam*, 554 F.2d 34 (2d Cir. 1977) (construing New York law). In *Philo Smith* two partners sued to recover a finder's fee resulting from a corporate acquisition. The partners, one James Rutherford and Rodney Hawes, over a period of several years, signed two successive fee agreements to assist defendant in the acquisition of All American Life & Financial Corporation [All American]. Prior to signing the first agreement, Hawes objected to its proposed termination date. Defendant's chief executive officer assured Hawes that the agreement would be extended if necessary. With this assurance, Hawes signed. *Id.* at 1269. The acquisition fell through and Hawes took no further action relating to the USLIFE acquisition of All American. Near the expiration date of the agreement, defendant's CEO told Hawes he was still interested in acquiring All American and sent him a copy of a new fee agreement, which extended the termination date. Hawes again expressed concern about the termination date, but after receiving assurances that defendant would extend the date if they still wanted to acquire All American, the plaintiffs signed. Hawes again met with the CEO near the termination date of the second agreement and suggested that defendant extend the December 31, 1972, termination date. Crosby replied that they would "take care of the paper work if we get anything going after the first of the year," *id.* at 1270, but shortly thereafter told Hawes' former employer that he had no intention of entering another agreement because he was unhappy with Hawes' performance. *Id.*

Some five months after the expiration date of the second agreement the prospects for the acquisition suddenly brightened and defendant's CEO called Rutherford, who had been working continuously on the acquisition, and asked him to arrange a meeting with All American. Rutherford agreed to do so, but suggested that the fee agreement be updated, to which Crosby replied there would be no problem. USLIFE ultimately consummated the acquisition, but refused to pay defendants a finder's fee.

court ultimately granting a directed verdict in defendant's favor.⁴⁹⁸ The court noted that plaintiffs' promissory estoppel claim faced "serious obstacles" in the forms of the Statute and the parol evidence rule,⁴⁹⁹ but acknowledged that the claim possibly could surmount these obstacles when it stated that plaintiffs "had to establish the presence of the elements of promissory or equitable estoppel if they were to avoid the combined effect of the statute of frauds and the parol evidence rule."⁵⁰⁰

The *Philo Smith* court listed five elements that plaintiffs must show before courts will apply promissory estoppel. First, plaintiffs must show "an oral promise . . . made *contemporaneously with or subsequent to* the making of a written agreement."⁵⁰¹ Second, plaintiffs must show that "the promise was fraudulently made."⁵⁰² Third, "defendant must have anticipated that the plaintiffs would rely on the oral promise and such reliance must have been reasonable on the plaintiffs' part."⁵⁰³ Fourth, "plaintiffs must have relied on [the] oral promise by engaging in acts which are 'unequivocally referable' to the oral promise."⁵⁰⁴ Last, "plaintiffs must have suffered substantial injury as a result"⁵⁰⁵ of their reliance. This list is highly objectionable. The first element leaves no real guidance as to the court's view about the effect that the parol evidence rule would have on promissory estoppel claims predicated upon promises made *prior* to the execution of the writing.⁵⁰⁶ While the third and fifth elements properly reside under the umbrella of promissory estoppel, the second element's requirement that the promise be "fraudulently made" sounds more like equitable estoppel,⁵⁰⁷ and the "unequivocally referable" test of the fourth element

498. *Id.* at 1274.

499. *Id.* at 1271.

500. The Court observed that the "impact of the statute of frauds on this case is strengthened by the effect of the parol evidence rule," since the parties had executed two integrated written agreements. *Id.*

501. *Id.* (emphasis added) (citations omitted).

502. *Id.* (citations omitted).

503. *Id.* at 1272 (citations omitted).

504. *Id.* (citations omitted).

505. *Id.* (citations omitted).

506. For Corbin's views on the worth of the "contemporaneous" concept in the integration context, see *supra* note 69 and accompanying text. Of course, the parol evidence rule should play no role whatsoever in the enforcement of extrinsic promises made subsequent to an integration. See *supra* note 65 and accompanying text.

507. See *supra* notes 215-16 and accompanying text. Of course, equitable estoppel would be inappropriate here because the plaintiffs alleged reliance on *promises*, not misrepresentations of *fact* by the defendant.

smacks of the equitable doctrine of part performance.⁵⁰⁸

Plainly, requiring relying promisees to meet such stringent requirements in order to raise successfully a promissory estoppel claim not only is questionable from the standpoint of traditional principles of promissory estoppel, but also presents an insurmountable obstacle in many parol evidence situations. Given the stacked list of elements, the court's finding that plaintiffs had failed to show that any of defendant's promises met the requisite elements⁵⁰⁹ is hardly surprising. The heart of the decision probably rests on the court's determination that "the plaintiffs have completely failed to demonstrate any injury resulting from acts of reliance or that any acts of reliance were 'unequivocally referable' to any oral promises."⁵¹⁰ Only the initial aspect of the court's determination should be relevant. The reliance of many promisees likely is equally referable to the parties' written agreement.⁵¹¹ The requirement of substantial injury stemming from the reliance, unlike the "unequivocally referable" requirement, is a universal element of any formulation of the promissory estoppel doctrine. Perhaps the best explanation of the result in *Philo Smith*, if not of its reasoning, lies with the court's conclusion that plaintiffs had not suffered substantial injury as a result of reliance. If the court had employed the standard enunciated by section 139 of the *Restatement (Second)*,⁵¹² it would have found that no "injustice" would result from denying enforcement to defendant's promises and the same result would have occurred.⁵¹³

508. See *supra* note 348 for the allegation that part performance is nothing more than an earlier, limited variant of the doctrine of equitable estoppel. One commentator doubts that any conduct could meet the "unequivocally referable" test and observes that acts which traditionally serve as evidence of part performance are nothing more than evidence of substantial reliance by the plaintiff. Note, *supra* note 216, at 286. Of course, under promissory estoppel the promise must cause the promisee's reliance. Boyer, *supra* note 222, at 470-71. The exact degree of causal relationship required, however, remains unclear. One author suggests that the promise at issue must play a "major role" in inducing reliance, but need not be the only causal factor. Comment, *Promissory Estoppel in Washington*, 55 WASH. L. REV. 795, 805 (1980). See also *Miller v. Lawlor*, 245 Iowa 1144, 1155, 66 N.W.2d 267, 274 (1954) (unnecessary for plaintiff seeking to circumvent Statute of Frauds via promissory estoppel to show reliance solely based on defendant's promise; sufficient that plaintiff would not have acted without promise).

509. 420 F. Supp. at 1272.

510. *Id.*

511. For a discussion of this point as it relates to the evidentiary value of reliance in parol evidence cases, see *infra* note 554 and accompanying text.

512. For the text of § 139, see *supra* note 421.

513. On appeal, the Second Circuit did little to clarify the trial court's decision. It affirmed the lower court's decision, but only addressed the Statute of Frauds issue, limiting

Five years after the *Philo Smith* decision, *Triology Variety Stores, Ltd. v. City Products Corp.*⁵¹⁴ presented the United States District Court for the Southern District of New York with the opportunity to rectify *Philo Smith's* legacy of conceptual disarray. Unfortunately, the court was only partially successful. In *Triology*, plaintiff filed suit seeking injunctive relief and damages. Plaintiff alleged five causes of action, three sounding in contract and two premised on promissory estoppel. The court dismissed the contract claims for, among other things, failure to comply with the Statute of Frauds, but allowed plaintiff's promissory estoppel claims to stand.⁵¹⁵ The court noted that the defendant "relies heavily" on *Philo Smith* in support of its motion to dismiss, but rejected *Philo Smith's* applicability and criticized its reasoning. The court rejected *Philo Smith's* requirement that reliance furnishing the basis for promissory estoppel be "unequivocally referable" to the alleged agreement⁵¹⁶ by observing that it had not found approval either by the Second Circuit or by the New York State courts, and by correctly noting that such a requirement more appropriately was asso-

its holding to the plaintiff's failure to prove substantial injury and expressing doubts about the lower court's other holdings. See *Philo Smith & Co. v. USLIFE Corp.*, 554 F.2d 34 (2d Cir. 1977) (per curiam).

514. 523 F.Supp. 691 (S.D.N.Y. 1981) (construing New York law). In *Triology* defendant City Products Corporation, a franchisor of "Ben Franklin" stores, secured a ten-year lease, including options to renew for either an additional ten-year period or two successive five-year periods, on certain premises located in New York City. City Products subsequently entered two contracts with plaintiff Triology Variety Stores. The first contract was a franchise agreement conferring upon plaintiff's owners the right to use the "Ben Franklin" name and to sell franchised goods to the public. The second contract was a sublease of the premises that defendant previously leased for the store. The sublease was to expire sixteen days prior to the expiration of City Products' lease. The parties renewed both the lease and sublease at the end of the first ten-year period for an additional five years. In November of 1978, with twenty-six months remaining on the lease and sublease, an investor, after City Products screened and approved him, purchased Triology and entered into a new franchise agreement with a termination date of December 31, 1983. Prior to the purchases, the investor expressed concern about a clause in the franchise contract that allowed the franchisor to terminate the franchise in the event that the franchised store moved to a new location. The investor feared that this clause could cause his business to terminate at the end of the sublease if City Products failed to renew the lease and sublease. He, therefore, sought and allegedly received promises from various agents and employees of City Products that as long as the franchise remained in good standing City Products would renew the sublease. City Products, however, later refused to renew either the lease or the sublease. 523 F. Supp. at 693-94.

515. *Id.* at 694-96. Notably, the court's action in dismissing the plaintiff's contract claims but allowing his promissory estoppel claims to survive is consistent with treating promissory estoppel as a theory of recovery independent of contract. See Metzger & Phillips, *supra* note 3, at 510, 512.

516. See *supra* notes 504, 508 & 510-11 and accompanying text.

ciated with the doctrine of part performance. The court similarly rejected *Philo Smith's* fraud requirement⁵¹⁷ with the observation that "New York State and federal courts have frequently applied the doctrine of promissory estoppel in the absence of a finding of fraud."⁵¹⁸ In the end, the court found that plaintiff and its owner "have set forth a valid claim of substantial injury for which it would be unconscionable to deny them the benefits of the promises upon which they allegedly relied."⁵¹⁹

Plainly, *Triology* serves to clarify the elements of promissory estoppel under New York law. Since the court's discussion solely concerns promissory estoppel as a means of defeating the Statute of Frauds, however, it sheds no light on the court's views about whether promissory estoppel likewise may serve to defeat the parol evidence rule. Arguably, no significant parol evidence question may be present in *Triology*.⁵²⁰ Promissory estoppel's future as a device for circumventing the parol evidence rule in the Southern District of New York, then, clearly awaits further judicial clarification.

Although the foregoing cases do not lend much significant support for the proposition that promissory estoppel is a worthy candidate for recognition in the parol evidence context, a number of considerations indicate a greater role for the doctrine in future parol evidence cases. The following section sets forth the arguments militating in favor of so expanding the reliance principle and attempts to anticipate and respond to counter-arguments that opponents likely will raise to any such expansion. The section, furthermore, briefly outlines the manner in which promissory estoppel's extension into the parol evidence context should proceed. This latter effort is, by nature, particularly exploratory and is only an at-

517. See *supra* notes 502 & 507 and accompanying text.

518. 523 F. Supp. at 698 (footnote omitted).

519. *Id.*

520. Certainly, conventional arguments support the proposition that the admission of evidence concerning defendant's promises would not violate the parol evidence rule. City Product's oral promises concerning the written sublease, which the new owner acquired when he purchased *Triology*, were made subsequent to the initial execution of the sublease; potentially therefore, introduction of evidence of these promises might not contravene the parol evidence rule. See *supra* note 72 and accompanying text; see also *supra* note 513. The oral promises concerning the written franchise, although clearly made prior to the execution of the writing, could have represented a "collateral" agreement that was separate and distinct from the franchise contract. See *supra* note 96 and accompanying text. Alternatively, the franchise contract arguably was not a complete integration of the parties' agreement on the franchisor's obligations concerning renewal of the sublease. Whether the court failed to discuss the parol evidence issue for one of these reasons or because defendant simply failed to raise it, is unclear.

tempt to raise some of the issues that future courts and commentators must clarify.

B. Should Promissory Estoppel Apply to Parol Evidence Cases?

The parol evidence rule and the Statute of Frauds, as mentioned earlier, are strikingly similar. The same underlying policies support both doctrines:⁵²¹ prevention of perjury by excluding presumptively untrustworthy oral testimony⁵²² and the desire to give judges increased control over the jury.⁵²³ Both are rules of form⁵²⁴ whose efficacy as such some commentators have questioned.⁵²⁵ The criticisms directed at the Statute and the rule are remarkably similar. Scholars criticize both for causing injustice because they may prevent proof of agreements that the parties actually made.⁵²⁶ Both, critics note, have been the source of an enormous volume of litigation.⁵²⁷ Similarly, observers argue that the reasons for adopting the Statute and the parol evidence rule no longer are compelling.⁵²⁸ Finally, the courts have meted out similar treatment to both the Statute and the rule: erosion by numerous judicially-created exceptions⁵²⁹ and frequent manipulation by judges seeking just results.⁵³⁰ In sum, critics of both legal doctrines are abundant and these critics frequently suggest consigning the Statute and the rule to oblivion.⁵³¹

The similarities between the Statute and the rule, the judicial trend of circumventing the Statute by promissory estoppel, and the historically expansionist nature of the reliance doctrine suggest that the parol evidence rule is a likely candidate for the application of promissory estoppel principles. Certainly, the reasoning

521. Wallach, *supra* note 17, at 653.

522. See *supra* notes 27-28 & 320-21 and accompanying text; see also J. CALAMARI & J. PERILLO, *supra* note 20, § 3-3, at 109-10.

523. See *supra* notes 32-38 & 322-25 and accompanying text; see also Sweet, *supra* note 21, at 1053.

524. See J. CALAMARI & J. PERILLO, *supra* note 20, § 3-3, at 108; see also Sweet, *supra* note 21, at 1053-54.

525. See *infra* notes 559-63 and accompanying text.

526. See *supra* notes 184-89, 327 & 333-36 and accompanying text. Seavey, *supra* note 209, at 924 ("refusal to admit such evidence causes the Statute of Frauds and the parol evidence rule to be a trap").

527. See *supra* notes 164 & 328 and accompanying text.

528. See *supra* notes 167-73, 190 & 329-31 and accompanying text.

529. See *supra* notes 124-52 & 339-40 and accompanying text.

530. See *supra* notes 153-61 & 341 and accompanying text.

531. See *supra* notes 190 & 332 and accompanying text.

that justified estoppel's use in the Statute of Frauds context is applicable equally in parol evidence cases. The argument that estoppel's use to circumvent the Statute was necessary to prevent the Statute from serving as an agent of injustice, an argument at the heart of early estoppel cases involving the Statute,⁵³² similarly could serve to justify estoppel's use in preventing the parol evidence rule from operating as an agent of injustice.⁵³³ As one recent commentator observed, "[i]f it is reasonable to rely on an entirely oral agreement, it must also be reasonable in some cases to rely on assurances that a writing is not necessary to preserve a particular term of [an] agreement."⁵³⁴ Furthermore, the authors' inclusion of section 139⁵³⁵ in the *Restatement (Second)* "indicates an increasing willingness to subordinate form to substance, where justice requires,"⁵³⁶ a tendency which led one observer to conclude that "it seems inevitable that another bastion of form-over-substance, the parol evidence rule, eventually will fall under similar attack."⁵³⁷

Several other considerations also militate in favor of promissory estoppel's expansion into the parol evidence context. As previously observed, a general trend exists toward liberality in the enforcement of the parol evidence rule itself,⁵³⁸ a trend which indicates a generally increased judicial willingness to admit evidence of terms extrinsic to written contracts and a diminished level of judicial concern for the doctrine's rule-of-form aspects.⁵³⁹ The courts' observed tendency to manipulate the rule to obtain just results further supports promissory estoppel's expansion.⁵⁴⁰ This tendency should make at least some courts consider applying a doctrine such as promissory estoppel, which openly and avowedly seeks just ends. An expansion of promissory estoppel's scope would be consistent with many observed tendencies in twentieth century contract law in general,⁵⁴¹ and with the modern fate of rules of form in particular.⁵⁴² Finally, if promissory estoppel ultimately

532. See *supra* notes 347-48 and accompanying text.

533. See *supra* notes 184-89 and accompanying text.

534. Knapp, *supra* note 5, at 78.

535. See *supra* notes 421-30 and accompanying text.

536. Knapp, *supra* note 5, at 78.

537. *Id.*

538. See *supra* note 118 and accompanying text.

539. See *infra* notes 558-65 and accompanying text.

540. See *supra* notes 119-23 and accompanying text.

541. See *supra* notes 207-08, 210, 213 & 287-88 and accompanying text.

542. See, e.g., Sweet, *supra* note 21, at 1054 ("Rules of form have a poor performance record in American law.").

gains recognition as a theory of recovery independent of contract,⁵⁴³ courts could argue that the parol evidence rule, a creature of contract, is simply inapplicable to claims premised on promissory estoppel.⁵⁴⁴

Given the similarities between the parol evidence rule and the Statute of Frauds, many of the concerns raised in criticism of promissory estoppel's circumvention of the Statute also should apply to extensions of the doctrine into the parol evidence context. Of course, the rebuttal arguments that support the use of promissory estoppel in the Statute of Frauds context apparently outweigh the concerns, and these arguments should prove equally persuasive when applied to the parol evidence rule. Indeed, because of the difference in the roles that the Statute and the rule play, proponents of promissory estoppel's application to parol evidence cases face an easier task than the proponents of the doctrine's application in the Statute of Frauds context. The judicial fears concerning abrogation of the Statute that observers often cite as a source of the resistance which promissory estoppel's incursion into the territory of the Statute has encountered⁵⁴⁵ should not be as great an obstacle to estoppel's circumvention of the judicially created parol evidence rule. The only major statutory manifestation of the rule is section 2-202 of the Uniform Commercial Code.⁵⁴⁶ The Code's section 1-103 facilitates estoppel's application to the Statute of Frauds;⁵⁴⁷ this section could justify promissory estoppel's application to the Code's parol evidence rule just as easily.

Nonetheless, application of promissory estoppel to the parol evidence rule is certain to provoke judicial fears, similar to the concerns raised in the Statute of Frauds context,⁵⁴⁸ concerning the loss of the protection that the rule affords written contracts. These fears include the increased threat of perjury, the loss of significant evidentiary value, the elimination of the rule's channeling and cautionary functions, the undermining of the rule's role as a source of predictability and certainty in commercial transactions, and the loss of the rule's capability of protecting the parties' intent to

543. See *supra* notes 297-308 & 431-35 and accompanying text.

544. See *supra* note 434 and accompanying text.

545. See *supra* notes 376-80 and accompanying text.

546. See *supra* notes 99-106 and accompanying text.

547. See *supra* notes 387-89 and accompanying text. See also J. WHITE & R. SUMMERS, *supra* note 27, § 2-11 (Judicially created exceptions to the parol evidence rule are still valid under § 1-103 except as specifically precluded by the language of § 2-202.).

548. See *supra* notes 390-91 and accompanying text.

finalize their agreement in an integrated writing. If the primary motivation for such fears is the prospect of perjury, the doubts jurists voiced about the efficacy of the Statute's writing requirement as a device for preventing perjury are equally applicable to the parol evidence rule. The writing as easily could be a forgery in a parol evidence case⁵⁴⁹ as in a Statute of Frauds case.⁵⁵⁰ Given the number of exceptions to both the Statute and the rule, a dishonest plaintiff as easily could "frame" a contract within an exception to the rule as he could within an exception to the Statute.⁵⁵¹ Arguably, less danger of perjury exists in parol evidence cases than in Statute of Frauds cases; in Statute of Frauds cases the basic issue concerns the existence of a contract, generally a given in parol evidence cases in which the central issue is whether the parties may vary or supplement an existing written contract with alleged extrinsic terms.⁵⁵²

While the existing writing, then, provides a convenient point of reference against which a court may measure the believability of the proffered extrinsic term in parol evidence cases,⁵⁵³ a promisee's reliance often may provide substantially less evidentiary value in parol evidence cases than in Statute of Frauds cases.⁵⁵⁴ Certainly, in some parol evidence cases the nature of the promisee's reliance plainly will be referable to the existence of the proffered extrinsic term; in many cases, however, reference to the parties' written contract may explain the parties' reliance equally well. Any sensible application of promissory estoppel in the parol evidence context plainly should consider the evidentiary value of the promisee's reliance in determining whether granting relief premised on reliance is appropriate.⁵⁵⁵ Notably, Corbin's view of integration, which has been influential in current interpretations of the parol evidence rule, already has eroded substantially whatever special evidentiary protection the rule once afforded written contracts.⁵⁵⁶ Thus, little real residual protection likely remains that application of promis-

549. See *supra* note 180 and accompanying text.

550. See *supra* note 394 and accompanying text.

551. See *supra* notes 392-98 and accompanying text.

552. See Note, *supra* note 15, at 984.

553. *Id.* Since some formulations of the parol evidence rule concern prior written terms, some parol evidence cases may exist in which the parties unquestionably agreed to the proffered extrinsic term at least at some point in the negotiation process. See *supra* notes 62-63 and accompanying text.

554. See *supra* notes 400-01 and accompanying text.

555. See *infra* notes 595-97 and accompanying text.

556. See *supra* notes 96-115 and accompanying text.

sory estoppel to circumvent the rule would erode.

Commentators who believe that the parol evidence rule properly functions as a rule of form⁵⁵⁷ likely will view any incursion by promissory estoppel into the rule's province as a development that deleteriously will affect the rule's ability to fulfill the normal rule-of-form functions.⁵⁵⁸ Although this undeniably would occur, critics raised similar objections to estoppel's circumvention of the Statute of Frauds. Proponents of this extension of the doctrine largely overcame the objections by noting that the Statute did not work well as a rule of form⁵⁵⁹—questioning its efficacy either as a cautionary⁵⁶⁰ or channeling device.⁵⁶¹ These arguments are more persuasive when applied to the parol evidence rule, which has no significant role as a cautionary device and whose channeling abilities⁵⁶² critics question.⁵⁶³ Modern interpretations of the rule, in any event, have reduced dramatically its potential to operate as a rule of form.⁵⁶⁴ Finally, under contemporary versions of promissory estoppel, partial enforcement of extrinsic agreements through reliance-based damage awards is available. This option can preserve a modest channeling function for the parol evidence rule in much the same manner as it has done for the Statute of Frauds.⁵⁶⁵

Critics of the doctrine's extension also may argue that application of promissory estoppel in parol evidence cases would undermine the rule's vaunted, but often criticized,⁵⁶⁶ role as a source of predictability and certainty in commercial transactions.⁵⁶⁷ This argument's veracity depends upon whether the rule as currently applied provides some real measure of stability. In view of the numerous ways by which courts now circumvent the rule, questions exist about whether the rule truly provides commercial transactions with a significant measure of predictability.⁵⁶⁸

The last and most significant challenge to using promissory es-

557. See *supra* note 39 and accompanying text.

558. Observers who view parol evidence's rule-of-form functions as significant also must view as deleterious the numerous existing means for circumventing the rule and its frequent distortion by the courts.

559. See *Sweet*, *supra* note 21, at 1053.

560. See *supra* note 345 and accompanying text.

561. See *supra* notes 406-07 and accompanying text.

562. See *supra* notes 40-42 and accompanying text.

563. See *supra* note 182 and accompanying text.

564. See *supra* note 98 and accompanying text.

565. See *supra* notes 429-30 and accompanying text.

566. See *supra* note 177 and accompanying text.

567. See *supra* notes 45-46 and accompanying text.

568. See *supra* note 178 and accompanying text.

toppel to circumvent the parol evidence rule arises because the rule's modern defenders justify it for its role in protecting the parties' intent to finalize their agreement in an integrated writing.⁵⁶⁹ The effect that circumventing the rule via promissory estoppel would have on its ability to perform this admittedly salutary function depends on how a particular court determines the parties' intent to integrate their agreement. In a jurisdiction that adopts a Willistonian approach to integration, courts focus on the writing itself as the major consideration in determining the parties' intent to integrate. Formal recognition of promissory estoppel as a device for circumventing the rule in appropriate cases could provide the courts in such a jurisdiction with doctrinal justification for admitting proof of extrinsic terms to which they believe the parties actually agreed, even though those terms are such that similarly situated parties might "naturally and normally" have included in the writing.⁵⁷⁰

Admission of such terms under the ambit of estoppel could serve to minimize the rule-of-form aspects of the Willistonian approach, which otherwise can operate to frustrate the parties' true intent in the name of their presumed intent.⁵⁷¹ Plainly, if properly applied, promissory estoppel would in some cases facilitate admission of extrinsic terms to which the parties in fact did agree. The rule otherwise would bar such evidence. Thus, promissory estoppel's application could effectuate the parties' actual intent. Courts, however, that consciously follow the Willistonian approach, as opposed to courts that simply follow Williston out of blind adherence to precedent, arguably do so because they still perceive either that the rule so interpreted serves a valuable function as a rule of form,⁵⁷² or that the rule provides written contracts with a degree of deserved protection. Such a court would be loath to circumvent the parol evidence rule via estoppel because doing so would erode the rule's perceived ability to perform this valuable function.

Courts embracing Corbin's views on integration already attempt to determine the parties' subjective intent; these courts—willing to consider all available evidence on the issue of intent and to base admission solely on the basis of the evidence's

569. See *supra* notes 48-57 and accompanying text.

570. See *supra* notes 80-87 and accompanying text.

571. See *supra* notes 88-90 and accompanying text.

572. Even the Willistonian approach represents a diminution of the rule-of-form aspects of the parol evidence rule when compared with the now defunct "four corners" doctrine. See Wallach, *supra* note 17, at 656-60.

credibility⁵⁷³—already largely have abandoned the parol evidence's rule-of-form aspects.⁵⁷⁴ In these jurisdictions, the issue becomes whether promissory estoppel really could offer any utility to such a court. The proponent of an alleged extrinsic term seeking to justify its admission on the grounds of estoppel presumably would have to offer the court convincing evidence of the existence of the term. Conceivably, the proponent also would have to offer evidence of either a promise to incorporate the term into the writing or an assurance that incorporation into the writing is unnecessary to preserve the term as an element of the parties' agreement.⁵⁷⁵ A court adopting the Corbin approach as easily could conclude that the writing was a partial integration of the parties' agreement. Such a court would then admit evidence of the proffered term on this more traditional ground without using promissory estoppel.

The Corbin-following court, however, could not handle similarly a proffered term that plainly "contradicts" the writing but to which the court, nonetheless, believes the parties assented. Some courts might be willing to follow Corbin's lead to its logical conclusion and reject such a writing on the grounds that it lacks finality.⁵⁷⁶ Others, however, might feel constrained by the longstanding general rule that only "consistent" extrinsic terms may supplement partially integrated writings.⁵⁷⁷ If the proffered term contradicted the writing, then the court either would have to reject the term on the basis of inconsistency or resort to some difficult, though far from uncommon,⁵⁷⁸ judicial gymnastics aimed at depicting the proffered term as consistent with the writing. The availability of promissory estoppel as a means of circumventing the parol evidence rule, in such a case, could provide doctrinal rationale for admitting contradictory extrinsic terms that the court, nonetheless, believes the parties intended to stand. The court thus would avoid frustrating the parties' true intent by excluding such a term and would obviate the need for judicial subterfuge justifying admission by pretending consistency.⁵⁷⁹

573. See *supra* notes 91-94 and accompanying text.

574. See *supra* note 98 and accompanying text.

575. See *infra* notes 595-97 & 599 and accompanying text.

576. Corbin advocated abandonment of the "partial integration" idea on the ground that parties rarely intend to consider an incomplete writing final. See 3 A. CORBIN, *supra* note 48, § 581, at 441. See also J. CALAMARI & J. PERILLO, *supra* note 20, § 3-2, at 101 n.16.

577. See *supra* note 120 and accompanying text. For a discussion of the pitfalls associated with this issue, see *supra* notes 121-22 and accompanying text.

578. See *supra* note 123 and accompanying text.

579. But see *Mack v. Earle M. Jorgenson Co.*, 467 F.2d 1177, 1179 (7th Cir. 1972).

In addition to these concerns for the potentially deleterious effects on the salutary functions of the rule, two other, more pragmatic, problems exist that proponents of the change must confront before promissory estoppel ultimately gains formal judicial recognition in parol evidence cases. First, few cases exist in which the courts have discussed promissory estoppel's availability in the parol evidence context; furthermore, these cases offer scant support for the idea that courts may utilize estoppel in proper circumstances, to circumvent the parol evidence rule. The scarcity of such cases, however, is due in part to the proposition's novelty and the consequent unfamiliarity of courts and counsel with the arguments that support it. Furthermore, although the few cases in which courts have raised the promissory estoppel issue provide little authority for the proposition's viability, none of these cases evidences any thorough judicial analysis of the various arguments either in favor of, or against, circumventing the parol evidence rule by estoppel. Thus, while the proponents of estoppel's eventual intrusion into the sphere of the parol evidence rule cannot show convincing precedent in favor of their proposition, they need not confront any well-considered authority in opposition to it.

The second problem, however, is a more troubling argument that opponents could advance both to explain the current failure of promissory estoppel principles to make significant inroads into the parol evidence context and to argue that any future formal recognition of estoppel as a device for circumventing the parol evidence rule is unlikely: certain *de facto* aspects of the parol evidence rule's operation exist that conceivably could render superfluous any application of promissory estoppel. If, as some of the rule's critics assert, the only genuine issue in most parol evidence cases is whether the parties actually agreed to the alleged extrinsic term⁵⁸⁰ and most courts in such cases seek "just" results by selecting as authority whichever formulation of the rule and its numerous exceptions reach the just result,⁵⁸¹ then a serious question arises concerning whether a controversial doctrine such as promissory estoppel offers any advantages. Certainly, an excellent explanation of why promissory estoppel, which has evidenced an historical tendency toward ever-increasing intrusion into the prov-

580. See *supra* note 58 and accompanying text.

581. See *supra* notes 153-54 and accompanying text. For a well-known assertion that this process of reasoning backwards to justify tentatively formed conclusions is actually descriptive of most judicial decision-making, see J. FRANK, *LAW AND THE MODERN MIND* 100-17 (1930).

inces of traditional contract law, has yet to gain significant formal recognition in the parol evidence context is that many courts long have afforded protection to parties relying on oral promises. Although rigorous application of the parol evidence rule would bar proof of these promises, courts reach just results by manipulating the rule itself.⁵⁸²

Several rejoinders to the argument of superfluosity are available to proponents of estoppel's intrusion into parol evidence rule cases. First, from an historical standpoint, judicial protection of reliance in a variety of contexts via various manipulations of traditional contract principles preceded the first *Restatement's* formal recognition of promissory estoppel.⁵⁸³ Promissory estoppel provided a clearer doctrinal explanation for the results in these contexts and probably could provide a similar service in many parol evidence cases. Thus, the availability of other, less straightforward means of protecting reliance did not prevent ultimate recognition of promissory estoppel's applicability in other legal contexts, and no compelling reason exists to suspect that it will do so for the parol evidence rule.

Second, by openly admitting that they will use promissory estoppel to circumvent the strictures of the parol evidence rule when necessary to avoid injustice, courts could provide the parol evidence rule with a much needed measure of conceptual clarity. Apart from the general benefits associated with conceptual clarity,⁵⁸⁴ formal recognition of the reliance principle in parol evidence cases could reduce the inconsistency that characterizes the parol evidence rule's current operation and thereby enhance the predictability that the rule affords. These benefits would help cure aspects of the rule that have been frequent targets of criticism.⁵⁸⁵ Clearly, the availability of estoppel would have a beneficial effect on judicial treatment of the various exceptions to the rule, which are themselves the product of judicial dissatisfaction with the rule's potential for injustice⁵⁸⁶ and serve as a frequent source of inconsistent judicial application.⁵⁸⁷ Courts whose judicial arsenal included

582. See, e.g., the cases discussed *supra* notes 470-83 and accompanying text.

583. See *supra* notes 223-37 and accompanying text.

584. "[A] body of law is more rational and more civilized when every rule it contains is referred articulately and definitely to an end which it subserves, and when the grounds for desiring that end are stated or are ready to be stated in words." Holmes, *supra* note 192, at 469.

585. See *supra* notes 164-66 and accompanying text.

586. See *supra* note 124 and accompanying text.

587. See *supra* notes 130, 132, 137 & 149 and accompanying text.

promissory estoppel as a device for furthering justice in parol evidence cases simply would have little need to indulge in the circuitous, and intellectually dishonest, process of torturing the facts of a case to bring the case within the ambit of an established exception to the rule. If ultimate results in parol evidence cases clearly arise from the court's perceptions of the situation's equities rather than from the arcana of the parol evidence rule, then predictability and certainty would result from the consequent focusing on the real issues involved.

Last, not all courts have indulged in manipulation of the parol evidence rule to further the interests of fairness. Some courts mechanically apply the rule to prevent proof of extrinsic terms that the court may believe the parties actually agreed to and intended to supplement the writing.⁵⁸⁸ To the extent that such judicial behavior reflects timidity, rule-orientation rather than result-orientation, or an unwillingness to participate in the subterfuges in which more result-oriented courts indulge, the formal extension of promissory estoppel's reach into the parol evidence context could provide such courts with an acceptable doctrinal justification for avoiding injustice.

The rudiments of a version of promissory estoppel suitable for application in parol evidence cases already exist in section 139 of the *Restatement (Second)*, which allows courts to enforce oral promises despite their failure to conform to the Statute of Frauds.⁵⁸⁹ Section 139 and its attendant comments bear witness to its drafters' careful consideration of the multiplicity of factors that a court should consider carefully before employing promissory estoppel to circumvent a rule of form such as the Statute or the parol evidence rule. A comment to the section admonishes courts that parties seeking to take advantage of the section should bear a greater burden than their counterparts seeking relief under section 90.⁵⁹⁰ The second subsection of section 139 lists numerous factors that are "significant" in "determining whether injustice can be avoided only by enforcement of the promise."⁵⁹¹ Most of these considerations have some relevance in parol evidence cases as well; the relevant factors include "the definite and substantial character of

588. See *supra* note 155 and accompanying text.

589. See *supra* note 421.

590. "Like § 90 this Section states a flexible principle, but the requirement of consideration is more easily displaced than the requirement of a writing." *RESTATEMENT (SECOND) OF CONTRACTS* § 139 comment b (1981).

591. See *supra* note 421.

the action or forbearance in relation to the remedy sought,"⁵⁹² "the reasonableness of the action or forbearance,"⁵⁹³ and the "extent to which the action or forbearance was foreseeable by the promisor."⁵⁹⁴

Of particular relevance in parol evidence cases, however, is "the extent to which the action or forbearance corroborates evidence of the making and terms of the promise, or the making and terms are otherwise established by clear and convincing evidence."⁵⁹⁵ While the corroboration element may not have as much practical utility in parol evidence cases as it does in most Statute of Frauds cases,⁵⁹⁶ the drafters could modify the subsection for parol evidence cases to allow the existing writing to serve as a yardstick against which courts may test the credibility of the proffered extrinsic term.⁵⁹⁷ In addition, any formal promulgation of a parol evidence-rule version of promissory estoppel could admonish courts to consider the degree of formality surrounding the execution of the writing, the relative sophistication of the parties, and any evidence indicating that the writing is the product of an abuse of unequal bargaining power. Such factors are of obvious potential relevance in reestablishing the environment in which the parties made their agreement, and have some bearing not only on the degree of credibility the court should afford evidence of alleged extrinsic terms, but also upon the justice of the result sought by the proponent of the extrinsic term. In any event, these factors arguably play a *sub rosa* role in many courts' decisions regardless of whether courts afford them formal recognition.⁵⁹⁸

Several other observations can be made about the form promissory estoppel is likely to assume in parol evidence cases and the impact its application is likely to have on the operation of the parol evidence rule. First, presumably courts that presently restrict

592. RESTATEMENT (SECOND) OF CONTRACTS § 139(2)(b) (1981).

593. *Id.* § 139(2)(d).

594. *Id.* § 139(2)(e).

595. *Id.* § 139(2)(c). Such an approach is obviously consistent with those proposals for reformation which suggest that courts compel proponents of extrinsic terms to prove their existence by "clear and convincing" evidence. *See supra* note 144 and accompanying text.

596. *See supra* note 554 and accompanying text.

597. *See supra* note 553 and accompanying text. Since these factors merely play an admonitory role rather than serving as conclusive determinants of the availability of relief, inconsistencies between the writing and the proffered extrinsic term should engender a measure of judicial wariness. Such inconsistencies, however, should not serve as a bar to enforcement if the court, nonetheless, believes in the legitimacy of the extrinsic term. *See supra* notes 576-79 and accompanying text.

598. *See supra* notes 158-61 and accompanying text.

promissory estoppel's application in Statute of Frauds cases to those cases concerning ancillary promises would impose similar restrictions on its application in parol evidence cases. Thus, these courts probably would insist on proof of a promise to incorporate the alleged extrinsic term into the writing or a promise not to raise the parol evidence rule as a defense to subsequent proof of such a term. Such a limitation of the reliance doctrine's application in the parol evidence context, however, would be subject to the same criticisms that critics have lodged against the limitation in the Statute of Frauds context.⁵⁹⁹ Courts, therefore, should avoid such limitations. Second, as previously observed,⁶⁰⁰ the potential for partial enforcement of an extrinsic term, which modern versions of promissory estoppel afford, could allow the parol evidence rule to continue to perform a channeling function, albeit at a reduced level. Last, the extension of promissory estoppel into the parol evidence rule's domain would not operate necessarily to deprive judges of the rule's jury control aspects, which allow the courts—without appearing to rule on the credibility of the party proffering evidence⁶⁰¹—to exclude from jury consideration evidence of extrinsic terms whose credibility the court doubts. A court could maintain some jury control because while the issues of whether the promisor made a promise that he reasonably should expect to induce reliance and whether the promise actually induced reliance are questions of fact for the jury, the issue of whether enforcement of the promise is the only way to avoid injustice is a discretionary question of policy for the court to decide.⁶⁰² Thus, a court that remains unconvinced in the face of evidence that the jury has found convincing still retains a point of entry to judicial corrective action, albeit at a later stage in the proceedings.⁶⁰³

599. See *supra* notes 358-70 and accompanying text.

600. See *supra* note 565 and accompanying text.

601. See *supra* notes 36-38 and accompanying text.

602. *Gruen Indus., Inc. v. Biller*, 608 F.2d 274, 280 (7th Cir. 1979); *Kramer v. Alpine Valley Resort, Inc.*, 108 Wis. 2d 417, 422, 321 N.W.2d 293, 296 (1982); *Hoffman v. Red Owl Stores, Inc.*, 26 Wis. 2d 683, 698, 133 N.W.2d 267, 275 (1965). At least one court has held that there is no right to a jury trial in promissory estoppel cases because promissory estoppel is an equitable doctrine. See *C & K Eng'g Contractors v. Amber Steel Co.*, 23 Cal. 3d 1, 11, 587 P.2d 1136, 1141, 151 Cal. Rptr. 323, 328 (1978). For a discussion of *C & K* and a criticism of its holding, see Note, *supra* note 208, *passim*.

603. While this obviously concerns judicial subterfuge of the sort this Article criticizes, observers who see the rule's primary utility as a device for permitting such subterfuge hardly can complain.

VI. CONCLUSION

The parol evidence rule long has been the deserving recipient of criticism—criticism aimed at the confusion surrounding its bases in policy, the inconsistencies in its formulation, and the vagaries in its application. Several extant considerations suggest that the rule is both a likely and a proper candidate for the meliorating ministrations of the doctrine of promissory estoppel. Prominent among these considerations are the inherently expansionist nature of the reliance principle, the already well-established tendency of courts to circumvent the Statute of Frauds using promissory estoppel, the numerous similarities between the Statute and the rule, the modern trend toward liberality in the interpretation and application of the rule, and the possibility that promissory estoppel eventually will gain recognition as a separate theory of recovery independent of contract.⁶⁰⁴

An examination of the issues related to any eventual intrusion by promissory estoppel into the domain of the parol evidence rule indicates that the consequences attendant to such a development generally would be salutary in nature. Estoppel could enhance the effectuation of the parties' true intent in a written contract, an objective shared by modern formulations of the parol evidence rule, and could minimize the injustice associated with mechanical applications of the rule. Estoppel also could provide a better doctrinal explanation for the results courts reach in many parol evidence cases than the traditional parol evidence rubric that courts currently employ affords. Such doctrinal clarity would result in the focusing of judicial attention in parol evidence cases on the real issues that ultimately will determine whether a court will give legal effect to an extrinsic promise, a consequence likely to provide a much needed measure of clarity and predictability in the administration of the parol evidence rule. Finally, promissory estoppel's application in parol evidence cases would be consonant with the tendency of twentieth century contract law to elevate substance over form in the pursuit of just results, a tendency of which the reliance principle is merely one manifestation. Thus, although the extant evidence is admittedly far from conclusive and exercises in prediction are fraught with obvious perils, one addressing the parol

604. For a discussion of the numerous factors upon which such an eventuality depends, see Metzger & Phillips, *supra* note 3, at 553-56.

evidence rule might borrow from Lord Tennyson and say:

There's a new foot on the floor, my
friend,
And a new face at the door, my
friend,
A new face at the door.⁶⁰⁵

605. A. TENNYSON, *The Death of the Old Year*, in *THE WORKS OF ALFRED LORD TENNYSON* 62 (1896).

The First Amendment and Nonpicketing Labor Publicity Under Section 8(b)(4)(ii)(B) of the National Labor Relations Act

*Lee Goldman**

Throughout the turbulent labor history of this nation, the primary economic strike has been considered the union's ultimate weapon for bringing about recognition, or for forcing an employer to accede to its collective bargaining demands. Recent events, however, suggest that, at least with regard to the most bitter labor-management disputes, the strike weapon may be replaced, or at least substantially supplemented, by secondary boycott activity.¹

I.	INTRODUCTION	1470
II.	BACKGROUND	1472
	A. <i>The Statute</i>	1472
	B. <i>The Cases</i>	1473
	1. The "Producer-Distributor" Requirement .	1473
	2. The "For the Purpose of" Requirement ..	1479
III.	THE FIRST AMENDMENT PROTECTION OF NONPICKET- ING LABOR PUBLICITY	1480
	A. <i>Labor Picketing versus Nonpicketing Labor Publicity</i>	1481
	1. "Speech Plus" Theory	1482
	2. "Unlawful Objective" Theory	1486
	B. <i>Commercial Speech versus Labor Speech</i>	1487
	1. The Commercial Speech Analogy	1488
	2. First Amendment Values	1489

* Visiting Assistant Professor of Law, University of Detroit School of Law; J.D., 1979, Stanford Law School; B.A., 1976, Queens College.

1. 117 Daily Lab. Rep. (BNA) D-1 (June 18, 1981) (remarks of NLRB member Don Zimmerman before New York University's National Conference on Labor, June 10-12, 1981). *Accord*, Kamer, *The 8(b)(4) Publicity Proviso and NLRB v. Servette: A Supreme Court Mandate Ignored*, 16 GA. L. REV. 575, 575 (1982).

(a)	<i>Individual Self-Fulfillment</i>	1490
(b)	<i>Advancement of Knowledge and Discovery of Truth</i>	1492
(c)	<i>Participation in Decisionmaking</i> . .	1494
(d)	<i>Balance Between Stability and Change</i>	1496
IV.	JUSTIFYING THE RESTRICTIONS ON PROTECTED SPEECH	1498
A.	<i>Content-Based Restrictions</i>	1499
B.	<i>The Government's Interest in the "For the Purpose of" Requirement</i>	1500
C.	<i>The Government's Interest in the "Producer- Distributor" Requirement</i>	1502
V.	A LESS RESTRICTIVE ALTERNATIVE	1506
VI.	CONCLUSION	1508

I. INTRODUCTION

Labor's increased efforts to induce secondary consumer boycotts² have drawn into question the constitutionality of restrictions on labor's power to use picketing and nonpicketing publicity for this purpose. The Supreme Court addressed the constitutionality of restrictions on a union's ability to seek consumer support by picketing³ in *NLRB v. Retail Store Employees, Local 1001 (Safeco)*,⁴ holding that Congress, without violating the first amendment, may prohibit a union from picketing any secondary employer when the effect of the picketing predictably would result in substantial loss to or ruin of the secondary employer. Neither the courts nor the National Labor Relations Board (the Board), however, has addressed directly the constitutionality of restrictions on nonpicketing labor publicity.⁵ Rather, they consistently have avoided deciding the issue,⁶ as illustrated most recently by the Su-

2. A secondary boycott is the application of economic pressure by a union upon a person with whom the union *does not* have a dispute (the secondary or neutral party), in order to induce that person to apply pressure upon a person with whom the union *does* have a dispute (the primary or nonneutral party). Economic pressure in the form of appeals to consumers to withhold patronage is a secondary consumer boycott; union appeals directed to the secondary's employees seeking a work stoppage is a secondary employee boycott. This Article considers only secondary consumer boycotts.

3. Picketing typically refers to speech that includes patrolling, the carrying of placards, and face-to-face confrontation. See *Chicago Typographical Union No. 16*, 151 N.L.R.B. 1666, 1669 (1965).

4. 447 U.S. 607 (1980).

5. Unless otherwise noted, nonpicketing labor publicity, as used in this Article, refers to nonpicketing publicity directed against the *secondary*.

6. See, e.g., *Solien v. United Steelworkers*, 593 F.2d 82, 88 n.3 (8th Cir.), *cert. denied*,

preme Court's decision in *Edward J. DeBartolo Corp. v. NLRB*.⁷ The reluctance to address the first amendment questions that restrictions on nonpicketing labor publicity have raised leaves the state of the law in doubt and provides little guidance to unions that increasingly contemplate employing this most effective economic weapon. The courts' reluctance is particularly disturbing since Congress enacted the relevant statutory provision to avoid unconstitutional restrictions on union speech.⁸ Thus, a proper statutory analysis should have included consideration of the first amendment.

This Article attempts to provide the appropriate constitutional analysis of restrictions on nonpicketing labor publicity. Part II describes the relevant statute and illustrative cases, including the Supreme Court's *DeBartolo* decision, that have raised but not resolved the first amendment issues concerning nonpicketing labor publicity. The cases focus attention on two restrictions the courts have imposed on nonpicketing labor publicity—the “producer-distributor” and the “for the purpose of” requirements. Part III analyzes the protected status of the nonpicketing labor speech by comparing nonpicketing labor publicity with labor picketing and commercial speech—two areas that bear superficial similarity to nonpicketing labor publicity and that do not receive *full* first amendment protection. Demonstrating that the justifications for restrictions on labor picketing and commercial speech are not applicable to nonpicketing labor publicity and that nonpicketing labor publicity satisfies traditional first amendment values, this Article reasons that nonpicketing labor publicity should trigger *full* first amendment protection. Part IV identifies the possible govern-

444 U.S. 828 (1979); *Hospital and Serv. Employees Union Local 399 (Delta Air Lines)*, 263 N.L.R.B. 153, 111 L.R.R.M. (BNA) 1159 (Sept. 10, 1982); *Florida Gulf Coast Bldg. Trades Council (DeBartolo)*, 252 N.L.R.B. 702, 705 n.3 (1980), *enforced sub nom. Edward J. DeBartolo Corp. v. NLRB*, 662 F.2d 264 (4th Cir. 1981), *rev'd*, 103 S.Ct. 2926, 2933 (1983); *United Steelworkers (Pet, Inc.)*, 244 N.L.R.B. 96, 102 n.33 (1979), *rev'd sub nom. Pet, Inc. v. NLRB*, 641 F.2d 545 (8th Cir. 1981).

7. 103 S. Ct. 2926, 2933 (BNA) (Sept. 10, 1983).

8. See 105 CONG. REC. 17,720 (1959), *reprinted in* 2 NLRB, LEGISLATIVE HISTORY OF THE LABOR MANAGEMENT REPORTING AND DISCLOSURE ACT OF 1959, at 1388-89 (1959) (statement of Sen. Kennedy) [hereinafter cited as LEGISLATIVE HISTORY]; 105 CONG. REC. 17,898-99 (1959), *reprinted in* LEGISLATIVE HISTORY, at 1432 (statement of Sen. Kennedy); 105 CONG. REC. 6233 (1959), *reprinted in* LEGISLATIVE HISTORY, at 1038 (statement of Sen. Humphrey); 105 CONG. REC. 18,133-34 (1959), *reprinted in* LEGISLATIVE HISTORY, at 1720-21 (statement of Rep. Thompson); see also *NLRB v. Servette*, 377 U.S. 46, 55 (1964) (“The proviso was the outgrowth of a profound Senate concern that the unions’ freedom to appeal to the public for support of their case be adequately safeguarded.”).

ment interests in the "for the purpose of" and "producer-distributor" requirements and shows that because these interests do not justify the infringements on nonpicketing labor publicity the requirements are unconstitutional. Finally, the Article in Part V concludes by recommending a less restrictive alternative that allows more, not less, speech and that equally can serve the government's interest.

II. BACKGROUND

A. *The Statute*

Section 8(b)(4)(ii)(B)⁹ of the National Labor Relations Act (the Act) governs the limits of a union's ability to induce a secondary boycott¹⁰ by prohibiting a labor organization from threatening, coercing, or restraining "any person engaged in commerce or in an industry affecting commerce" for the purpose of forcing the person to "cease doing business with any other person. . . ." Congress, however, in response to the first amendment concerns of its members¹¹ provided an exception to the prohibition of secondary boycotts. The exception, known as the publicity proviso, exempts union appeals to the public when the publicity (1) does not include picketing; (2) is truthful; (3) is for the purpose of informing the public that the primary party to the dispute produces a product that is distributed by the secondary; and (4) does not result in a work stoppage.¹²

The construction and constitutionality of the third element of the publicity proviso has fostered great controversy and Board and

9. 29 U.S.C. § 158 (b)(4)(ii)(B) (1976). Congress enacted this provision in 1959 as part of the Landrum-Griffin amendments to the Act.

10. Section 8(b)(4)(ii)(B) does not restrict labor activity aimed directly against the primary employer or directly against others sufficiently related to the primary's business to lose their neutral or secondary status.

11. See sources cited *supra* note 8.

12. The complete text of the publicity proviso is:

Provided further, That for the purposes of this paragraph (4) only, nothing contained in such paragraph shall be construed to prohibit publicity, other than picketing, for the purpose of truthfully advising the public, including consumers and members of a labor organization, that a product or products are produced by an employer with whom the labor organization has a primary dispute and are distributed by another employer, as long as such publicity does not have an effect of inducing any individual employed by any person other than the primary employer in the course of his employment to refuse to pick up, deliver, or transport any goods, or not to perform any services, at the establishment of the employer engaged in such distribution;

National Labor Relations Act, § 8(b)(4), 29 U.S.C. § 158(b)(4) (1976).

court attention.¹³ Recent decisions have established that the third element contains a two-prong requirement that union activity must satisfy for publicity proviso exemption. First, a producer-distributor relationship must exist between the union's primary and secondary targets (the "producer-distributor" requirement). Second, the union's publicity must relate solely to the primary dispute or to the secondary's relationship to the primary dispute (the "for the purpose of" requirement). This interpretation of the publicity proviso, in turn, raises the unanswered constitutional question that is the focus of this Article—whether the first amendment permits Congress to restrict truthful nonpicketing labor publicity that does not result in an employee work stoppage solely on the ground that the labor activity does not satisfy the "producer-distributor" or the "for the purpose of" requirements.

B. The Cases

1. The "Producer-Distributor" Requirement

The constitutionality of the "producer-distributor" requirement arose in *United Steelworkers (Pet, Inc.)*,¹⁴ and in *Florida Gulf Coast Building Trades Council (DeBartolo)*.¹⁵ In *Pet, Inc.*, the United Steelworkers Union (the Union) had a labor dispute with Hussmann, a wholly owned subsidiary of Pet, Inc. When the Union struck and Hussmann hired permanent replacements¹⁶ the Union responded by advertising in newspapers and distributing handbills in the St. Louis area urging the public to boycott Pet products and *all* Pet subsidiaries.¹⁷ The Union conducted its publicity in an orderly and peaceful manner, did not picket,¹⁸ and truthfully disclosed that its primary dispute was with Hussmann, whose connection to Pet the Union accurately indicated.¹⁹ The publicity did not cause any work stoppage.²⁰

Pet filed a petition with the National Labor Relations Board charging that the Union's activity violated section 8(b)(4)(ii)(B) of

13. See, e.g., cases cited *supra* note 6.

14. 244 N.L.R.B. 96 (1979), *rev'd sub nom.* *Pet, Inc. v. NLRB*, 641 F.2d 545 (8th Cir. 1981).

15. 252 N.L.R.B. 702 (1980), *enforced sub nom.* *Edward J. DeBartolo Corp. v. NLRB*, 662 F.2d 264 (4th Cir. 1981), *rev'd*, 103 S. Ct. 2926 (1983).

16. *Solien v. United Steelworkers*, 593 F.2d 82, 83 (8th Cir.), *cert. denied*, 444 U.S. 828 (1979). *Solien* was a related injunction proceeding in federal court. See *infra* note 25.

17. 244 N.L.R.B. at 99.

18. *Id.*

19. *Id.* at 100.

20. *Id.* at 99.

the Act and that the publicity proviso did not protect the activity because the requisite "producer-distributor" relationship between Pet's subsidiaries and Hussmann did not exist.²¹ The Union countered that its conduct did not violate the Act, came within the language and purpose of the publicity proviso, and in any event, was conduct protected by the first amendment.²²

The Board ruled in favor of the Union, finding that the requisite "producer-distributor" relationship existed because "Hussmann applies capital, enterprise, and service to Pet and its other subsidiaries" in the form of profits and goodwill.²³ The Board therefore held that regardless of whether the Union's activity violated the prohibitions of section 8(b)(4)(ii)(B), the publicity proviso protected the activity.²⁴ Despite specific instructions from the court of appeals in a related proceeding,²⁵ the Board declined to address the first amendment issue.²⁶

21. *Id.*

22. *Id.* at 99-100.

23. *Id.* at 101.

24. *Id.* at 100 n.23. The Union asserted that its conduct had not violated the Act because Pet and its subsidiaries were not neutral secondaries and therefore did not fall within the Act's protection. The Union also asserted that it had not exercised any restraint or coercion within the meaning of the Act. One may wonder why the Board did not quickly dispose of the complaint by finding that Pet, Inc. was not a neutral secondary since it wholly owned Hussman. The neutrality issue, however, is not as simple as it may appear. The Board has held that divisions of the same company may be neutral; mere potential for control does not rob a parent of its neutrality. *See* Los Angeles Newspaper Guild Local 69, 185 N.L.R.B. 303, 304-05 (1970), *enforced*, 443 F.2d 1173 (9th Cir. 1971), *cert. denied*, 404 U.S. 1018 (1972); American Fed'n of Television and Radio Artists, Washington-Baltimore Local, 185 N.L.R.B. 593 (1970), *enforced*, 462 F.2d 887 (D.C. Cir. 1972). For a general discussion of the Board's neutrality doctrine, see Siegel, *Conglomerates, Subsidiaries, Divisions and the Secondary Boycott*, 9 GA. L. REV. 329 (1975); Comment, *The Single Employer Doctrine as Applied to Section 8(b)(4) of the National Labor Relations Act*, 28 CATN. U. L. REV. 555 (1979); Comment, *Unions, Conglomerates, and Secondary Activity Under the NLRA*, 129 U. PA. L. REV. 221 (1980).

25. After Pet, Inc. filed its complaint, the Board assumed jurisdiction of the case and sought an injunction against the Union in federal court pursuant to the National Labor Relations Act § 10(l), 29 U.S.C. § 160(l)(1976). The issues presented were identical to those later argued to the Board. The district court dismissed the Board's petition, finding that the Union's actions were outside the prohibition section of the Act. *See* Solien v. United Steelworkers, 449 F. Supp. 580 (E.D. Mo. 1978), *rev'd*, 593 F.2d 82 (8th Cir.), *cert. denied*, 444 U.S. 828 (1979). The court of appeals reversed, finding reasonable cause to believe that a violation of the Act had occurred. Without deciding whether an injunction would violate the Union's first amendment rights, the court of appeals directed the district court to grant an injunction. The court of appeals called the first amendment claim "not insubstantial" and directed the Board to consider the claim in the administrative proceeding. 593 F.2d at 88 n.3.

26. 244 N.L.R.B. at 102 n.33.

On appeal, the Eighth Circuit reversed.²⁷ According to the court, the connection between Pet's products and Hussmann was "highly attenuated," and the Board's finding that Hussmann was a producer of Pet's products because "Hussman's profits inure to the benefit of Pet" was "totally at odds with any normal interpretation of the word 'produce.'"²⁸ Thus, the court concluded that the publicity proviso did not protect the Union's conduct because Hussmann was not a producer of the products of Pet or Pet's subsidiaries and remanded the case to the Board to decide whether the union activity came within the prohibitions of section 8(b)(4)(ii)(B). The court recognized, but like the Board, declined to reach the first amendment issue.²⁹

Thus, in *Pet, Inc.*, Hussmann permanently replaced the Union members who engaged in primary activity and the court enjoined the Union's attempt through secondary activity to take its appeal to the public. At present, three courts and the National Labor Relations Board have heard the Union's case, yet the Union still has not received a decision on, or even a discussion of, its first amendment claims. The most recent case in which the constitutionality of the "producer-distributor" requirement arose is *DeBartolo*,³⁰ which concerned a primary labor dispute between the respondent Florida Gulf Coast Building Trades Council, AFL-CIO (the Union), and High Construction Company (High) over the payment of allegedly substandard wages and fringe benefits.³¹ The H.J. Wilson Company (Wilson) previously had hired High to build a retail store in a mall³² that the Edward J. DeBartolo Corp. (DeBartolo) managed. The Union, in support of its primary dispute with High,³³ circulated handbills urging the public not to shop at the mall or patron-

27. *Pet, Inc. v. NLRB*, 641 F.2d 545 (8th Cir. 1981).

28. *Id.* at 549.

29. *Id.* at 550.

30. 252 N.L.R.B. 702 (1980), *enforced sub nom. Edward J. DeBartolo Corp. v. NLRB*, 662 F.2d 264 (4th Cir. 1981), *rev'd*, 103 S. Ct. 2926 (1983).

31. *See id.* at 703.

32. *See id.*

33. The handbill appeal, *see infra* note 34, might suggest that the Union's primary dispute was with DeBartolo, not High, and accordingly that the Union's handbilling was primary and not secondary activity. The Union, however, possibly to avoid potential anti-trust liability, *see, e.g., Connell Constr. Co. v. Plumbers Local 100*, 421 U.S. 616 (1975), stipulated that its primary dispute was with High and not with DeBartolo. The charging party, rather than contest that stipulation, claimed that the handbill was misleading and untruthful by omitting High's name and focusing on DeBartolo. The Board rejected the claim, finding that the handbill had not "substantially departed from fact or intended to deceive." 252 N.L.R.B. at 704 n.2. The Board never questioned the stipulation that the primary dispute was with High.

ize the stores in the mall until "the mall's owner publicly promise[d] that all construction at the mall [would] be done using contractors who [paid] their employees fair wages and fringe benefits."³⁴ The Union conducted the handbilling peacefully and did not cause any employee work stoppage.³⁵

Before the Board, DeBartolo charged that the Union had violated section 8(b)(4)(ii)(B) of the Act. DeBartolo further asserted that since DeBartolo and the mall tenants had no relationship to High, no "producer-distributor" relationship existed and the publicity proviso of section 8(b)(4)(ii)(B) could not apply.³⁶ The Union responded that the publicity proviso exempted the Union's activity from the prohibitions of section 8(b)(4)(ii)(B) and that even if the proviso did not apply, the first amendment protected the Union's activity.³⁷ The Board ruled in favor of the Union and dismissed the complaint. Relying on its decision in *Pet, Inc.*,³⁸ the Board held

34. 252 N.L.R.B. at 703. The complete text of the handbill read:

PLEASE DON'T SHOP AT EAST LAKE SQUARE MALL PLEASE

The FLA. GULF COAST BUILDING TRADES COUNCIL, AFL-CIO is requesting that you do not shop at the stores in the East Lake Square Mall because of the Mall ownership's contribution to substandard wages.

The Wilson's Department Store under construction on these premises is being built by contractors who pay substandard wages and fringe benefits. In the past, the Mall's owner, The Edward J. DeBartolo Corporation, has supported labor and our local economy by insuring that the Mall and its stores be built by contractors who pay fair wages and fringe benefits. Now, however, and for no apparent reason, the Mall owners have taken a giant step backwards by permitting our standards to be torn down. The payment of substandard wages not only diminishes the working persons's [sic] ability to purchase with earned, rather than borrowed, dollars, but it also undercuts the wage standard of the entire community. Since low construction wages at this time of inflation means [sic] decreased purchasing power, do the owners of East Lake Mall intend to compensate for the decreased purchasing power of workers of the community by encouraging the stores in East Lake Mall to cut their prices and lower their profits?

CUT-RATE WAGES ARE NOT FAIR UNLESS MERCHANDISE PRICES ARE ALSO CUT-RATE.

We ask for your support in our protest against substandard wages. Please do not patronize the stores in the East Lake Square Mall until the Mall's owner publicly promises that all construction at the Mall will be done using contractors who pay their employees fair wages and fringe benefits.

IF YOU MUST ENTER THE MALL TO DO BUSINESS, please express to the store managers your concern over substandard wages and your support of our efforts.

We are appealing only to the public—the consumer. We are not seeking to induce any person to cease work or to refuse to make deliveries.

Id.

35. *See id.* at 703-04.

36. *See id.* at 704.

37. *See id.*

38. At the time of the Board's decision in *DeBartolo*, the Eighth Circuit had not reversed the Board's decision in *Pet, Inc.*

that "a union could lawfully handbill a neutral employer, urging a total consumer boycott of that employer, 'so long as the primary employer has at some stage produced, in the sense of applying capital, enterprise, or service, a product of the neutral employer.'"³⁹ The Board, employing rather impressive sophistry, found that High's relationship with DeBartolo and the mall tenants satisfied the "producer-distributor" requirement as a result of the "mutual obligations" and "sympiotic" relationship of the parties.⁴⁰ The contribution of High (the primary employer and producer) to the mall enterprise was application of "its labor to a product, i.e., the Wilson's store, from which DeBartolo and its tenants [the secondary employers and distributors] will derive substantial benefit."⁴¹ The Board again did not address the first amendment issue.

On appeal, a divided panel of the Fourth Circuit, according great respect to the expertise of the Board⁴² and rejecting the analysis of the Eighth Circuit in *Pet, Inc.*, affirmed. The court clarified the Board's reasoning by supplying the missing "distributed by" link of the producer-distributor chain and thus found that a producer-distributor relationship existed between the mall tenants and High.⁴³ The court's clarification, however, made the Board's construction of the statutory language no less tortuous. Despite urging by the Board's appellate counsel, by intervenor Florida Gulf Coast Building Trades Council, and by amicus Building and Construction Trades Department, the Fourth Circuit also declined to address the first amendment issue.⁴⁴

39. 252 N.L.R.B. at 705.

40. *Id.*

41. *Id.*

42. 662 F.2d at 269.

43. The court summarized the Board's finding:

The Board concluded that High is a producer of the Wilson's store at the mall because, by its employees' labor, it adds value to the store. High is also a producer of the mall enterprise itself because, by helping to build the Wilson's store, it has applied capital, enterprise, and service to the mall enterprise and has thus added value to that enterprise. Both the Wilson's store and the mall enterprise itself, therefore, are products of High's labor.

Although the Board does not elaborate, its reasoning suggests that it considered DeBartolo and the mall's tenants to be distributors of both of High's products. DeBartolo and the tenants help "distribute" the new Wilson's store and its inventory by attracting shoppers, helping to maintain common areas, and participating in joint advertising. They help "distribute" the mall enterprise and the goods sold through that enterprise simply by conducting business.

Id.

44. *See id.* at 271 n.6.

The Supreme Court reversed *DeBartolo*.⁴⁵ The Court had little trouble with the Board's finding that a person who, at some stage, adds capital, service, or enterprise to a product is a producer within the meaning of the Act. The Court, however, faulted the Board for failing to give even a "fairly possible" meaning to the "distributed by" language of the publicity proviso.⁴⁶ The Court reasoned that by focusing on the relationship between two secondary employers, *DeBartolo* and the cotenants, and testing that relationship by a standard so generous that virtually any employer could satisfy it, the Board effectively had eliminated the "distributed by" language from the publicity proviso.⁴⁷ The Court held that the secondary parties in *DeBartolo* did not satisfy the "distributed by" language because they neither had a business relationship with nor sold any products whose chain of production could reasonably be said to include the primary, High.⁴⁸ The Court also declined to address the first amendment issue, remanding the case for the Board to determine whether the Union conduct fell within the prohibitions of section 8(b)(4)(ii)(B).⁴⁹ Thus the consti-

45. *Edward J. DeBartolo Corp. v. NLRB*, 103 S. Ct. 2926 (1983).

46. *Id.* at 2933.

47. *Id.* at 2932. One can make an entirely reasonable argument that Congress, in drafting the "producer-distributor" requirement did not intend to restrict the scope of the publicity proviso with its "distributed by" language. Congress may have chosen such language as an example since the normal context of a secondary boycott is when the secondary party distributes the goods of the primary manufacturer. Thus, reading the "distributed by" language out of the statute arguably is completely appropriate. *See, e.g., Kamer, supra* note 1, at 581-84. Of course such a construction would obviate the need for any constitutional analysis and was impliedly rejected by the Court.

48. 103 S. Ct. at 2932-33. The Court could have held that the Union appeal did not satisfy the "producer-distributor" requirement *solely* because the secondary did not sell any product whose chain of production included the primary. The Court, by also suggesting that the existence of a business relationship can satisfy the "producer-distributor" requirement, left open the possibility that a parent-subsidiary relationship as in *Pet, Inc.* might fall within the proviso's coverage. Whether the Court intended the proviso to include within its coverage parent-subsidiary relationships is unclear.

49. *Id.* at 2933. The Court cited *NLRB v. Retail Store Employees Union Local 1001 (Safeco)*, 447 U.S. 607 (1980) and *NLRB v. Fruit Vegetable Packers Local 706 (Tree Fruits)*, 377 U.S. 58 (1964), suggesting that the appropriate question on remand was whether the Union advocated a boycott of a "substantial portion" of the secondaries' business. 103 S. Ct. at 2933. The *DeBartolo* handbill, however, clearly presented a Union appeal for a total boycott.

Two additional arguments that avoid the first amendment issue are available, however, to establish that the Union's conduct did not violate § 8(b)(4)(ii)(B). The first argument is that the Union's message did not "threaten, coerce, or restrain" the secondaries, *DeBartolo* and the cotenants, since the secondaries' relationship to the primary, High, was so remote that even advocacy of a total boycott of the secondaries was not "reasonably calculated" to induce customers not to patronize the secondaries at all. *See NLRB v. Retail Store Employees Local 1001*, 447 U.S. 607, 614 (1980). The second argument is that the object of the

tutionality of the "producer-distributor" requirement remains unaddressed.

2. The "For the Purpose of" Requirement

The question of the constitutionality of the publicity proviso's "for the purpose of" requirement arose in *Hospital and Service Employees Union Local 399 (Delta Air Lines)*.⁵⁰ In *Delta Air Lines*, the Union had a primary dispute with Statewide, a nonunion contractor that Delta Air Lines hired to perform janitorial services. In furtherance of its primary dispute with Statewide, the Union, in front of Delta's airport and downtown facilities distributed handbills urging passengers not to fly Delta.⁵¹ One of the handbills, in addition to publicizing the facts of the labor dispute, disclosed Delta's flight safety record implying that Delta was unsafe.⁵²

Delta filed a complaint with the Board charging that the Union violated section 8(b)(4)(ii)(B) and argued that the publicity proviso did not protect the Union conduct since the conduct was not solely "for the purpose of" advising the public of the primary labor dispute. A divided Board agreed with Delta and held that the handbill, by including information about Delta's flight safety record, was not "for the purpose of" advising the public of a labor dispute—only information related *solely* to the primary dispute could satisfy that requirement.⁵³ The Board declined to decide the constitutionality of this interpretation and took the position that

handbilling was not to force a secondary employer "to cease using, selling, handling, transporting, or otherwise dealing in the products of any other producer," 29 U.S.C. § 158 (b)(4)(ii)(B) (1976), because neither DeBartolo nor the cotenants had *any* business relationship with or power to control the labor policies of the primary, High, or any *substantial* business relationship with or control over the contracting decisions of the initial secondary, Wilson. The Board by adopting either argument would once again avoid confronting the first amendment issue. The conflict between the "producer-distributor" requirement and the first amendment would nevertheless remain, in light of the Supreme Court's holding in *DeBartolo* that Congress intended the "producer-distributor" requirement to restrict the scope of the publicity proviso. See 103 S. Ct. at 2932. The merits of the statutory arguments are beyond the scope of this Article.

50. 263 N.L.R.B. No. 153, 111 L.R.R.M. (BNA) 1159 (Sept. 10, 1982).

51. See *id.* at 1160.

52. The Union had distributed a total of four different handbills and republished the text of two in two Union newspapers. *Id.* The statutory issues concerning all but the one handbill discussed in the text are not relevant here.

53. See *id.* at 1161-62. The Board indicated that it would not prohibit publication of noncoercive information unrelated to the primary dispute. *Id.* at 1162 n.11. The publicity proviso, however, does not exempt such information. The prohibition section simply does not proscribe the information in the first instance.

the Board could "presume the constitutionality of the Act . . . absent binding court decisions to the contrary."⁵⁴

Thus, section 8(b)(4)(ii)(B), as interpreted by the Board and the courts, imposes at least two restrictions on truthful nonpicketing labor publicity that does not result in an employee work stoppage. The secondary must have a specified relationship with the primary—the "producer-distributor" requirement—and the publicity must relate solely to the primary dispute or the secondary's relationship to the primary dispute—the "for the purpose of" requirement. Curiously, however, the courts have avoided addressing the constitutionality of these restrictions on nonpicketing labor publicity despite the fact that unions increasingly are engaging in this activity and have presented the Board and courts with the opportunity to decide the issue. An appropriate constitutional analysis of the two restrictions on nonpicketing labor publicity as presented below necessitates a two-prong analysis—focusing first on the protected status of the speech infringed and second on the government's justification for the infringement. The analysis demonstrates that the "producer-distributor" and "for the purpose of" restrictions on nonpicketing labor publicity are unconstitutional.

III. THE FIRST AMENDMENT PROTECTION OF NONPICKETING LABOR PUBLICITY

The first inquiry in analyzing the constitutionality of the "producer-distributor" and "for the purpose of" requirements is the extent to which the first amendment protects nonpicketing labor publicity.⁵⁵ To make this determination, the courts may look to the

54. *Id.* at 1163 (footnote omitted).

55. The question whether nonpicketing labor publicity is entitled to *any* first amendment protection is not in issue, because it is clear that handbilling and other methods of peaceful publicity are forms of protected speech, *see, e.g.*, *United States v. Grace*, 103 S. Ct. 1702, 1706 (1983); *Carey v. Brown*, 447 U.S. 455, 460 (1980); *Organization for a Better Austin v. Keefe*, 402 U.S. 415, 419 (1971), and that these forms remain within the area of free expression when their subject is a labor dispute. *See, e.g.*, *Thornhill v. Alabama*, 310 U.S. 88, 102 (1940) ("In the circumstances of our times the dissemination of information concerning the facts of a labor dispute must be regarded as within that area of free discussion that is guaranteed by the Constitution."). Members of the Court, however, have begun to order the degree of protection afforded different classes of speech. *See, e.g.*, *FCC v. Pacific Found.*, 438 U.S. 726, 746-47 (1978) (indecent speech) (Stevens, J., concurring); *Young v. American Mini Theatres, Inc.*, 427 U.S. 50, 70-71 (1976) (nonobscene "adult movies") (plurality opinion); *Virginia State Bd. of Pharmacy v. Virginia Citizens Consumer Council, Inc.*, 425 U.S. 748 (1976) (commercial speech); *see also Connick v. Myers*, 103 S. Ct. 1684, 1689 (1983) (quoting *NAACP v. Claiborne Hardware Co.*, 102 S. Ct. 3409, 3426 (1982)) ("speech on public issues occupies the highest rung of the hierarchy of First Amendment values, and is entitled to special protection"); Emerson, *First Amendment Doctrine and the Burger*

law of the seemingly related activities of labor picketing and commercial speech. The courts may treat nonpicketing labor publicity as speech plus conduct similar to labor picketing, or as economically motivated similar to commercial speech, and therefore relegate the publicity to lesser first amendment protection. A review of the justifications for judicial restrictions on labor picketing and commercial speech and a comparison of such speech with nonpicketing labor publicity, however, illustrate the inappropriateness of subjecting nonpicketing labor publicity to similar restrictions. Nonpicketing labor publicity satisfies traditional first amendment values and is entitled to full first amendment protection.

A. Labor Picketing vs. Nonpicketing Labor Publicity

The Supreme Court has treated labor picketing with increasing disrespect. The broad pronouncements concerning the importance of labor speech contained in *Thornhill v. Alabama*⁵⁶ have eroded to the point that the Court has upheld restrictions on labor picketing with the most " cursory " first amendment analysis.⁵⁷ The Court has justified its cavalier treatment of labor picketing speech through the use of two related doctrines—the "speech plus"⁵⁸ and "unlawful objective"⁵⁹ theories. The two theories, however appropriate they may be in the picketing context, are not applicable to

Court, 68 CAL. L. REV. 422, 452-53 (1980); Schauer, *Categories and the First Amendment: A Play in Three Acts*, 34 VAND. L. REV. 265, 282-96 (1981); Stephan, *The First Amendment and Content Discrimination*, 68 VA. L. REV. 203, 212-14 (1982). The relevant inquiry therefore is the extent of protection nonpicketing labor publicity deserves, i.e., whether nonpicketing labor publicity is entitled to full first amendment protection. Whether the protection afforded different categories of speech should be ordered at all is beyond the scope of this Article.

56. 310 U.S. 88 (1940). See *supra* note 55.

57. See *NLRB v. Retail Store Employees Union Local 1001 (Safeco)*, 447 U.S. 607, 616 (1980) (Blackmun, J., concurring).

58. See, e.g., *Safeco*, 447 U.S. at 619 (Stevens, J., concurring); *American Radio Ass'n v. Mobile S.S. Ass'n*, 419 U.S. 215, 229-30 (1974); *NLRB v. Fruit & Vegetable Packers Local 760 (Tree Fruits)*, 377 U.S. 58, 77 (1964) (Black, J., concurring); *id.* at 93-94 (Harlan, J., dissenting); *International Bhd. of Teamsters Local 695 v. Vogt, Inc.*, 354 U.S. 284, 289 (1957); *Hughes v. Superior Court*, 339 U.S. 460, 464-65 (1950); *International Bhd. of Teamsters Local 309 v. Hanke*, 339 U.S. 470, 474 (1950) (plurality opinion); *Building Serv. Employees Int'l Union Local 262 v. Gazzam*, 339 U.S. 532, 537 (1950); *Giboney v. Empire Storage & Ice Co.*, 336 U.S. 490, 503 n.6 (1949) (quoting *Bakery & Pastry Drivers & Helpers Local 802 v. Wohl*, 315 U.S. 769, 776-77 (1942) (Douglas, J., concurring)).

59. See, e.g., *Safeco*, 447 U.S. at 616; *International Bhd. of Teamsters Local 695 v. Vogt, Inc.*, 354 U.S. at 293-94; *International Bhd. of Elec. Workers Local 501 v. NLRB*, 341 U.S. 694, 705 (1951); *Hughes v. Superior Court*, 339 U.S. at 465-66; *Building Serv. Employees Int'l Union Local 262 v. Gazzam*, 339 U.S. at 539-40; *Giboney v. Empire Storage & Ice Co.*, 336 U.S. at 501.

nonpicketing labor publicity.

1. "Speech Plus" Theory

The "speech plus" theory focuses on the physical component of labor picketing, arguing that picketing is "something more" than speech. The "speech plus" aspect of labor picketing thus justifies additional restrictions on labor speech—restrictions that are not applicable to other forms of noncommercial speech. "Speech plus" theorists contend that labor picketing constitutes, at least in part, conduct designed to coerce and to "induce action of one kind or another, quite irrespective of the nature of the ideas which are being disseminated."⁶⁰ According to these theorists, coercion may take the form of (1) threats of physical harm, (2) threats of economic sanctions, (3) unreasoned appeals to class or union loyalties, or (4) psychological pressure in the form of possible social embarrassment.⁶¹ Regardless of their merits as applied to labor picketing, these arguments cannot justify restrictions on nonpicketing labor publicity because this form of publicity lacks the conduct necessary to effectively *coerce* the *listener*—i.e. the consumer.⁶²

60. *Hughes v. Superior Court*, 339 U.S. at 465 (quoting *Bakery & Pastry Drivers & Helpers Local 802 v. Wohl*, 315 U.S. at 775-76 (Douglas, J., concurring)). The first amendment does not protect from regulation speech that merely coerces the listener. *See, e.g., NLRB v. Gissel Packing Co.*, 395 U.S. 575, 618 (1969); *NLRB v. Virginia Elec. & Power Co.*, 314 U.S. 469, 477 (1941); *see also Baker, Scope of the First Amendment Freedom of Speech*, 25 U.C.L.A. L. Rev. 964, 998-1002 (1978).

61. *See cases cited supra* note 58; *see also Gregory, Constitutional Limitations on the Regulation of Union and Employer Conduct*, 49 MICH. L. REV. 191, 198-210 (1950); Teller, *Picketing and Free Speech*, 56 HARV. L. REV. 180, 200-02 (1942).

62. Nonpicketing publicity may attempt to coerce *nonlisteners*, i.e. secondary employers. The union, by applying economic pressure to the secondary, hopes to coerce the secondary to pressure the person with whom the union has a primary dispute. The reviewing body considers this coercion when it determines if there is a statutory violation. A finding of coercion that satisfies the statutory requirements of § 8(b)(4)(ii)(B), however, cannot control the outcome of the first amendment question. Once the court reaches the first amendment issue, the analysis must focus on the possible coercion of the *listener* i.e., the consumer. *See Baker, supra* note 60, at 998; Jones, *Picketing and Coercion: A Jurisprudence of Epithets*, 39 VA. L. REV. 1023, 1047-51 (1953).

Under the first amendment, the government cannot regulate secondary consumer boycotts on the grounds of nonlistener coercion any more than it can regulate, for example, the solicitation of voters by special interest groups attempting to "coerce"—lobby—legislators or the appeals to consumers by civil rights organizations seeking a boycott of stores to "coerce" employers to cease engaging in racial discrimination. *See NAACP v. Claiborne Hardware Co.*, 102 S. Ct. 3409 (1982) (The first amendment protects a secondary boycott by civil rights organizers.). Unless the union coerces the listener, the government must base its regulation solely on the expected result of the uncoerced response of listeners to the union's message, precisely the kind of content-based restriction that is suspect under the first amendment. *See Central Hudson Gas & Elec. Corp. v. Public Serv. Comm'n*, 447 U.S. 557,

First, nonpicketing labor publicity does not threaten—either implicitly or explicitly—physical harm or violence. Although observers at one time may have associated labor picketing with regular outbursts of violence,⁶³ a similar association never has been present with nonpicketing labor publicity.⁶⁴ Nonpicketing publicity neither involves a procession of patrolling bodies that physically may intimidate consumers nor requires consumers seeking access to the secondary to contend with stationary crowds.⁶⁵ Many forms of nonpicketing publicity include no physical presence and, consequently, no face-to-face confrontation that even remotely could threaten violence.⁶⁶ Moreover, even if a slight chance of physical harm existed—for example, from a single overly enthusiastic handbiller—the appropriate response would be to regulate the offending conduct, not to regulate the speech that only remotely may produce the conduct.⁶⁷

Second, nonpicketing appeals to consumers to boycott a secondary employer do not threaten economic sanctions. A union—no matter how powerful—cannot punish or impose sanctions upon a consumer. The fear of economic sanctions is relevant only when a union directs its appeal to the employees of the secondary.⁶⁸ While

575 (1980) (Blackmun, J., concurring); *Linmark Assoc., Inc., v. Township of Willingboro*, 431 U.S. 85, 96-97 (1977); *Cohen v. California*, 403 U.S. 15, 23-24 (1971); *Tinker v. Des Moines Indep. Community School Dist.*, 393 U.S. 503, 508-09 (1969); see also, L. TRIBE, *AMERICAN CONSTITUTIONAL LAW* 580-81 (1978); Cox, *The Supreme Court 1979 Term—Foreword: Freedom of Expression in the Burger Court*, 94 HARV. L. REV. 1, 33-34 (1980); Ely, *Flag Desecration: A Case Study in the Roles of Categorization and Balancing in First Amendment Analysis*, 88 HARV. L. REV. 1482, 1496-1502 (1975); Stone, *Restrictions of Speech Because of Its Content: The Peculiar Case of Subject-Matter Restrictions*, 46 U. CHI. L. REV. 81 (1978).

63. See Jones, *supra* note 62, at 1030; Note, *Peaceful Labor Picketing and the First Amendment*, 82 COLUM. L. REV. 1469, 1490-91 (1982) [hereinafter cited as Note, *Peaceful Picketing*].

64. The association of labor picketing to sporadic violence does not suggest that the government properly may restrict labor picketing because of a perceived potential for physical intimidation. Not all picketing presents this concern, thus the government's restriction on all picketing may violate the first amendment. See *Cafeteria Employees Union Local 302 v. Angelos*, 320 U.S. 293, 295-96 (1943); Note, *Labor Picketing and Commercial Speech: Free Enterprise Values in the Doctrine of Free Speech*, 91 YALE L.J. 938, 952-53 (1982) [hereinafter cited as Note, *Labor Picketing*]; Note, *Peaceful Picketing*, *supra* note 63, at 1491; see also *infra* note 140.

65. When disruptive handbillers block public access, the publicity may constitute picketing. See *Service and Maintenance Employees Union Local 399*, 136 N.L.R.B. 431 (1962).

66. No one would argue, for example, that a newspaper advertisement physically abused its reader.

67. See *infra* note 140 and accompanying text.

68. See T. EMERSON, *THE SYSTEM OF FREEDOM OF EXPRESSION* 445-46 (1970); Cox,

consumer appeals inadvertently may reach the secondary's employees, nonpicketing publicity is less likely to induce employees to engage in a secondary boycott than is consumer picketing. Consumer picketing is more visible and conjures up the historical battle between management and labor and thus carries much greater symbolic force than does the distribution of handbills. Moreover, even if nonpicketing publicity directed at consumers occasionally induces a secondary employee boycott, the proper first amendment response is not to prohibit the broad class of nonpicketing publicity but rather to treat separately the occasional instances of resultant secondary employee boycotts.⁶⁹

Third, while nonpicketing publicity does appeal to consumer loyalties this without more does not constitute coercion. In many disputes the union must rely not only on its presentation of its grievances but also on the public's existing class biases. The union may seek to convince the public that the union's individual dispute is part of labor's greater cause. Such appeals, however, are no different than a political candidate's declaration of party affiliation to garner machine and party support despite possible differences on individual issues. In short, the government and courts must not limit first amendment protection to speech that seeks only to change the listener's mind but must extend protection to speech that encourages the listener to act on existing beliefs and loyalties.

Strikes, Picketing and the Constitution, 4 VAND. L. REV. 574, 594-95 (1951); Note, *The Invisible Hand and the Clenched Fist: Is There a Safe Way to Picket Under the First Amendment*, 26 HASTINGS L.J. 167, 177-78 (1974).

69. See *infra* note 140 and accompanying text. A prohibition of nonpicketing publicity that results in a work stoppage may have its own constitutional problems. See Note, *Labor Picketing*, *supra* note 64, at 945 n.40 ("the economic effect, actual or intended, of a peaceful consumer picket has no bearing on whether it is communication and should not be relevant in determining its level of protection under the First Amendment"). Some obvious differences exist, however, between nonpicketing publicity that does result in a work stoppage and publicity that does not produce a stoppage which may justify different constitutional treatment. When a work stoppage results, it is more likely that the union implicitly or explicitly threatened economic sanctions or promised benefits (e.g., support in any future work stoppage by the secondary employees), see *NLRB v. Gissel Packing Co.*, 395 U.S. 575 (1969), and it is more likely that the union had an objective that was illegal or contrary to a public policy unrelated to the expression (e.g., violation of a no-strike clause). See Note, *Labor Picketing*, *supra* note 64, at 938 n.2. Moreover, the court may presume that the union intended an employee work stoppage when a consumer appeal results in a work stoppage. See *American Radio Ass'n v. Mobile S.S. Ass'n, Inc.*, 419 U.S. 215, 231-32 (1974) (rational inferences arising from the effect of picketing may be relevant to a determination of its purpose). If the union did not have such an intent, the union could request that the secondary employees return to work. The secondary employees probably would honor the union's request, particularly if the secondary employees knew that the court otherwise would hold the union liable.

Finally, while nonpicketing publicity may produce social embarrassment⁷⁰ and cause some persons to shop elsewhere rather than suffer some minor discomfort, the Supreme Court has recognized that social embarrassment alone cannot justify restrictions on protected speech.⁷¹ The purposes of the first amendment would be undermined severely if speech became unprotected merely because it caused people to feel uncomfortable. The novel, the radical, the religious, and the political all may cause discomfort to one group or another.

Thus, nonpicketing publicity does not contain the same elements of conduct that the Court has considered sufficient to justify reduced protection of labor picketing.⁷² The isolated incidents of coercion that could arise do not justify restrictions on first amendment rights.

A second argument that "speech plus" adherents might advance is that nonpicketing publicity is part of an integrated course of union conduct that is designed to put pressure on the secondary. This argument would claim that the appeal to boycott transforms what otherwise would be protected speech into economic conduct subject to regulation.⁷³ This view has two problems. First, as the Supreme Court repeatedly has acknowledged, the first amendment "extends to more than abstract discussion, unrelated to action. . . . 'Free trade in ideas' means free trade in the opportunity to persuade to action, not merely to describe facts."⁷⁴ Second, distinguishing between information that discloses a secondary's participation in or connection with a labor dispute and appeals to consumers to boycott the secondary is not practicable. The former necessarily implies a request for the latter. Thus, this second justi-

70. To the extent one considers social embarrassment relevant, labor picketing carries a much greater potential for embarrassment than nonpicketing publicity. Picketing is more visible, carries greater symbolic force, requires face-to-face confrontation, and generally includes larger numbers of persons than does nonpicketing publicity.

71. See *NAACP v. Claiborne Hardware Co.*, 102 S. Ct. 3409, 3424 (1982).

72. The Supreme Court in dicta has recognized some of the common sense differences between labor picketing and nonpicketing publicity. See, e.g., *Safeco*, 447 U.S. at 618-19 (Stevens, J., concurring); *Babbitt v. United Farm Workers Nat'l Union*, 442 U.S. 289, 311 n.17 (1979); *Hughes v. Superior Court*, 339 U.S. 460, 464-65 (1950).

73. See *Giboney v. Empire Storage & Ice Co.*, 336 U.S. 490, 502 (1949) ("it has never been deemed an abridgment of freedom of speech or press to make a course of conduct illegal merely because the conduct was in part initiated, evidenced, or carried out by means of language . . .").

74. *NAACP v. Claiborne Hardware Co.*, 102 S. Ct. 3409, 3424 (1982) (quoting *Thomas v. Collins*, 323 U.S. 516, 537 (1945)); see also *In re Primus*, 436 U.S. 412, 432 (1978); *Organization for a Better Austin v. Keefe*, 402 U.S. 415, 419 (1971).

fication for characterizing nonpicketing publicity as "speech plus" also fails.

2. "Unlawful Objective" Theory

The "unlawful objective" theory, as originally formulated, denied protection to any picketing whose object or imminent effect was the violation of a valid state law.⁷⁵ The courts, however, soon expanded the doctrine to preclude from protection picketing whose purpose was to induce any action that either the legislature or the courts judged to be against state or congressional policy.⁷⁶

The "unlawful objective" doctrine is not applicable to nonpicketing labor publicity for two reasons. First, the doctrine evolved from the Court's finding that labor picketing was "speech plus," which permitted application of a less strict first amendment analysis.⁷⁷ While nonpicketing labor publicity may have the same ultimate objective as secondary labor picketing—that is, to force the secondary to apply pressure to the primary—nonpicketing labor publicity does not share the necessary "speech plus" elements upon which the "unlawful objective" theory depends.⁷⁸

Second, a union does not seek an unlawful objective when it appeals to consumers to boycott a secondary employer. The public violates no law if it heeds the union's advice and refuses to purchase from the secondary employer; the secondary employer does not act illegally if, as a result of the consumer pressure, it ceases to deal with the primary; and the primary does not violate the law if in response to such pressure it agrees to the union's demands concerning the terms and conditions of employment.⁷⁹ A

75. *See, e.g.*, *Giboney v. Empire Storage & Ice Co.*, 336 U.S. 490 (1949) (union picketed in an attempt to obtain agreement from wholesale ice distributor to refuse to sell to nonunion ice peddlers; the agreement would violate state antitrust statute so the picketing was not protected under the first amendment).

76. *See, e.g.*, *Hughes v. Superior Court*, 339 U.S. 460 (1950) (the Court affirmed an injunction prohibiting picketing for purpose of inducing employer to hire blacks in proportion to their population in the area because the picketing violated judicial policy against hiring on the basis of race); *see also* *International Bhd. of Teamsters Local 695 v. Vogt, Inc.*, 354 U.S. 284, 289-93 (1957) (tracing the evolution of the illegal objective doctrine).

77. *See, e.g.*, *American Radio Ass'n v. Mobile S.S. Ass'n, Inc.*, 419 U.S. 215, 229-30 (1974); *International Bhd. of Teamsters Local 695 v. Vogt, Inc.*, 354 U.S. 284, 289-93 (1957); *Hughes v. Superior Court*, 339 U.S. 460, 464-66 (1950); *Building Serv. Employees Int'l Union Local 262 v. Gazzam*, 339 U.S. 532, 537 (1950).

78. *See supra* notes 60-74 and accompanying text.

79. *Cox, supra* note 62, at 36-37; *see Note, Labor Picketing, supra* note 64, at 945-46. *But cf.* *Connell Constr. Co., Inc. v. Plumbers & Steamfitters Local 100*, 421 U.S. 616 (1975) (contractor's agreement with union to subcontract only to union subcontractors was subject to the Federal Antitrust Laws).

court, nevertheless, could find that the objective of nonpicketing publicity contravenes a valid law or policy by determining that section 8(b)(4)(ii)(B) itself prohibits secondary boycotts. This reasoning, however, clearly constitutes bootstrapping and, if valid, would permit states or Congress to impose by legislation any restriction on speech simply by declaring invalid the purpose of the speech.⁸⁰

Thus, neither the "unlawful objective" nor the "speech plus" theories, by which courts justify restrictions on labor picketing, are applicable to nonpicketing labor publicity. Accordingly, the labor picketing cases, although indicative of the Court's unsympathetic view toward union appeals, are not controlling in determining the constitutionality of restrictions on nonpicketing publicity.⁸¹

B. Commercial Speech v. Labor Speech

The Supreme Court has manifested its view that labor speech may be unworthy of full first amendment protection not only by a differential treatment of labor picketing, but also by making unflattering characterizations of labor speech in dicta in commercial

80. See Z. CHAFEE, *FREE SPEECH IN THE UNITED STATES* 14 (1941); T. EMERSON, *THE SYSTEM OF FREEDOM OF EXPRESSION* 446 (1970). But see *NLRB v. Retail Store Employees Union Local 1001 (Safeco)*, 447 U.S. 607, 616 (1980) (plurality opinion).

The Court in *Safeco* apparently applied this fallacious reasoning when it upheld the restrictions on labor picketing contained in § 8(b)(4)(ii)(B). The plurality opinion held that the secondary consumer picketing furthered an unlawful objective and therefore was subject to regulation. The *Safeco* opinion marked the greatest departure from the early illegal objective cases and was the first consumer picketing case in which the court found the illegal objective in the very statute that made the speech illegal.

The plurality opinion in *Safeco* garnered only four votes and has been sharply criticized. See Cox, *supra* note 62, at 36-39; Note, *Labor Picketing*, *supra* note 64, at 944-46; Note, *Peaceful Picketing*, *supra* note 63, at 1480-81; Comment, *Picketing at the Secondary: Retail Store and the Right to Publicize*, 30 *BUFFALO L. REV.* 405 (1981). No reason supports the extension of the discredited *Safeco* reasoning to nonpicketing publicity.

81. The picketing precedents also are inapplicable to nonpicketing publicity because, while restrictions on picketing leave other channels of communication open, see *NLRB v. Fruit Vegetable Packers Local 706 (Tree Fruits)*, 377 U.S. 58, 93 (1964) (Harlan, J., dissenting), a prohibition extending to nonpicketing publicity would cut off the entire message. Thus, the uniqueness of the medium or the manner of presentation could not justify reduced first amendment protection of nonpicketing publicity. See also *infra* note 125. The Supreme Court often has acknowledged the difference between regulations that leave alternate channels of communication open and those that do not. Compare *Young v. American Mini Theatres, Inc.*, 427 U.S. 50, 62 (1976) (time-place-manner regulation of adult movie theatres valid because those wanting to view adult movies still had access) with *Schad v. Borough of Mt. Ephraim*, 452 U.S. 61 (1981) (absolute ban on one form of adult entertainment unconstitutional). See also *International Longshoremen's Ass'n v. Allied Int'l, Inc.*, 456 U.S. 212, 227 (1982); *Friedman v. Rogers*, 440 U.S. 1, 15-16 (1979); *FCC v. Pacifica Found.*, 438 U.S. 726, 750 n.28 (1978); *Lloyd Corp., Ltd. v. Tanner*, 407 U.S. 551, 566-67 (1972).

speech cases.⁸² The first amendment values underlying labor speech, particularly nonpicketing publicity, however, indicate such speech is not analogous to commercial speech and is deserving of full first amendment protection.

1. The Commercial Speech Analogy

The Court in *Virginia State Board of Pharmacy v. Virginia Citizens Consumer Council, Inc.*⁸³ reversed the longstanding rule of *Valentine v. Christensen*⁸⁴ and extended limited first amendment protection to commercial speech. The Court based its decision in part upon the similarity between commercial and labor speech, finding "no satisfactory distinction between the two kinds of speech."⁸⁵ The Court explicitly refused to express any views on the constitutionality of restrictions on labor speech.⁸⁶ The Court's inability to distinguish labor and commercial speech coupled with its characterization of some labor speech as "of an entirely private and economic character,"⁸⁷ however, raises the question whether labor speech should enjoy any greater first amendment protection than commercial speech.

The Court's decision four years after *Virginia State Board of Pharmacy* in *Central Hudson Gas & Electric Corp. v. Public Service Commission*⁸⁸ raised further doubts about the protected status of labor speech. The Court in *Central Hudson* struck down a regulation that prohibited electric utilities from advertising to promote the purchase of utility services, characterizing commercial speech as "expression related solely to the economic interests of the speaker and its audience."⁸⁹ As the concurrence warned, this broad definition—unlike the Court's usual definition of commercial speech as speech proposing no more than a commercial transaction⁹⁰—arguably includes labor speech.⁹¹ Although labor speech is

82. See *infra* notes 83-94 and accompanying text.

83. 425 U.S. 748 (1976).

84. 316 U.S. 52 (1942) (the first amendment does not protect purely commercial advertising).

85. 425 U.S. at 763.

86. *Id.* at 763 n.17.

87. *Id.* Several later cases draw upon this characterization of labor speech. See, e.g., *NAACP v. Claiborne Hardware Co.*, 102 S. Ct. 3409, 3425-26 (1982); *Bates v. State Bar*, 433 U.S. 350, 364 (1977); see also *Carey v. Brown*, 447 U.S. 455, 466-67 (1980) (rejecting the claim that labor picketing deserves more first amendment protection than other picketing and suggesting the contrary was true).

88. 447 U.S. 557 (1980).

89. *Id.* at 561.

90. See, e.g., *Bolger v. Youngs Drug Prods. Corp.*, 51 U.S.L.W. 4961, 4963 (June 24,

not related solely to the economic interests of its audience, the union's economic interests certainly are a motivating factor. The *Central Hudson* definition of commercial speech uncomfortably parallels the Court's characterization of some labor speech as "of an entirely private and economic character."⁹² Thus, *Central Hudson* and *Virginia State Board of Pharmacy* provide precedent for extending to any form of labor speech no greater first amendment protection—and perhaps even less—than that afforded commercial speech.⁹³

2. First Amendment Values

An analysis of how commercial⁹⁴ and labor speech satisfy the values underlying the first amendment—(a) individual self-fulfill-

1983); *Ohralik v. Ohio State Bar Ass'n*, 436 U.S. 447, 456 (1978); *Virginia State Bd. of Pharmacy v. Virginia Citizens Consumer Council, Inc.*, 425 U.S. 748, 762 (1976); *Pittsburgh Press Co. v. Pittsburgh Comm'n on Human Relations*, 413 U.S. 376, 385 (1973).

The Court in *Central Hudson* restated the traditional definition of commercial speech later in its opinion, 447 U.S. at 562, but whether the Court believes the definitions are equivalent or prefers one definition is unclear.

91. See *Central Hudson*, 447 U.S. at 579-80 (Stevens, and Brennan, JJ., concurring). The concurring justices opined that the majority's definition of commercial speech was too broad because the definition "arguably could include labor speech." Labor speech, according to the concurrence, should not receive the lesser commercial standard of first amendment protection, but the maximum protection available under the first amendment. *Id.*

92. More ominous to Professor Cox was the fact that the Court's decision in *Central Hudson* seemed to afford even greater respect to commercial speech than to labor speech. See Cox, *supra* note 62, at 36-39; see also Note, *Peaceful Picketing*, *supra* note 64, at 1487. In *Central Hudson*, as in *Safeco*, which the Court decided the same day, the government proscribed a form of expression appealing to readers or observers because the government believed that the message would persuade its audience to behave in an undesirable way. Yet the Court, without providing an explanation for the different treatment, invalidated the nonlabor regulation in *Central Hudson* and upheld the labor regulation in *Safeco*. See Cox, *supra* note 62, at 37.

93. Some commentators have suggested that the Court already may have relegated labor speech to no greater than commercial speech status. See, e.g., Cox, *supra* note 62, at 39; Note, *Labor Picketing*, *supra* note 64, at 959-60.

94. This discussion considers commercial speech to be that speech which does no more than propose a commercial transaction. See *supra* note 90 and accompanying text; see also Baker, *Commercial Speech: A Problem in the Theory of Freedom*, 62 IOWA L. REV. 1, 35 (1976). This Article does not consider the categorization problems here. See *Bolger v. Youngs Drug Prod. Corp.*, 51 U.S.L.W. 4961, 4962-63 (June 24, 1983); *Virginia State Bd. of Pharmacy v. Virginia Citizens Consumer Council, Inc.*, 425 U.S. 748, 764-65 (1976). But see Jackson & Jeffries, *Commercial Speech: Economic Due Process and the First Amendment*, 65 VA. L. REV. 1, 21-24 (1979) (categorization has not been and should not be a problem).

This Article does not examine the values that commercial speech satisfies to analyze the merits of the Court's decision in *Virginia State Bd. of Pharmacy*, because that analysis is beyond the scope of this Article. See generally *id.*; Baker, *supra*; Farber, *Commercial Speech and First Amendment Theory*, 74 Nw. U. L. REV. 372 (1979); Redish, *The First Amendment in the Marketplace: Commercial Speech and the Values of Free Expression*,

ment, (b) the advancement of knowledge and discovery of truth, (c) participation in decisionmaking by all members of society, and (d) the maintenance of the proper balance between stability and change⁹⁵—compels the conclusion that labor speech, particularly nonpicketing labor speech, deserves greater protection than commercial speech.⁹⁶

(a) *Individual Self-Fulfillment*

Professor Emerson argues that “every man . . . has the right to form his own beliefs and opinions” and “to express these beliefs and opinions. . . . For expression is an integral part of the development of ideas, of mental exploration and of the affirmation of self.”⁹⁷ Thus, this first amendment value fosters individual self-re-

39 GEO. WASH. L. REV. 429 (1971); Note, *Constitutional Protection of Commercial Speech*, 82 COLUM. L. REV. 720 (1982). Rather, the Article addresses the values satisfied by commercial speech to provide a framework for analysis and comparison of the value of the labor speech upon which the “producer-distributor” and “for the purpose of” requirements infringe.

95. Professor Emerson first described the values the accompanying text analyzes and other noted commentators generally have accepted these values. T. EMERSON, *THE SYSTEM OF FREE EXPRESSION* 6-7 (1970); see also G. GUNTHER, *CONSTITUTIONAL LAW* 1044 n.3 (9th ed. 1975); Baker, *supra* note 60, at 990-91. Some scholars emphasize one value over or to the exclusion of another. See, e.g., A. MEIKLEJOHN, *FREE SPEECH AND ITS RELATION TO SELF GOVERNMENT* (1948) (emphasizing the value of political participation, but defining it sufficiently broadly to encompass other values); Bork, *Neutral Principles and Some First Amendment Problems*, 47 IND. L.J. 1 (1971) (emphasizing exclusively the decisionmaking value, arguing that the first amendment extends only to explicitly political speech); BeVier, *The First Amendment and Political Speech: An Inquiry into the Substance and Limits of Principle*, 30 STAN. L. REV. 299 (1978) (emphasizing only political speech, but accepting extension of protection to some additional speech for pragmatic and institutional reasons); Jackson & Jeffries, *supra* note 94 (emphasizing the value of self-government, although also acknowledging a possible interest in individual self-fulfillment). The Emerson values, however, do not exclude any important value. Moreover, while some commentators criticize the Emerson choices as overinclusive, the overinclusiveness may be more apparent than real. Some values derive from others, see, e.g., Baker, *supra* note 60, at 991, and Emerson himself recognized that the four values have meaning only as an integrated set. Emerson, *First Amendment Doctrine and the Burger Court*, 68 CALIF. L. REV. 422, 423 (1980); accord Note, *supra* note 94, at 731 n.80.

96. Indeed, since the courts have subjected restrictions on commercial speech to an increasingly rigorous first amendment analysis, see *infra* notes 155 and 156, the courts should provide full first amendment protection to nonpicketing labor publicity.

97. T. EMERSON, *TOWARD A GENERAL THEORY OF THE FIRST AMENDMENT* 5 (1966); see also Baker, *supra* note 60, at 995 (the value of self-expression is the most important value underlying the first amendment).

The Supreme Court also has recognized that the first amendment protects the individual's interest in self-expression. See, e.g., *Consolidated Edison Co. v. Public Serv. Comm'n*, 447 U.S. 530, 534 n.2 (1980); *First Nat'l Bank v. Bellotti*, 435 U.S. 765, 777 n.12 (1978); *id.* at 804 (White, J., dissenting); *Cohen v. California*, 403 U.S. 15, 24-26 (1971).

alization and self-determination.

Providing first amendment protection to commercial speech will not promote these goals.⁹⁸ A seller hawking his wares is not engaged in an expression of self. The advertisements giving the price of cigarettes or toilet tissue are not instrumental to the development of anyone's personality and do not give the speaker a feeling of self-fulfillment.⁹⁹

Labor speech, by contrast, is an expression of self—of the individual's frustrations and aspirations. The individual pleading "Do Not Patronize"—like the person wearing the jacket bearing the words "Fuck the Draft"¹⁰⁰ or a war protestor chanting "Stop This War Now"¹⁰¹—is expressing his emotions and views on a subject of special importance to him. Providing first amendment protection to such speech "comport[s] with the premise of individual dignity"¹⁰² and channels emotions away from more socially undesirable behavior—e.g. violence—that may undermine the individ-

98. See Baker, *supra* note 94, at 9-25; Jackson & Jeffries, *supra* note 94, at 14-15. But see Note, *supra* note 94, at 744. The form of an advertisement, in some limited cases, may be an expression of self, e.g., a humorously presented commercial or nouveau poster. The protection given the commercial message, however, may be separable from that given the form of communication. Even if not separable, uniquely expressive advertisements are rare and their existence only suggests the problems of categorization (i.e. whether particular speech "does no more than propose a commercial transaction.") While problems of categorization may justify coverage of *all* commercial speech, see Farber, *supra* note 94, at 384, the Court in *Virginia State Bd. of Pharmacy* did not rely on this reasoning, finding instead that speech which does no more than propose a commercial transaction deserves protection for its own sake. 425 U.S. 748; see Jackson & Jeffries, *supra* note 94, at 18-20.

99. Without a self-fulfillment interest, an individual's commercial speech, unlike labor speech, does not deserve first amendment protection unless the speech satisfies other first amendment values *and* the speech of *other* speakers could not satisfy those values. This follows because only the self-fulfillment value furthers the *speaker's* first amendment interests. See Note, *supra* note 94, at 738.

100. See *Cohen v. California*, 403 U.S. 15 (1971).

101. Baker, *supra* note 60, at 994. Professor Baker argues that:

[t]o engage voluntarily in a speech act is to engage in self-definition or expression. A Vietnam war protestor may explain that when she chants "Stop This War Now" at a demonstration, she does so without any expectation that her speech will affect the continuance of war or even that it will communicate anything to people in power; rather, she participates and chants in order to *define* herself publicly in opposition to the war. This war protestor provides a dramatic illustration of the importance of this self-expressive use of speech, independent of any effective communication to others, for self-fulfillment or self-realization. Generally, any individually chosen, meaningful conduct, whether public or private, expresses and further defines the actor's nature and contributes to the actor's self-realization.

Id. (emphasis in original).

102. *Cohen v. California*, 403 U.S. 15, 24 (1971); see Z. CHAFEE, *supra* note 80, at 33 ("There is an individual interest, the need of many men to express their opinions on matters vital to them if life is to be worth living . . ."); see also Cox, *supra* note 62, at 1.

ual's self-actualization as well as society's stability.¹⁰³

Labor speech also is connected intimately with the individual's associational rights. The union member acts as part of a group—the union—and learns to interact with other members of that group. He achieves his goals most effectively when he coordinates his efforts with those of his co-workers in an attempt to persuade his employer and often the public. Thus, labor speech not only leads to the development of the individual *qua* individual, but also contributes to the individual's growth as a member of the group and of the larger society.¹⁰⁴

(b) *Advancement of Knowledge and Discovery of Truth*

The second value is essentially a formulation of the classic "marketplace of ideas" metaphor that litigants and the courts frequently invoke to protect challenged speech.¹⁰⁵ As Justice Holmes explained, "the ultimate good desired is better reached by free trade in ideas—that the best test of truth is the power of the thought to get itself accepted in the competition of the market"¹⁰⁶

Commercial speech only remotely promotes this value. Commercial speech does not convey ideas, is not a response to the ideas of another, is not susceptible to opposing argument, and is not conducive to dialogue that will ferret out facts to yield some ultimate

103. See *infra* note 125 and accompanying text.

104. See Baker, *supra* note 60, at 1033 ("the contribution of associations or assemblies as the embodiment of people's power to the first amendment values of self-fulfillment or participation in change should be obvious").

A trade association promoting its product also may include some element of association rights. Since the individual's message lacks an expression of self, however, the combined efforts of the group conveying the commercial message make a reduced contribution to the individual's growth. Moreover, unlike labor speech, a trade association's commercial speech, although formulated jointly, is rarely communicated to the public through group activity. Perhaps most important, trade association speech is not representative of commercial speech. Whereas *all* labor speech contains an element of association, only a very small portion of commercial speech has that element of joint communication.

105. See, e.g., Board of Educ., Island Trees Union Free School Dist. No. 26 v. Pico, 102 S. Ct. 2799, 2808 (1982); Gertz v. Robert Welch, Inc., 418 U.S. 323, 339-40 (1974); New York Times Co. v. Sullivan, 376 U.S. 254, 272 n.13, 279 n.19 (1964); see also L. TRIBE, *supra* note 62, at 576-77. For a criticism of the market place of ideas model, see Baker, *supra* note 60, at 974-81.

106. Abrams v. United States, 250 U.S. 616, 630 (1919) (Holmes, J., dissenting); accord Consolidated Edison Co. v. Public Serv. Comm'n, 447 U.S. 530, 534 (1980); see also 2 SELECTED LITERARY AND POLITICAL PAPERS AND ADDRESSES OF WOODROW WILSON 333 (1926), quoted in Thomas v. Collins, 323 U.S. 516, 546 n.1 (1945) (Jackson, J., concurring); Stone, *supra* note 62, at 101-03.

truth. Commercial speech presents facts and little more.¹⁰⁷

Some labor speech, like commercial speech, may suffer from a failure to express ideas. The union member carrying a picket sign saying no more than "on strike" provides a new fact, but may not contribute to the marketplace of ideas any more than a classified advertisement that offers minimum wage.¹⁰⁸ Most labor speech, however, provides greater impact on the formation of opinions and beliefs than an "on strike" sign. In particular, nonpicketing publicity typically will provide information and express opinions about the labor dispute.¹⁰⁹ In contrast to the use of a picket sign that the public may have difficulty reading if all the details of the labor dispute are presented, a handbill or advertisement is circulated precisely to allow for a more thorough discussion of the facts.¹¹⁰

Finally, since labor unions most often communicate verbally, by handbill or by picketing, a representative sharing the expressed opinion generally is available to respond to questions. The representative can indicate the causes of the dispute, the reasons for the action, and the goals the union seeks. When such information is

107. The claim that the facts conveyed in commercial speech are essential to the proper working of the economic marketplace, see *Virginia State Bd. of Pharmacy v. Virginia Citizens Consumer Council, Inc.*, 425 U.S. 748, 763, 765 (1976), is not sufficient to satisfy the second value in the first amendment analysis. The marketplace depends on much more than just commercial speech. For example, the absence of automotive transportation would cripple delivery systems. Yet the first amendment does not prevent governmental restrictions on and regulation of motor vehicles. Quite simply, arguments based upon the needs of the marketplace merely resurrect economic due process "in the ill-fitting garb of the first amendment." Jackson & Jeffries, *supra* note 94, at 30; see also Note, *Labor Picketing*, *supra* note 64, at 957-58. Furthermore, the argument that these commercial facts help to form opinions about important public issues also falls short of satisfying the second value in the first amendment analysis. See *infra* note 114.

108. Both the picketer and the classified advertisement may help to form the opinion that labor is oppressed. Neither, however, directly makes that argument; instead, the public must arrive at its own conclusion based on the interaction of the individual's life experience with the new fact.

109. For an example, see the *DeBartolo* handbill, *supra* note 34.

110. The difference in the detail with which labor picketing and nonpicketing publicity describe a labor dispute arguably could provide a justification for treating the two forms of publicity differently. That is, one could attempt to treat labor speech that presents views or opinions (typically nonpicketing publicity) differently from so-called "signal" labor speech (presumably picketing), which does not add to the marketplace of ideas, but merely manifests that a labor dispute exists. This approach, however, presents several problems. First, determining what constitutes enough information to contribute to the marketplace of ideas would be a difficult task. Second, when union activity provides few details the message is less persuasive, proportionately reducing the government interest in suppressing the communication. Last, a detail-based approach does not accord sufficient weight to the contribution that a picket sign makes to the value of self-expression. Cf. Baker, *supra* note 60, at 995 (self-expressive speech more fully promotes fundamental first amendment values than does speech that communicates propositions and attitudes).

provided, the employer who is the object of the dispute has something to which he may respond. He may post signs in his store window, speak to customers personally, or counteradvertise. Labor publicity therefore permits the dialogue that the marketplace of ideas model envisions.

(c) *Participation in Decisionmaking*

The third main function of free expression is to provide for participation in decisionmaking through a process of open discussion that is available to all members of the community.¹¹¹ This function has particular significance in the political realm in which the state makes most of the immediate decisions on its survival, welfare, and progress, and in which the state has a special incentive and often more effective power to suppress free expression.¹¹² Thus, to many, the most important societal interest that the first amendment advances is this third value, self-government.¹¹³

Commercial speech does not provide information directly relevant to the individual's participation in the political process.¹¹⁴

111. See T. EMERSON, *supra* note 97, at 8-11.

112. See *id.* at 9.

113. See, e.g., *Connick v. Myers*, 51 U.S.L.W. 4436, 4437 (Apr. 20, 1983); *NAACP v. Claiborne Hardware Co.*, 102 S. Ct. 3409, 3426 (1982); *Carey v. Brown*, 447 U.S. 455, 467 (1980); *First Nat'l Bank v. Bellotti*, 435 U.S. 765, 776-77 (1978); *Buckley v. Valeo*, 424 U.S. 1, 14-15, 39 (1976). Indeed, some commentators argue that the political process principle should be the *only* premise employed when construing the limits of the first amendment. See also *supra* note 95.

114. "The typical newspaper advertisement or television commercial makes no comment on governmental personnel or policy. It does not marshal information relevant to political action, nor does it focus public attention on questions of political significance." Jackson & Jeffries, *supra* note 94, at 15; accord *BeVier*, *supra* note 95, at 353-55.

Commercial speech arguably transmits information that contributes to the public's formation of political opinions and thereby indirectly contributes to the individual's participation in the political process. See Note, *supra* note 94, at 745; see also *Virginia State Bd. of Pharmacy v. Virginia Citizens Consumer Council, Inc.*, 425 U.S. 748, 765 (1976); *id.* at 780 n.8 (Stewart, J., concurring). For example, Justice Stewart argued in *Virginia State Bd. of Pharmacy* that drug price information might affect a person's views concerning price control issues, government subsidy proposals, special health care, consumer protection, or tax legislation. *Id.* at 780 n.8. The ban on drug prices challenged in *Virginia State Bd. of Pharmacy*, however, also may help the public form opinions on those same issues and has the identical political significance. See Jackson & Jeffries, *supra* note 94, at 15 n.53. Similarly, an individual viewing a street fight may conclude that the government should do more to reduce crime, should make sentencing or parole laws stricter, should reform prisons, or should improve police service. No one, however, would seriously argue that the ban on drug price information in *Virginia State Bd. of Pharmacy* or the street fight were protected speech. In short, all of life's experiences may affect a person's views concerning some political issues. Accordingly, a more direct link to the political process than that Justice Stewart postulates is necessary to trigger first amendment protection for commercial speech.

Commercial speech also does not further the individual's ability to participate in the political process. Unlike speech in such fields as philosophy, logic, or literature, commercial speech neither expands the limits of the mind nor improves upon the individual's overall decisionmaking processes.¹¹⁵

Labor speech also generally makes no comment on governmental personnel or policy and, like commercial speech, often is economically motivated. Labor speech, however, unlike commercial speech, does concern political and social choices. The union appeals to the public for support in its particular dispute and in its larger struggle.¹¹⁶ The success of the appeal depends upon the public's decision to forego what it had perceived as in its economic self-interest and to support the political and social values the union represents.¹¹⁷ The meaningfulness of these political and social choices ultimately rests upon the union's ability to communicate its message. A public deprived of information concerning the facts of a labor dispute is deprived of the ability to participate knowledgeably in the decisionmaking process.¹¹⁸ Furthermore, re-

115. Commercial speech, of course, does enable the consumer to engage more rationally in economic decisionmaking. See *Virginia State Bd. of Pharmacy*, 425 U.S. at 765; see also Redish, *supra* note 94, at 443-45. The first amendment, however, was not designed to further strictly economic interests. See *supra* note 107.

116. See *Thornhill v. Alabama*, 310 U.S. 88, 103 (1940); Note, *Labor Picketing*, *supra* note 64, at 955.

The Court in *Thornhill* reasoned that free discussion concerning labor disputes is indispensable to the proper use of governmental processes to shape the destiny of modern industrial society. 310 U.S. at 103. This reasoning arguably suffers from the same defect implicit in the *Virginia State Bd. of Pharmacy* opinion—the lack of a direct link between the speech and the political decision. See *supra* note 114. If the only political decision at issue was a future electoral decision on the roles of labor and management in industrial relations, the argument might have merit. A union's message and its act of transmitting it through group association, however, unlike the commercial advertisement, have more immediate political and social repercussions. See *infra* note 117 and accompanying text.

117. Moreover, the union members themselves engage in political decisionmaking by the very practice of bonding together with their fellow workers to voice their concerns on public issues. See *Citizens Against Rent Control v. Berkeley*, 454 U.S. 290, 294 (1981).

While the political and social choices the union and the public make do not immediately register in the voting booth, only the most restrictive definition of political speech would require electoral significance. See Bork, *supra* note 95, at 29-35. Judge Bork would limit first amendment protection to political speech that deals "explicitly, specifically and directly with politics and government." *Id.* at 26. Although this interpretation of political speech would exclude labor speech from first amendment protection, it also would exclude scientific and literary expression as well. *Id.* at 28. The Supreme Court never has endorsed the Bork definition of political speech. See *Connick v. Myers*, 51 U.S.L.W. 4436, 4438 (Apr. 20, 1983); *Abood v. Detroit Bd. of Educ.*, 431 U.S. 209, 231-32 (1977).

118. See *Carpenters and Joiners Union Local 213 v. Ritter's Cafe*, 315 U.S. 722 (1942) (Black, J., dissenting).

Disputes between one or two unions and one contractor over the merits and justice of

pression of labor's grievances also might affect the political processes by undermining the confidence and faith of the public in the democratic system. Justice Brandeis recognized that "repression breeds hate; that hate menaces stable government; that the path of safety lies in the opportunity to discuss freely supposed grievances and proposed remedies."¹¹⁹

(d) *Balance Between Stability and Change*

The final value the first amendment promotes is a "method of achieving a more adaptable and at the same time more stable community, of maintaining the precarious balance between healthy cleavage and necessary consensus."¹²⁰ Expression therefore allows society to avoid the natural tendency toward rigidity on the one hand and the violent response such rigidity may produce on the other.¹²¹

Commercial speech does not satisfy this value. Commercial speech neither calls for nor addresses social, political, or even economic change nor allows dissidents to "let off steam,"¹²² nor constitutes a response to such minority views. Rather, commercial speech is merely a means by which the speaker seeks to benefit from the existing economic structure.¹²³ Indeed, observers have ex-

union as opposed to non-union systems of employment are but a part of the nationwide controversy over the subject. I can see no reason why members of the public should be deprived of any opportunity to get information which might enable them to use their influence to tip the scales in favor of the side they think is right.

Id. at 730.

119. *Whitney v. California*, 274 U.S. 357, 375 (1927) (Brandeis, J., concurring). The need to prevent repression closely parallels the due process requirements of our judicial system. Both further the individual's self-dignity and the operation of the underlying system, whether judicial or political, by maintaining confidence in its equity and fairness. The Brandeis argument does not apply to commercial speech because this form of speech lacks a self-expression value and thus repression would not breed hatred.

120. T. EMERSON, *supra* note 97, at 11; *cf.* *Connick v. Myers*, 51 U.S.L.W. 4436, 4437 (Apr. 20, 1983) ("The First Amendment 'was fashioned to assure unfettered interchange of ideas for the bringing about of political and social changes desired by the people.'"); *Stromberg v. California*, 283 U.S. 359, 369 (1931) ("The maintenance of the opportunity for free political discussion to the end that government may be responsive to the will of the people and that changes may be obtained by lawful means . . . is a fundamental principle of our constitutional system.").

121. *See* T. EMERSON, *supra* note 97, at 11-12.

122. *Id.* at 12.

123. Certain businesses will benefit over others through the process of consumers registering their preferences in the marketplace. This process may result in a gradual change in our economic institutions. The consequent change, however, is a mere shifting of interests within a given segment of the existing economic structure, resulting in neither a fundamental change in the distribution of wealth between existing classes, nor a restructuring of the

pressed concern that commercial speech may help to rigidify the prevailing distribution of wealth and economic structure.¹²⁴

By contrast, advocacy of political and social change is at the core of labor speech. The essence of the labor movement is the realignment of individual wealth from one segment of society to another. The union member appeals to the public to support his cause in a particular dispute. The union member's plea, however, is also directed to the public's sentiments towards labor's larger struggle, which the individual dispute is, in fact, the first step toward resolving. Silencing labor speech not only would impede social change, but also would risk channeling union frustrations into more destructive avenues of release.¹²⁵ The union will accept more readily the common judgment when given an opportunity to persuade consumers to adopt the union's cause. If, however, the law permits complete repression of the union's voice, the union may resort to force, rather than consumers' marketplace votes, to persuade the secondary employer. Thus, allowing labor speech strikes the delicate balance between stability and change.

In summary, neither the labor picketing nor the commercial speech precedents can justify reduced first amendment protection of nonpicketing labor publicity. Nonpicketing labor publicity satisfies all the values underlying the first amendment and does so more fully than commercial speech. Indeed, since the Supreme Court recently has subjected restrictions on commercial speech to increasingly thorough first amendment analysis, courts should apply an even more rigorous analysis to restrictions on nonpicketing labor publicity. Under an appropriate analysis, the "for the purpose of" and "producer-distributor" requirements clearly infringe fully protected nonpicketing labor speech. The finding that the "producer-distributor" and "for the purpose of" requirements are an infringement on fully protected speech, however, does not end the relevant inquiry.

larger economic system. The fourth value contemplates the balance between stability and change in these larger institutions. See *id.* at 11-14.

124. See Baker, *supra* note 60, at 979-80 and n.59.

125. See T. EMERSON, *supra* note 97, at 12-13. This justification for protecting labor speech is particularly strong for nonpicketing publicity. An individual restrained from picketing may have alternate avenues of communication to convey his message, but a prohibition extended to nonpicketing publicity cuts off an entire message. While other avenues may be available to pressure the *primary* employer, the union member nevertheless may feel frustrated by his inability to make the *secondary* bear the full responsibility of his relationship to and possible support of—whether by action or inaction—the primary.

IV. JUSTIFYING THE RESTRICTIONS ON PROTECTED SPEECH

The second stage in the analysis of the constitutionality of the "producer-distributor" and "for the purpose of" requirements is an evaluation of the government's justification for infringing upon nonpicketing labor speech. To regulate fully protected speech¹²⁶ on the basis of its content, the government must show that the regulation is a narrowly tailored means to promote a compelling government interest.¹²⁷ The "producer-distributor" and "for the purpose

126. Congress may regulate fully protected speech if the regulation is "a reasonable, content-neutral, time, place or manner regulation." *See, e.g.,* Grayned v. City of Rockford, 408 U.S. 104 (1972); Adderley v. Florida, 385 U.S. 39 (1966); Kovacs v. Cooper, 336 U.S. 77 (1949). The "for the purpose of" and "producer-distributor" requirements, however, prohibit certain messages and consumer appeals at all times, at all places, and in every manner, *see infra* notes 128-32 and accompanying text, and, therefore, lack the content neutrality required of reasonable time, place, and manner regulations. *See, e.g.,* Metromedia, Inc. v. City of San Diego, 453 U.S. 490, 516-17 (1981) (plurality opinion); Consolidated Edison Co. v. Public Serv. Comm'n, 447 U.S. 530, 536 (1980); Virginia State Bd. of Pharmacy v. Virginia Citizens Consumer Council, Inc., 425 U.S. 748, 771 (1976).

The government also may regulate fully protected speech if the communication is part speech and part conduct and the regulation promotes an important or substantial government interest, is unrelated to the suppression of free expression, and imposes incidental restrictions on first amendment freedoms no greater than is essential to further the government's interest. *See, e.g.,* United States v. O'Brien, 391 U.S. 367 (1968). The *O'Brien* test—whether viewed as limited to expressive conduct or as representing a subcategory of time, place, and manner regulations—and regardless of the feasibility of separating conduct from expression, *see generally, L. TRIBE, supra* note 62, at 599-601; Baker, *supra* note 60, at 1009-12; Cox, *supra* note 62, at 48-49; Ely, *supra* note 62, at 1493-95; Farber, *Content Regulation and the First Amendment: A Revisionist View*, 68 GEO. L.J. 727, 743-46 (1980); Redish, *The Content Distinction in First Amendment Analysis*, 34 STAN. L. REV. 113, 148 n.204 (1981)—has no application to the nonpicketing publicity upon which the "for the purpose of" and the "producer-distributor" requirements infringe. The restrictions resulting from those requirements are not "unrelated to the suppression of expression." The government's concern arises from the nonpicketing publicity's communicative impact—the way the public is expected to respond to the union's message—and is explicitly content-based. *See Ely, supra* note 62, at 1497; *infra* notes 128-32 and accompanying text.

The government also may regulate fully protected speech that occurs in a nonpublic or special forum. *See, e.g.,* Connick v. Myers, 51 U.S.L.W. 4436 (Apr. 20, 1983) (office environment of government employer); Jones v. North Carolina Prisoners' Labor Union, Inc., 433 U.S. 119 (1977) (prison); Greer v. Spock, 424 U.S. 828 (1976) (military base); Lehman v. City of Shaker Heights, 418 U.S. 298 (1974) (interior advertising space of mass transit vehicle). This justification for restricting protected speech is not even arguably applicable to the "for the purpose of" and "producer-distributor" requirements, which apply to publicity conducted through traditional channels of communication that are open to the public. When labor publicity occurs on private property, an employer nondiscriminatorily may enforce valid trespass laws, regardless of the "for the purpose of" and "producer-distributor" requirements. *See* Sears, Roebuck & Co. v. San Diego County Dist. Council of Carpenters, 436 U.S. 180 (1978); *Hudgens v. NLRB*, 424 U.S. 507 (1976).

127. *See* Bolger v. Youngs Drug Prods. Corp., 51 U.S.L.W. 4961, 4962 (June 24, 1983); United States v. Grace, 51 U.S.L.W. 4444, 4446 (Apr. 20, 1983); Consolidated Edison Co. v. Public Serv. Comm'n, 447 U.S. 530, 540 (1980); First Nat'l Bank v. Bellotti, 435 U.S. 765,

of” requirements, however, do not satisfy any of the judicially approved bases for regulating protected speech. The two restrictions on nonpicketing labor publicity are content-based and are not the least restrictive means of achieving a sufficiently compelling governmental interest.

A. Content-Based Restrictions

The “for the purpose of” and “producer-distributor” requirements most assuredly are content-based. The two restrictions do not merely regulate labor speech in the context of a statute balancing labor and management rights¹²⁸ or present difficult questions concerning the content-neutrality of *subject-matter* regulation.¹²⁹ Rather, the “for the purpose of” and “producer-distributor” requirements restrict speech that presents a particular point of view—the union’s. Moreover, the restriction is triggered only when the union’s speech has a particular message.¹³⁰ Even in labor picketing cases members of the Court have recognized that restrictions on appeals to consumers to boycott a secondary are content-based.¹³¹ A fortiori when the restriction completely cuts off a particular message, regardless of whether the communication is speech only or speech plus conduct the government has regulated on the basis of content.¹³²

786 (1978); *NAACP v. Button*, 371 U.S. 415, 438-39 (1963); see also Redish, *supra* note 126, at 113; Stone, *supra* note 62, at 82.

The government subjects certain categories of speech to content regulation on a regular basis. See, e.g., *Gertz v. Robert Welch, Inc.*, 418 U.S. 323 (1974) (libelous speech); *Miller v. California*, 413 U.S. 15 (1973) (obscenity); *Chaplinsky v. New Hampshire*, 315 U.S. 568 (1942) (fighting words). These categories, however, are not exceptions to the general rule against content-based restrictions on protected speech, but are examples of speech that the Court has found not worthy of first amendment protection.

128. See Farber, *supra* note 126, at 736; Stephan, *supra* note 55, at 204.

129. See generally Farber, *supra* note 126; Karst, *Equality as a Central Principle in the First Amendment*, 43 U. CHI. L. REV. 20 (1975); Stone, *supra* note 62.

130. For example in *Pet, Inc.*, 244 N.L.R.B. 96, *rev'd sub nom. Pet, Inc. v. N.L.R.B.* 641 F.2d 82 (8th Cir. 1981), the Union could have circulated handbills in the same location requesting that the public not shop at Hussmanns. The Union, however, apparently lost its first amendment rights by including other Pet subsidiaries in its plea. Similarly, in *Delta Air Lines*, 263 N.L.R.B. 153, 111 L.R.R.M. (BNA) 1159 (Sept. 10, 1982) the Union handbill would not have violated the Act if it had not included information about Delta’s flight safety record. The inclusion of that specific message caused the Union to lose its protection.

131. See, e.g., *NLRB v. Retail Store Employees Local 1001 (Safeco)*, 447 U.S. 607, 618 (1980) (Stevens, J., concurring); *NLRB v. Fruit & Vegetable Packers Local 760 (Tree Fruits)*, 377 U.S. 58, 79 (1964) (Black, J., concurring) (“[W]e have a case in which picketing, otherwise lawful, is banned only when the picketers express particular views.”).

132. See also Ely, *supra* note 62, at 1498. (A regulation is content-based if the need for the regulation would disappear when the audience could not read English).

B. The Government's Interest in the "For The Purpose Of" Requirement

The acknowledged government interest in enacting the secondary boycott provisions of the Act was to protect neutral employers and employees from disputes not of their own making.¹³³ The "for the purpose of" requirement as interpreted by the Board in *Delta Air Lines* does not promote that interest. The Board clearly indicated that it would have found no fault with a Union handbill without the flight safety information and would have permitted the Union to enmesh Delta in its primary dispute by requesting customers to withhold patronage from Delta.¹³⁴ Thus, the "for the purpose of" requirement, as interpreted by *Delta Air Lines*, will not protect neutral employers from labor disputes not of their own making.

A second government interest that arguably justifies the *Delta Air Lines* interpretation of the "for the purpose of" requirement is preventing severe economic harm to the business of an individual employer. That strictly economic interest, however, does not constitute, in the first amendment context, a substantial—much less compelling—government interest.

Furthermore, the government's interest in regulating one activity is not sufficiently compelling when the government does not regulate other activity that jeopardizes the same interest.¹³⁵ The arguable interest in preventing severe economic harm to an individual business suffers from this infirmity. The government fails to regulate not only some other activity, but *all* other activity that produces the same "evil" as the union's publicity.¹³⁶ The decision

133. 93 CONG. REC. 4323 (1947), reprinted in 2 NLRB LEGISLATIVE HISTORY OF THE LABOR MANAGEMENT RELATIONS ACT, 1947, at 1106 (1948) (Statement of Sen. Taft); see also *DeBartolo Corp. v. NLRB*, 103 S. Ct. 2926, 2932 (1983); *NLRB v. Denver Bldg. & Constr. Trade Council*, 341 U.S. 675, 692 (1951); *Service Employees Local 399 (Delta Air Lines)*, 263 N.L.R.B. No. 153, 111 L.R.R.M. (BNA) 1159, 1161-62 (Sept. 10, 1982).

134. *Service Employees Local 399 (Delta Air Lines)*, 263 N.L.R.B. No. 153, 111 L.R.R.M. (BNA) 1159, 1162 (Sept. 10, 1982).

135. See, e.g., *Metromedia, Inc. v. City of San Diego*, 453 U.S. 490, 520 (1981) (plurality opinion); *Carey v. Brown*, 447 U.S. 455, 465 (1980); *First Nat'l Bank v. Bellotti*, 435 U.S. 765, 793 (1978); *Police Dept. of Chicago v. Mosley*, 408 U.S. 92, 100 (1972).

136. The defect of the "for the purpose" requirement is not simply that too little speech falls within the prohibition. Rather, by ignoring some speech that is indistinguishable from the speech prohibited, the government has undermined the likelihood that a genuine state interest exists for restriction of the prohibited speech. As explained by Justice White, "The exceptions do not create the infringement, rather the general prohibition does. But the exceptions to the general prohibition are of great significance in assessing the strength of the city's interest . . ." *Metromedia, Inc. v. City of San Diego*, 453 U.S. 490,

in *Delta Air Lines* would have allowed any person other than a Union member to publicize the same information contained in the Union's handbill without incurring liability. For example, Ralph Nader or even a competing business would not violate the law by publishing Delta's flight safety record,¹³⁷ yet the injury to Delta would be as great as, if not greater than, the injury the Union's handbilling caused.¹³⁸

A final arguable justification for the "for the purpose of" requirement is that the restriction serves the government interest in preventing the dissemination of misleading information. The concern is that union publication of information about the secondary—such as Delta's safety record—may present the data misleadingly or lead the public to believe that a primary dispute exists with the secondary. While the Board's "for the purpose of" requirement may prevent the dissemination of misleading data by prohibiting all information unrelated to the primary dispute, the restriction obviously is overbroad. Handbills and advertisements are memorialized on paper—the government easily can police their truthfulness on an individual basis.¹³⁹ Since "broad prophylactic rules in the area of free expression are suspect,"¹⁴⁰ the solution to

520 (1981) (plurality opinion).

137. See RESTATEMENT (SECOND) OF TORTS § 768 (1982). First amendment rights should not depend solely upon who the speaker is. See, e.g., *First Nat'l Bank v. Bellotti*, 435 U.S. 765, 777 (1978). The "for the purpose of" restriction is particularly objectionable because it makes first amendment rights depend not only on who the speaker is, but also on whether the speaker is a union member. The restriction therefore also may violate first amendment associational rights. See *Baker*, *supra* note 60, at 1031-35 (first amendment protects association as well as assembly); see also *NAACP v. Claiborne Hardware Co.*, 102 S. Ct. 3409, 3423 (1982); *NAACP v. Alabama*, 357 U.S. 449, 460-61 (1958).

138. When the Union publishes Delta's flight safety record, the public, believing that a vindictive motive prompted the Union to publish the information, may discount the record's credibility. See *NLRB v. Local Union 1229, Int'l Bhd. of Elec. Workers*, 346 U.S. 464, 477 (1953).

139. Cf. *Ohralik v. Ohio State Bar Ass'n*, 436 U.S. 447, 466 (1978) ("in-person solicitation is not visible or otherwise open to public scrutiny," therefore, the government may justify more easily prophylactic regulation).

140. See, e.g., *Village of Schaumburg v. Citizens for a Better Env't*, 444 U.S. 620, 637 (1980); *In re Primus*, 436 U.S. 412, 432 (1978); *NAACP v. Button*, 371 U.S. 415, 438 (1963). But see *Ohralik v. Ohio State Bar Ass'n*, 436 U.S. 447 (1978).

In *Ohralik*, a commercial speech case, the Court endorsed the use of broad prophylactic regulations. 436 U.S. at 464. Labor speech, however, is not commercial speech, nor should it be treated similarly. See *supra* notes 97-125 and accompanying text. Moreover, the more demanding inquiry into the government's interest and the precision of its regulation that the Supreme Court has made in other commercial speech cases indicates that *Ohralik* may be limited to its facts. See, e.g., *Bolger v. Youngs Drug Prods. Corp.*, 103 S. Ct. 2875 (1983); *In re R.M.J.*, 455 U.S. 191 (1982); *Central Hudson Gas & Elec. Corp. v. Public Serv. Comm'n*, 447 U.S. 557 (1980); *In re Primus*, 436 U.S. 412 (1978); *First Nat'l Bank v. Bel-*

the problem of potentially misleading messages is not to enjoin all messages as did the Board in *Delta Air Lines*, but to restrain only the misleading ones¹⁴¹ or perhaps to require greater disclosure.¹⁴² Thus, the "for the purpose of" requirement is not narrowly tailored to accomplish any sufficiently compelling government interest.¹⁴³

C. The Government's Interest in the "Producer-Distributor" Requirement

The constitutional validity of the "producer-distributor" requirement also is suspect. The producer-distributor language attempts to further the acknowledged government interest of protecting neutral parties from labor disputes not of their own making¹⁴⁴ by focusing on the secondary's relationship to the primary: the more remote the relationship, the more "neutral" the secondary, and the less the union can justify the publicity. In contrast, when the secondary distributes the primary's product, a close relationship exists between the parties, and the secondary is less neutral, and a secondary boycott directed against the primary's product¹⁴⁵ injures the secondary roughly¹⁴⁶ to the extent the secon-

lotti, 435 U.S. 765 (1978).

141. The Board in *Delta Air Lines* discussed the possibility that the safety data might be misleading, but did not rule on that basis. The Board's injunction prohibited *all* publication of information unrelated to the primary dispute, not merely nontruthful or misleading publicity. See 263 N.L.R.B. 996, 111 L.R.R.M. 1159 (BNA) (Sept. 10, 1982). Moreover, the majority expressly rejected the administrative law judge's finding that the publicity proviso did not cover the handbill because it contained "misleading unrelated" information. *Id.* at 998, 111 L.R.R.M. at 1161. Finally, to the extent the handbill misled the public about the parties to the primary dispute, the Union's appeal probably was less effective and the government's interest not as great. See *supra* note 138.

142. See *infra* notes 156-66 and accompanying text.

143. The Court's decision in *NLRB v. Local Union 1229, Int'l Bhd. of Elec. Workers (Jefferson Standard)*, 346 U.S. 464 (1953), ostensibly supports the constitutionality of the restriction in *Delta Air Lines*. In *Jefferson Standard*, technicians employed by, and in a labor dispute with a broadcasting company, distributed handbills attacking the quality of the company's product. The Board and the Supreme Court found that the employer was justified in terminating the responsible technicians since the disparaging information was unrelated to the primary dispute. No constitutional issue, however, arose in *Jefferson Standard*. The Court did not prohibit the employees from publicizing their message, but merely found that the employees' activity had no affirmative statutory protection. The case involved state action.

144. See *supra* note 133.

145. A boycott directed against only the primary's product is a product boycott. A union's appeal to the consumer to withhold *all* patronage from the secondary is a total boycott. A product boycott in some cases may "merge" into a total boycott when the primary product is inseparable from the secondary's remaining business. See, e.g., *Hoffman v. Cement Masons Local 337*, 468 F.2d 1187 (9th Cir. 1972) (struck product was homes, secon-

dary benefits from its relationship with the primary. While this focus on the relationship between the primary and secondary may appear tailored to the government's interest,¹⁴⁷ the Board and courts have interpreted the prohibitions of section 8(b)(4)(ii)(B) and the producer-distributor language of the publicity proviso in a manner inconsistent with the government's acknowledged goal. The Board and courts have not focused properly on the injury to the secondary vis-a-vis its relationship to the primary; rather, they have adopted a bifurcated approach when construing section 8(b)(4)(ii)(B) that belies the government's acknowledged interest.

The inquiry under the "producer-distributor" requirement is whether some business relationship or chain of production nexus exists between the primary and secondary.¹⁴⁸ When some minimal relationship exists, the law permits total boycotts as well as product boycotts.¹⁴⁹ The law does not attempt to tailor the extent of permissible injury to the nature of the secondary's relationship

dary employer was real estate developer), *cert. denied*, 411 U.S. 986 (1973); *American Bread Co. v. NLRB*, 411 F.2d 147 (6th Cir. 1969) (struck product was bread, secondary employer was restaurant owner). When a merger occurs, the Board and the courts have treated the appeal as one for a total boycott. *Id.* at 154.

146. Some consumers may withhold all patronage from the secondary even when the union requests only a product boycott. *See NLRB v. Fruit & Vegetable Packers Local 760 (Tree Fruits)*, 377 U.S. 58, 82-83 (1964) (Harlan, J., dissenting). With that exception, however, the secondary boycott would have no greater effect on the secondary employer than an effective primary strike.

147. The statute, even if construed literally, still might be insufficiently tailored to the government's interest. By focusing on the lines of production and distribution existing between primary and secondary, the proviso ignores other situations in which the secondary's contributions to the primary are far greater than that of a distributor. *But see supra* note 48. For example, when the secondary is the parent of the primary but does not have apparent actual control, the Board might consider the parent neutral, even though the parent company may substantially influence the labor policy of the primary. *See supra* note 24; Comment, *Unions, Conglomerates, and Secondary Activity Under the NLRA*, *supra* note 24, at 239 ("A conglomerate's power in labor negotiations stems from its size and diversification, which give it the ability to 'cross-subsidize between industries and plants and whipsaw different unions at . . . various facilities—supported by substantially enhanced financial staying power.'") (quoting Alexander, *Conglomerate Mergers and Collective Bargaining*, 24 *INDUS. & LAB. REL. REV.* 354, 362 (1971)); *see also* Comment, *Consumer Picketing of Economically Interdependent Parties: Retail Store Employees Local 1001 v. NLRB (Safeco Title Insurance Co.)*, 32 *STAN. L. REV.* 631, 643 (1980). This criticism, of course, may apply more strongly to the Board's neutrality definition than to the "producer-distributor" requirement.

148. *See supra* note 48.

149. *See, e.g., United Steelworkers (Pet, Inc.)*, 244 N.L.R.B. 96, 101 (1979), *rev'd on other grounds sub nom. Pet, Inc. v. NLRB*, 641 F.2d 545 (8th Cir. 1981); *American Fed'n of Television and Radio Artists Local 55 (Great Western Broadcasting Corp.)*, 150 N.L.R.B. 467, 470-72 (1964), *aff'd sub nom. Great Western Broadcasting Corp. v. NLRB*, 356 F.2d 434 (9th Cir.), *cert. denied*, 384 U.S. 1002 (1966).

with the primary, and as a result, a secondary employer can incur injury far beyond that which its relationship with the primary warrants. The inquiry, when the publicity proviso is found inapplicable, even more starkly belies a government interest in protecting secondary parties to the extent of their neutrality. The reviewing body asks whether the boycott threatens substantial loss or financial ruin to the secondary,¹⁵⁰ protecting the secondary only when the boycott affects a substantial portion of the secondary's business. Thus, rather than protecting relatively innocent parties, the courts have construed the statute to provide more protection to those secondary distributors who are *less neutral*.¹⁵¹

The chief problem, however, with the "producer-distributor" requirement is not its failure to focus on the harm to the secondary vis-a-vis the extent of its neutrality. Rather, it is that the restriction is not the least restrictive alternative necessary to achieve the government's goal. The government can achieve its goal equally well by ensuring that the public learns of the actual relationship between the secondary and primary employers.

A union is unlikely to conduct a consumer boycott of a secondary that has no business relationship with the primary. The possible pressure a secondary could apply to the primary would be minimal. More importantly, even if informal means of pressure did exist—through, for example, personal relationships of officers or directors of the primary and secondary—the secondary likely could not be persuaded to exercise the pressure through a consumer boycott since the consumer appeal would have little effect on the secondary *if* the public was aware of the secondary's neutrality.¹⁵²

On the other hand, if the secondary does have a business relationship with the primary, the government's interest does not re-

150. See *supra* note 49.

151. This, of course, assumes the Court has not determined the secondary to be non-neutral, see *supra* note 24; see also *NLRB v. Business Mach. & Office Appliance Mechanics Local 459 (Royal Typewriter)*, 228 F.2d 553 (2d Cir. 1955) (ally doctrine), *cert. denied*, 351 U.S. 962 (1956), or otherwise outside the prohibitions of § 8(b)(4)(ii)(B). See generally *supra* note 49. As the *Safeco* case illustrates, however, the secondary can have a substantial relationship with the primary without falling within either of these categories.

152. This highlights yet another infirmity with the Act's restrictions on nonpicketing publicity. The Act by not requiring the charging party to show *actual* harm to the secondary or imminent coercion of the primary before the trier issues an injunction thus sweeps within its scope harmless as well as harmful speech and speech that does not produce a "clear and present danger" of the feared substantive evil. See Jones, *The Right to Picket—Twilight Zone of the Constitution*, 102 U. Pa. L. Rev. 995, 1020-27 (1954). Since the Act does require a showing of harm for a damage award, see National Labor Relations Act § 303, 29 U.S.C. § 187 (1976), the problems of proof can not justify their overbreadth.

quire that the secondary receive full protection from the union's message since the secondary no longer would be completely neutral. While the interested secondary arguably should not incur injury far beyond its relationship to the primary, here again, the best judge of the extent of the secondary's neutrality is the public. A public that is aware of the secondary's relationship can decide whether the relationship is significant enough to warrant the pressure the union urges.¹⁵³ Thus, the key to achieving the government's interest is full and accurate disclosure of the secondary's relationship with the primary.

The central failing of both the "producer-distributor" and "for the purpose of" requirements ultimately is the government's unwillingness to allow truthful speech for fear that the public may not perceive its true merit or give it proper weight. Such paternalistic regulation cannot meet the exacting scrutiny of the compelling interest standard,¹⁵⁴ and likely would be infirm even under the lesser commercial speech standard.¹⁵⁵

153. The collective judgment may distort the individual's judgment—that is, many consumers may feel that a secondary's limited relationship with the primary justifies limited withholding of patronage and therefore decide to shop elsewhere, unaware that many other consumers are doing likewise. The result, of course, would be a reduction in the secondary's sales beyond that which the individual consumer desired. This "very distant possibility of harm", however, does not justify infringing first amendment rights. *In re Primus*, 436 U.S. 412, 436 (1978).

154. *See id.* at 432; *Buckley v. Valeo*, 424 U.S. 1, 44-45 (1976); *Cohen v. California*, 403 U.S. 15, 24 (1971); *see also Redish, supra* note 126, at 143; *Stone, supra* note 62, at 82.

155. In *Central Hudson Gas & Elec. Corp. v. Public Serv. Comm'n*, 447 U.S. 557 (1980), the Court announced a four-part analysis for restrictions on commercial speech:

At the outset, we must determine whether the expression is protected by the First Amendment. For commercial speech to come within that provision, it at least must concern lawful activity and not be misleading. Next, we ask whether the asserted governmental interest is substantial. If both inquiries yield positive answers, we must determine whether the regulation directly advances the governmental interest asserted, and whether it is not more extensive than is necessary to serve that interest.

Id. at 566.

The Court has applied these requirements rigorously, *see, e.g., Bolger v. Youngs Drug Prods. Corp.*, 103 S. Ct. 2875 (1983); *Central Hudson Gas & Elec. Corp. v. Public Serv. Comm'n*, 447 U.S. 557 (1980); *First Nat'l Bank v. Bellotti*, 435 U.S. 765 (1978); *Bates v. State Bar of Arizona*, 433 U.S. 350 (1977); *Virginia State Bd. of Pharmacy v. Virginia Citizens' Consumer Council, Inc.*, 425 U.S. 748 (1976), and repeatedly has acknowledged that the government may not suppress commercial speech entirely merely because the government fears the misuse of such speech. *See, e.g., Central Hudson*, 447 U.S. at 566 n.9; *First Nat'l Bank*, 435 U.S. at 791-92 & n.31; *Virginia State Bd. of Pharmacy*, 425 U.S. at 770. The restrictions on nonpicketing labor publicity suffer from precisely that infirmity.

Moreover, even in the case of commercial speech, the preferred solution to potential harm is more—not less—speech, *see infra* note 159, the simple expedient that has eluded the government in its restrictions on nonpicketing publicity. Thus, even under the lesser protection due commercial speech, which is the minimum protection due nonpicketing labor

V. A LESS RESTRICTIVE ALTERNATIVE

The "producer-distributor" and "for the purpose of" requirements are not the least restrictive means to achieve the government's goal to ensure that the union's appeals do not unjustly injure secondary employers.¹⁵⁶ The government can achieve its interest as effectively and less restrictively by more, not less, disclosure, a remedy that also accords well with the theoretical underpinnings of the first amendment.¹⁵⁷ Full disclosure would provide an informed public that could decide best if and the degree to which the public should boycott the secondary. Full disclosure would protect substantially innocent parties without trampling first amendment rights. The government could place responsibility for disclosure on either the secondary, the union, or both.

First, the law might require that the secondary employers themselves provide the additional disclosure of information about the secondary-to-primary relationship. Secondary employers can counteradvertise, speak to consumers personally, or display signs in their company windows. Placing the burden of disclosure on the secondary alone, however, may be a more attractive than practical solution. A secondary employer may not know of the union's activity or what the union's message is until after the union has begun its appeal. Moreover, counteradvertising may not reach the identical audience that the union appeal reaches.

speech, *see supra* notes 96-125 and accompanying text, the "for the purpose of" and "producer-distributor" requirements violate the first amendment.

156. Of course, least restrictive analysis can be a "slippery slope" for, as some justices have recognized, "[a] judge would be unimaginative indeed if he could not come up with something a little less 'drastic' or a little less 'restrictive' in almost any situation, and thereby enable himself to vote to strike legislation down." *Widmar v. Vincent*, 454 U.S. 263, 279 n.3 (1981) (Stevens, J., concurring) (quoting *Illinois Elections Bd. v. Socialist Workers Party*, 440 U.S. 173, 188-89 (1979) (Blackmun, J., concurring)). By the same token, however, less restrictive alternative analysis requires something more than merely deferring to the legislative judgment when the alternate regulation is only marginally less effective than the regulations that restrict first amendment rights. *See Ely, supra* note 62, at 1486-87; *see also Landmark Communications, Inc. v. Virginia*, 435 U.S. 829, 843 (1978).

The Court has often applied a rigorous, less restrictive alternative analysis even in the commercial context. *See, e.g., Consolidated Edison Co. v. Public Serv. Comm'n*, 447 U.S. 530, 540-41 (1980); *Virginia State Bd. of Pharmacy v. Virginia Citizens Consumer Council, Inc.*, 425 U.S. 748, 770 (1976). Accordingly, a rigorous analysis is appropriate here, particularly since nonpicketing labor publicity is a traditional mode of communication, *see Ely, supra* note 62, at 1488, and since the proposed remedy relies on opening, not closing, the channels of communication. *See In re R.M.J.*, 455 U.S. 191, 201 (1982) ("preferred remedy is more disclosure, rather than less"); *Virginia State Bd. of Pharmacy v. Virginia Citizens Consumer Council, Inc.*, 425 U.S. 748, 770 (1976).

157. *See Linmark Associates, Inc. v. Township of Willingboro*, 431 U.S. 85, 97 (1977) (quoting *Whitney v. California*, 274 U.S. 357, 377 (1927) (Brandeis, J., concurring)).

The second and preferred solution is to require the union to make the disclosure. The Board already treats publicity that does not sufficiently identify the primary as untruthful.¹⁵⁸ A requirement that the union disclose additional information to prevent an appeal from misleading the public would be within the limits of this precedent.¹⁵⁹

The union disclosure solution, however, does not resolve the issue of what constitutes enough disclosure. The current Board truthfulness requirement prohibits publicity that substantially departs from fact or that is intentionally deceptive,¹⁶⁰ but because of difficult questions of intent, this standard may not suffice to ensure full disclosure.¹⁶¹ The Board or the court, however, could engraft onto the truthfulness requirement a simple burden-shifting procedure to alleviate this problem and still provide the union with some measure of predictability. For example, the Board could give the union the burden of asking the secondary employer to provide

158. See, e.g., *Local 732, Int'l Bhd. of Teamsters (Servair)*, 229 N.L.R.B. 392, 392 (1977); *Honolulu Typographical Union, No. 37*, 167 N.L.R.B. 1030, 1032 (1967), *aff'd*, *Honolulu Typographical Union, No. 37 v. NLRB*, 401 F.2d 952, 957-58 (D.C. Cir. 1968).

159. Questions arise whether the proposed disclosure requirements would be found constitutional. To the extent the omission makes the publicity untruthful, the answer is simple—untruthful speech is subject to regulation. See *Virginia State Bd. of Pharmacy v. Virginia Citizens Consumer Council, Inc.*, 425 U.S. 748, 771 (1976); *Gertz v. Robert Welch, Inc.*, 418 U.S. 323, 340 (1974); *Linn v. United Plant Guard Workers Local 114*, 383 U.S. 53, 67 (1966). Merely because the Board defines misleading speech as untruthful, however, does not make it so. The Supreme Court, nevertheless, has broadly endorsed disclosure requirements at least in the commercial speech context, to prevent misleading, even if not untruthful, messages. See *In re R.M.J.*, 455 U.S. 191, 203 (1982) (the remedy in the first instance for potentially deceptive speech is not necessarily a prohibition but preferably a requirement of disclaimers or explanation); see also *id.* at 200 & n.11; *Central Hudson Gas & Elec. Corp. v. Public Serv. Comm'n*, 447 U.S. 557, 570-71 (1980); *First Nat'l Bank v. Bellotti*, 435 U.S. 765, 792 n.32 (1978); *Bates v. State Bar of Arizona*, 433 U.S. 350, 384 (1977). But see *Friedman v. Rogers*, 440 U.S. 1, 12 n.11 (1979). The Court has reasoned that the commercial speaker is in a position to know the truth of his message and has a strong financial motive that would overcome any chilling effect that requiring disclosure would produce. See *Central Hudson Gas & Elec. Corp. v. Public Serv. Comm'n*, 447 U.S. 557, 564 n.6 (1980); *Friedman v. Rogers*, 440 U.S. 1, 10 (1979); *Virginia State Bd. of Pharmacy v. Virginia Citizens Consumer Council, Inc.*, 425 U.S. 748, 771-72 & n.24 (1976). Whatever the merits of a category-by-category approach to chilling effects, see *Farber, supra* note 126, at 759, the Court likely would find the same arguments applicable to labor speech. The union is a participant in the labor dispute and therefore is likely to know the facts. Moreover, the union's strong motivation to present its message frequently leads it not only to ignore threats of an injunction, but also to speak in violation of an injunction. But see *T. EMERSON, supra* note 95, at 449 (full protection theory applied to labor publicity would mean that the government could not impose a truthfulness requirement).

160. See *DeBartolo*, 252 N.L.R.B. at 704 n.2; *International Bhd. of Teamsters Local 537 (Lohman Sales Co.)*, 132 N.L.R.B. 901, 906 (1961).

161. See *supra* note 33.

a disclaimer about its connection to the primary. The union's request would shift the burden of providing such a disclaimer to the secondary. The union, exercising its option to include the secondary's disclaimer in its publicity, would raise the potentially irrebuttable presumption that both parties had made sufficient disclosure.¹⁶² The presumption would serve as the union's "safe harbor." The union's choice not to include the secondary's statement in the union publicity would raise a presumption that the union "intended to deceive." The union could rebut by explaining why it excluded the disclaimer¹⁶³ and by demonstrating that the message nonetheless was truthful and not misleading.

This remedy, admittedly, may not prevent all harm to a neutral secondary. For example, consumers may not read the union's entire message yet may respond to the appeal to withhold patronage. Alternatively, the initial receptors of the union message may perceive the secondary's true relationship to the primary yet convey that message incorrectly to others.¹⁶⁴ The secondary, however, partially may reduce these harms by counteradvertising. Unlike the voter in an election, a consumer may recast his economic vote.¹⁶⁵ More fundamentally, "the people in our democracy are entrusted with the responsibility for judging and evaluating the relative merits of conflicting arguments. . . . [I]f there be any danger that the people cannot evaluate the information and arguments advanced . . . , it is a danger contemplated by the Framers of the First Amendment."¹⁶⁶

VI. CONCLUSION

Labor increasingly has used the secondary boycott as a weapon to apply economic pressure upon noncooperating employers. The secondary boycott raises statutory issues that the Na-

162. Cf. *International Bhd. of Teamsters, Local 537 (Lohman Sales Co.)*, 132 N.L.R.B. 901, 906 (promptly remedied handbill was evidence that the union did not intend to deceive).

163. The union might decide to exclude the secondary employer's disclaimer from the union's handbill because the disclaimer was too long, too costly, or inaccurate.

164. Nonpicketing publicity also may produce certain harms such as violence or restricted access that are unrelated to the union's message. State law, however, provides sufficient remedy for those injuries. See *UAW v. Wisconsin Employment Relations Bd.*, 351 U.S. 266, 274-75 (1956); see also *supra* note 126.

165. But see Teller, *supra* note 61, at 201; Note, *Freedom of Expression in a Commercial Context*, 78 HARV. L. REV. 1191, 1208 (1965).

166. *First Nat'l Bank v. Bellotti*, 435 U.S. 765, 791-92 (1978); see also *Virginia State Bd. of Pharmacy v. Virginia Citizens Consumer Council, Inc.*, 425 U.S. 748, 770 (1976); Stone, *supra* note 62, at 104.

tional Labor Relations Board and the courts have analyzed extensively. The courts also have considered the constitutionality of restrictions on one form of secondary pressure, consumer picketing. The Board and the courts, however, repeatedly have avoided addressing the constitutionality of restrictions on nonpicketing forms of secondary consumer pressure. The legitimacy of such restrictions remains in doubt.

This Article argues that nonpicketing labor publicity should receive full first amendment protection. Nonpicketing labor publicity shares few of the qualities that have justified the lesser protection accorded labor picketing. Despite its economic motivation, nonpicketing labor publicity also is more deserving of full first amendment protection than commercial speech, an area of speech in which restrictions have come under increasingly rigorous scrutiny. Unlike commercial speech, nonpicketing publicity fully satisfies all the traditional values that underlie the First Amendment.

The restrictions that recent Board and court decisions impose on nonpicketing publicity therefore infringe fully protected speech. That infringement remains unjustified. The "for the purpose of" and "producer-distributor" requirements are content-based restrictions that are not drawn narrowly to further a compelling government interest. By ignoring the simple expedient of requiring more, not less, speech, the restrictions are likely infirm under the reduced commercial speech standard which is the minimum standard that courts can apply to nonpicketing labor publicity. In an era of complex economic relationships, the use of secondary consumer pressure is likely to increase and take novel forms. Accordingly, the failure of the Board and courts to provide guidance to unions and their advisors on the constitutional protection due such activity is most unfortunate. The courts should adopt the analysis that this Article recommends and remedy their failure to date to provide this guidance.

Defending the Citadel: The Dangerous Attack of "Reasonable Good Faith"

Stanley Ingber*†

During the last half century the proper scope of the fourth amendment¹ and its related exclusionary rule has provoked heated discussion among jurists² and scholars.³ The Supreme Court has interpreted the amendment as providing constitutional recognition of the individual's right to privacy.⁴ The amendment, therefore,

* Professor of Law, University of Florida, Holland Law Center. B.A., 1969, Brooklyn College; J.D., 1972, Yale University.

† My appreciation to Penelope Ingber, my wife, for her immeasurable help in writing this Article.

1. The fourth amendment guarantees,

The right of the people to be secure in their persons, houses, papers, and effects, against unreasonable searches and seizures, shall not be violated, and no Warrants shall issue, but upon probable cause, supported by Oath or affirmation, and particularly describing the place to be searched, and the persons or things to be seized.

U.S. CONST. amend. IV.

2. For a sampling of this debate in recent cases, see *United States v. Place*, 103 S. Ct. 2637 (1983); *Illinois v. Gates*, 103 S. Ct. 2317 (1983); *Florida v. Royer*, 103 S. Ct. 1319 (1983); *United States v. Ross*, 456 U.S. 798 (1982); *New York v. Belton*, 453 U.S. 454 (1981); *Michigan v. Summers*, 452 U.S. 692 (1981); *Rawlings v. Kentucky*, 448 U.S. 98 (1980); *United States v. Mendenhall*, 446 U.S. 544 (1980); *Rakas v. Illinois*, 439 U.S. 128 (1978).

3. See, e.g., Amsterdam, *Perspectives on the Fourth Amendment*, 58 MINN. L. REV. 349 (1974); Ashdown, *The Fourth Amendment and the "Legitimate Expectation of Privacy"*, 34 VAND. L. REV. 1289 (1981); Canon, *Is the Exclusionary Rule in Failing Health? Some New Data and a Plea Against a Precipitous Conclusion*, 62 KY. L.J. 681 (1974); Kamisar, *Is the Exclusionary Rule an "Illogical" or "Unnatural" Interpretation of the Fourth Amendment?*, 62 JUDICATURE 66 (1978); Kaplan, *The Limits of the Exclusionary Rule*, 26 STAN. L. REV. 1027 (1974); LaFave, *The Fourth Amendment in an Imperfect World: On Drawing "Bright Lines" and "Good Faith"*, 43 U. PITT. L. REV. 307 (1982); McMillian, *Is there Anything Left of the Fourth Amendment?*, 24 ST. LOUIS U.L.J. 1 (1979); Mertens & Wasserstrom, *The Good Faith Exception to the Exclusionary Rule: Deregulating the Police and Derailing the Law*, 70 GEO. L.J. 365 (1981); Miles, *Decline of the Fourth Amendment: Time to Overrule Mapp v. Ohio?*, 27 CATH. U.L. REV. 9 (1977); Oaks, *Studying the Exclusionary Rule in Search and Seizure*, 37 U. CHI. L. REV. 665 (1970); Schlesinger, *The Exclusionary Rule: Have Proponents Proven That it is a Deterrent to Police?*, 62 JUDICATURE 404 (1979); Wilkey, *The Exclusionary Rule: Why Suppress Valid Evidence?*, 62 JUDICATURE 214 (1978).

4. See *Katz v. United States*, 389 U.S. 347, 350-51 (1967); see also *id.* at 360-61 (Harlan, J., concurring).

protects the individual from unwarranted governmental intrusion upon his person and domain. This protection is essential to the maintenance of the proper balance between the life of a person as an individual and his life as a member of society.⁵ Accordingly, police normally must have probable cause, and often a search warrant, before they lawfully can invade an individual's privacy to search for or seize evidence.⁶ Few scholars question the value of so protecting individual privacy. In contrast, however, the wisdom of using the exclusionary rule as a procedural remedy for governmental violations of fourth amendment guarantees has been a topic of extensive debate.

Under the exclusionary rule courts must exclude at trial any evidence seized during searches made in violation of fourth amendment standards.⁷ Since its earliest pronouncement,⁸ critics have condemned this procedure that keeps relevant and reliable evidence from the trial fact-finder. Justice (then Judge) Cardozo, for example, described the rule as a mechanism permitting "[t]he criminal . . . to go free because the constable has blundered."⁹

More recently the exclusionary rule's detractors have marshalled their forces to reduce the rule's impact by carving out a good faith exception. The exception provides for admission of illegally obtained evidence if the offending officer acted with the reasonable good faith belief that his conduct conformed to the fourth amendment. In their frustration with the exclusionary rule a num-

5. T. EMERSON, TOWARD A GENERAL THEORY OF THE FIRST AMENDMENT 74 (1966). Justice Felix Frankfurter recognized that the fourth amendment held "a place second to none in the Bill of Rights." *Harris v. United States*, 331 U.S. 145, 157 (1947) (Frankfurter, J., dissenting). Even the first amendment cannot claim greater importance because "speech activity is [but] a small albeit precious part of the lives of most citizens, and even its fullest protection leaves the police unfettered to deal as they please with most of us most of the time." *Amsterdam*, *supra* note 3, at 377-78.

6. See *Katz v. United States*, 389 U.S. at 356-58 (1967); *United States v. Ventresca*, 380 U.S. 102 (1965); cf. *Camara v. Municipal Court*, 387 U.S. 523 (1967). In some situations the search may be sufficiently limited as not to require probable cause. *E.g.*, *Terry v. Ohio*, 392 U.S. 1 (1968). Exigencies also may prevent an officer from obtaining a warrant. *E.g.*, *Carroll v. United States*, 267 U.S. 132 (1925).

7. This statement is somewhat too broad. As discussed in part I, *see infra* text accompanying notes 38-42, illegally obtained evidence is admissible in certain types and stages of various proceedings. But, as a general proposition, courts must exclude illegally seized evidence.

8. The Supreme Court first articulated the exclusionary rule in *Weeks v. United States*, 232 U.S. 383 (1914), but imposed the rule only upon federal courts. Nearly fifty years passed before the Court extended the rule's application to encompass state courts as well. *Mapp v. Ohio*, 367 U.S. 643 (1961).

9. *People v. Defore*, 242 N.Y. 13, 21, 150 N.E. 585, 587 (1926).

ber of courts,¹⁰ jurists,¹¹ state legislatures,¹² legal associations,¹³ and scholars¹⁴ have adopted or supported some form of this good faith exception.¹⁵

Given this climate, the good faith exception unsurprisingly has gained the attention of the Supreme Court. Four, or perhaps even five, of the Justices publicly have acknowledged their unease with the exclusionary rule's effect and their sympathy with suggested reforms.¹⁶ At the Court's request, the attorneys in *Illinois v. Gates*¹⁷ briefed and argued the issue of a good faith exception during the October 1982 term. The Court, however, sidestepped the issue by deciding the case on other grounds.¹⁸ Nevertheless, within

10. The Fifth Circuit Court of Appeals specifically has recognized a good faith exception. *United States v. Williams*, 622 F.2d 830, 846-47 (5th Cir. 1980) (en banc), *cert. denied*, 449 U.S. 1127 (1981). Since the new Eleventh Circuit was carved out of the old Fifth Circuit, *Williams* likely is binding in the Eleventh Circuit as well.

11. See, e.g., *Commonwealth v. Sheppard*, 387 Mass. 488, _____, 441 N.E.2d 725, 733-34 (1982) (Wilkins, J.), *cert. granted sub nom. Massachusetts v. Sheppard*, 103 S. Ct. 3534 (1983).

12. At least two states have adopted by legislation some form of a good faith exception. See ARIZ. REV. STAT. ANN. § 13-3925 (1982); COLO. REV. STAT. § 16-3-308 (1982).

13. The prestigious American Law Institute has recommended the exclusion of evidence only in those cases in which the police violation was "substantial." In determining whether the violation, indeed, was substantial a court should consider all the circumstances, including the willfulness of the violation. MODEL CODE OF PRE-ARRAIGNMENT PROCEDURE § SS 290.2(2), (4) (Official Draft 1975).

14. See, e.g., Ball, *Good Faith and the Fourth Amendment: The "Reasonable" Exception to the Exclusionary Rule*, 69 J. CRIM. L. & CRIMINOLOGY 635 (1978); Bernardi, *The Exclusionary Rule: Is a Good Faith Standard Needed to Preserve a Liberal Interpretation of the Fourth Amendment?*, 30 DE PAUL L. REV. 51 (1980); Sunderland, *The Exclusionary Rule: A Requirement of Constitutional Principle*, 69 J. CRIM. L. & CRIMINOLOGY 141 (1978).

15. In its final report Attorney General William French Smith's bipartisan Task Force on Violent Crime urged the Attorney General to support a good faith exception through court argument and legislative recommendation. ATTORNEY GENERAL'S TASK FORCE ON VIOLENT CRIME, FINAL REP. 55 (Aug. 17, 1981).

16. Four Justices have authored opinions supporting some form of a good faith exception. See *Illinois v. Gates*, 103 S. Ct. 2317, 2340-47 (1983) (White, J., concurring); *Stone v. Powell*, 428 U.S. 465, 501 (1976) (Burger, C.J., concurring); *Brown v. Illinois*, 422 U.S. 590, 611-12 (1975) (Powell, J., concurring in part); *United States v. Peltier*, 422 U.S. 531, 542 (1975) (Rehnquist, J.). Furthermore, Justice O'Connor, at her nomination hearings before the Senate Judiciary Committee, testified about her belief that courts do not have to exclude illegally obtained evidence under the fourth amendment if the court takes into account the good faith of police. *Hearings Before the Senate Comm. on the Judiciary, Nomination of Sandra Day O'Connor*, 97th Cong., 1st Sess. 79-80 (1981) (unbound transcripts on file in Senate Document Room).

17. 103 S. Ct. 2317 (1983).

18. In *Gates* the Court reversed the Illinois court's ruling suppressing evidence obtained under a warrant supported predominately by an anonymous tip. Justice Rehnquist, writing for the majority, found sufficient police corroboration of the tipster's information to justify the magistrate's finding of probable cause. The Justice apologized for the Court's avoidance of the good faith argument, but claimed such avoidance appropriate upon a re-

three weeks of the *Gates* decision the Court accepted for argument during the October 1983 term three cases concerning good faith claims.¹⁹ Obviously, the Justices have signaled their readiness to enter the fray.

Opponents of the good faith exception, therefore, now must voice their arguments. This Article presents arguments that are designed to influence the Court's deliberations, to create a basis for critiquing the Court's opinions once rendered, and to provide guidance for state courts, which soon may need to decide whether a good faith exception is consistent with their state constitutions and procedures.²⁰ To place the subsequent discussion in context, part I

cord which showed the prosecution never had raised or addressed the question before the state tribunals and that none of the Illinois courts had considered the question. *Id.* at 2323. Justice White, concurring separately, chastised the Court for avoiding the issue under a totally unconvincing rationale. Noting that the Supreme Court's pronouncements bound the Illinois courts regarding the federal exclusionary rule, he found requiring a litigant to request a state court to overrule or modify a United States Supreme Court precedent pointless. *Id.* at 2338 (White, J., concurring). Such a practice would undercut *stare decisis*. "Either the presentation of such issues to the lower courts will be a completely futile gesture or the lower courts are now invited to depart from this Court's decisions whenever they conclude such a modification is in order." *Id.*

19. The *United States Law Week* noted the cases and described their good faith issues as follows:

(1) *Colorado v. Quintero*, ___ Colo. ___, 657 P.2d 948, cert. granted, 103 S. Ct. 3535 (1983). "Did Colorado Supreme Court mistakenly interpret Fourth Amendment's prohibition against unreasonable search and seizure when it applied exclusionary rule to undisputed good-faith seizure of stolen property?" 51 U.S.L.W. 3914 (June 28, 1983).

(2) *United States v. Leon*, No. 82-1093 (9th Cir. Jan. 19, 1983), cert. granted, 103 S. Ct. 3535 (1983). "Should Fourth Amendment exclusionary rule be modified so as not to bar admission of evidence seized in reasonable, good-faith reliance on search warrant that is subsequently held to be defective?" 51 U.S.L.W. 3914 (June 28, 1983).

(3) *Commonwealth v. Sheppard*, 387 Mass. 488, 441 N.E.2d 725 (1982), cert. granted *sub nom.* *Massachusetts v. Sheppard*, 103 S. Ct. 3534 (1983). "Does Fourth Amendment require inflexible application of exclusionary rule when police officer has reasonably and in good faith relied upon search warrant to seize items specified in his application for warrant, but warrant is subsequently invalidated for magistrate's error in failing to specify items in warrant itself?" 51 U.S.L.W. 3913 (June 28, 1983).

Subsequent to the writing of this Article, on December 12, 1983, the Court dismissed the writ of certiorari in *Quintero*. 52 U.S.L.W. 3460 (Dec. 13, 1983). The Court explained that "it appear[ed] that respondent died on November 27, 1983." *Id.* Ironically, *Quintero* could have been the most significant of the three cases since it involved the only search by an officer not acting pursuant to any warrant.

20. Assuming the Supreme Court approves a good faith exception to the federal exclusionary rule, state courts, consistent with the United States Constitution, may interpret their state constitutional provisions as granting broader protections than federal law mandates. Brennan, *State Constitutions and the Protection of Individual Rights*, 90 HARV. L. REV. 489, 502 (1977). See, e.g., *People v. Sporleder*, ___ Colo. ___, 666 P.2d 135 (1983) (warrantless use of a pen register violates the Colorado Constitution even if it is consistent with the federal Constitution under *Smith v. Maryland*, 442 U.S. 735 (1979)); *State v. Alston*, 88 N.J. 211, 440 A.2d 1311 (1981) (New Jersey retains defendant's automatic standing

of the Article briefly sketches the historical development of the exclusionary rule. Part II develops the general arguments against the exclusionary rule and the specific arguments in favor of a good faith exception. Part III exposes the conceptual flaws of the exception by demonstrating its inconsistency with a constitutional concept of right, its disastrous effect on the substance of the fourth amendment, and its interference with a defendant's right to effective counsel. The Article considers practical problems associated with implementation of such an exception in part IV. Finally the Article, in its entirety, demonstrates the hidden agenda behind the movement supporting the good faith exception: minimizing the fourth amendment's significance and substantive protection.

I. THE HISTORICAL CONTEXT

An understanding of the arguments for and against a good faith exception requires some basic familiarity with the evolution of the exclusionary rule and its traditional justifications. In 1914 the Supreme Court decided *Weeks v. United States*,²¹ which prohibited the use in federal courts of evidence obtained in violation of the fourth amendment. The Court stated in defense of this position:

The tendency of those who execute the criminal laws of the country to obtain conviction by means of unlawful seizures . . . should find no sanction in the judgments of the courts which are charged at all times with the support of the Constitution and to which people of all conditions have a right to appeal for the maintenance of such fundamental rights.²²

The Court's concern in *Weeks*, therefore, was not nearly as much with preventing illegal seizures by law enforcement officers as it was both with preventing courts from becoming accomplices in this illegal conduct²³ and with dissipating any impression of judicial imprimatur.²⁴ The Court viewed exclusion of illegally ob-

to invoke the exclusionary rule in search and seizure cases in spite of federal court decisions such as *Rakas v. Illinois*, 439 U.S. 128 (1978)); *People v. Bustamante*, 30 Cal. 3d 88, 634 P.2d 927, 177 Cal. Rptr. 576 (1981) (California suspects have a right to counsel at pre-indictment lineups even if federal suspects have no comparable right under *Kirby v. Illinois*, 406 U.S. 682 (1972)). A recent amendment to the Florida Constitution, however, forbids Florida courts from excluding evidence in those circumstances when similarly obtained evidence would be admissible in federal courts. FLA. CONST. art. 1, § 12.

21. 232 U.S. 383 (1914).

22. *Id.* at 392.

23. The Court used precisely this language in the much more recent case of *Elkins v. United States*, 364 U.S. 206, 223 (1960).

24. See *Terry v. Ohio*, 392 U.S. 1, 13 (1968). William Mertens and Silas Wasserstrom identify four normative principles underlying the *Weeks* exclusionary rule: first, convicting a

tained evidence as necessary to fulfill its responsibility under the fourth amendment. Thus, the Court considered the substantive scope of the fourth amendment and the procedural remedy of the exclusionary rule inseparably intertwined.

Thirty-five years passed before the Court suggested the exclusionary rule was not essential to enforcement of fourth amendment guarantees. In *Wolf v. Colorado*²⁵ the Court allowed the admission in state criminal proceedings of evidence that would have been inadmissible in federal proceedings under the exclusionary rule. In so doing, however, the Supreme Court did recognize that the fourth amendment limitations on governmental searches and seizures bound the states,²⁶ but refused to require exclusion of evidence as a remedy for state violation of the amendment's standards. Consequently, the Court severed the exclusionary rule from its mooring in the substantive amendment and severely restricted the exclusionary rule's range and effect. Therefore, free to develop their own remedies for fourth amendment violations,²⁷ the states and their law enforcement agencies found the Court's fourth amendment rhetoric unobjectionable. Justice Murphy in dissent, however, impliedly recognized that acknowledgement of a right without provision of a specific remedy for violation likely would result in no right at all: "Alternatives are deceptive. Their very statement conveys the impression that one possibility is as effective as the next. In this case their statement is blinding. For there is but one alternative to the rule of exclusion. That is no sanction at all."²⁸

Over time the wedge *Wolf* drove between the fourth amendment and the exclusionary rule became unbearable to the Court.

defendant on the basis of unlawfully seized evidence is simply unfair; second, admission of such evidence at trial constitutes an additional invasion of privacy; third, government is not to be advantaged by the wrongdoing of its agents; and last, admission of unlawfully seized evidence compromises the integrity of the courts. *Mertens & Wasserstrom, supra* note 3, at 377-78.

25. 338 U.S. 25 (1949).

26. *Id.* at 27-28. Because the Bill of Rights as initially conceived applied only to the federal government, *Barron v. Baltimore*, 32 U.S. (7 Peters) 242, 246 (1833), from early times the Court interpreted the fourth amendment as inapplicable to the states. *See Smith v. Maryland*, 59 U.S. (18 Howard) 71, 76 (1855). With the advent of the post-Civil War amendments, however, the Court gradually imposed the Bill of Rights upon the states by incorporation into the due process clause of the fourteenth amendment. Accordingly, *Wolf* held that the security of one's privacy against arbitrary government intrusion was "implicit in 'the concept of ordered liberty' and as such enforceable against the States through the Due Process Clause" of the fourteenth amendment. 338 U.S. at 27-28.

27. *See id.* at 30-33.

28. *Id.* at 41 (Murphy, J., dissenting).

Many states ignored the Court's plea for them to design and implement remedies for fourth amendment violations. Twelve years after *Wolf*, the Court's dissatisfaction with the states' general unresponsiveness culminated in *Mapp v. Ohio*.²⁹ The Justices no longer were willing to accept the existence of a right without an effective remedy for its violation. All evidence obtained by state officials in violation of the fourth amendment, therefore, became inadmissible at trial.³⁰

According to *Mapp*, the exclusionary rule was to perform two functions: first, protection of judicial integrity by avoiding the appearance of impropriety created by permitting use of tainted evidence;³¹ and second, deterrence of illegal police searches by compelling "respect for the constitutional guaranty in the only effectively available way—by removing the incentive to disregard it."³² Protecting judicial integrity was firmly grounded in exclusionary rule history dating back almost fifty years to the *Weeks* opinion. The purpose of deterring illegal police behavior, however, was of more recent vintage. In fact, the deterrence language that the Court approvingly quoted first had been set forth in a decision that the Court had rendered just one year earlier.³³

While the *Mapp* decision failed to indicate which of these functions was weightier, later decisions clearly illustrate the Court's view that police deterrence is the preeminent purpose of the exclusionary rule.³⁴ Although occasionally alluding to the "im-

29. 367 U.S. 643 (1961).

30. *Id.* at 655.

31. *Id.* at 659; see *Elkins v. United States*, 364 U.S. 206, 222-23 (1960).

32. 367 U.S. at 656.

33. See *Elkins v. United States*, 364 U.S. 206, 217 (1960). *Elkins* marked the demise of the so-called "silver platter" doctrine, whereby federal authorities could use in federal prosecutions evidence unlawfully obtained by state police.

34. In *Linkletter v. Walker*, 381 U.S. 618 (1965), the Court refused to give retroactive effect to *Mapp's* extension of the exclusionary rule to the states:

[A]ll of the cases . . . requiring the exclusion of illegal evidence have been based on the necessity for an effective deterrent to illegal police action. . . . We cannot say that this purpose would be advanced by making the rule retrospective. The misconduct of the police . . . has already occurred and will not be corrected by releasing the prisoners involved.

Id. at 636-37. In *United States v. Calandra*, 414 U.S. 338 (1974), the Court held illegally obtained evidence admissible at grand jury proceedings under the theory that forbidding the evidence's use at trial already achieved all possible deterrence. *Id.* at 349-52. But even assuming both *Linkletter* and *Calandra* were correct vis-a-vis deterrence, judicial integrity surely was at stake in both cases. Evidently, the Court considered only police deterrence of sufficient importance to mandate exclusion.

perative of judicial integrity,"³⁵ the Court has stressed the limited role this justification plays in determining whether the rule should be applied in a particular case.³⁶ Thus, the exclusionary rule has taken on an instrumental appearance, warranting application only when its consequence is significantly to enhance deterrence of illegal searches and seizures. From this perspective, before a decision to exclude evidence is justifiable a court must answer the following two questions: First, will the suppression of evidence in a specific case enhance deterrence of illegal police behavior? Second, is the deterrence sufficient to outweigh the damage to other goals of the criminal justice system? The decisional process, hence, becomes utilitarian: simply balancing costs against benefits.³⁷

Accepting this utilitarian approach to exclusionary rule decisions, the Court has refused to apply the rule to those situations in which its deterrent effect would be minimal. Accordingly, prosecutors may use illegally obtained evidence in grand jury proceedings,³⁸ in impeaching a criminal defendant's testimony at trial,³⁹ in civil proceedings brought by a different sovereign than the one that actually seized the evidence,⁴⁰ and in the criminal trial of a defendant who lacks standing to invoke the rule's protection because he was not the actual victim of the improper search.⁴¹ Furthermore, a state prisoner no longer may seek federal habeas corpus relief on fourth amendment grounds if he had a "full and fair opportunity" to litigate the issue at the state level because, at least in the Supreme Court's view, the "additional incremental deterrent effect" of exclusion at the habeas stage is insignificant.⁴²

35. See, e.g., *Terry v. Ohio*, 392 U.S. 1, 12 (1968).

36. See *Stone v. Powell*, 428 U.S. 465, 485 (1976). Some scholars now view deterrence as essentially the "sole theory" behind the exclusionary rule. E.g., Wilkey, *supra* note 3, at 220.

37. Utilitarian decisional modes are arguably inappropriate when dealing with constitutional rights rather than with societal gratuities. See *infra* text accompanying notes 103-07.

38. See *United States v. Calandra*, 414 U.S. 338 (1974).

39. See *United States v. Havens*, 446 U.S. 620 (1980).

40. See *United States v. Janis*, 428 U.S. 433 (1976).

41. See *United States v. Payner*, 447 U.S. 727 (1980).

42. See *Stone v. Powell*, 428 U.S. 465, 494 (1976). This consequentialist perspective has become so prevalent in the exclusionary rule context that some of the rule's opponents claim it now influences judicial analysis of the fourth amendment's substantive rights. Professor Steven Schlesinger, for example, insists that the exclusionary rule forces the judiciary to expand dangerously the meaning of a legal search in order to avoid exclusion of evidence the judges are reluctant to suppress. Schlesinger, *supra* note 3, at 405. Professor Schlesinger's perspective may explain the courts' increasing number of refusals to suppress evidence by claiming no, or only a limited, legitimate expectation of privacy existed. See

The pervasive deterrent justification for the exclusionary rule, however, makes it highly vulnerable to attack. Whether suppression of improperly obtained evidence actually deters illegal police behavior is subject to question.⁴³ With its deterrent benefit in doubt then, opponents of the rule argue its costs are too high and its benefits too low. The rule's detractors now use this cost benefit argument to justify a good faith exception to the exclusionary rule.

II. THE CASE FOR A GOOD FAITH EXCEPTION

Advocates of a good faith exception first focus on the general costs of any exclusionary rule application. They stress that physical evidence is reliable whether or not illegally obtained. The public, consequently, perceives excluding such reliable evidence as merely freeing countless guilty defendants⁴⁴ who then are able to commit further offenses.⁴⁵ The rule further impairs the truth-find-

United States v. Payner, 447 U.S. 727 (1980) (bank depositor has no fourth amendment privacy interest in copies of bank checks and deposit slips retained by bank); Smith v. Maryland, 442 U.S. 735 (1979) (no expectation of privacy in phone numbers dialed); Chambers v. Maroney, 399 U.S. 42 (1970) (the expectation of privacy in automobiles is sufficiently reduced to justify a search with probable cause but without warrant).

43. As Professor Francis Allen observed, "the case for the rule as an effective deterrent of police misbehavior has proved, at best, to be an uneasy one." Allen, *The Judicial Quest for Penal Justice: The Warren Court and the Criminal Cases*, 1975 U. ILL. L.F. 518, 537.

44. The fourth amendment exclusionary rule, therefore, may have significantly more harmful effects than the exclusion of incriminating statements obtained in violation of fifth amendment guarantees. Police tactics violating the fifth amendment purportedly threaten the reliability of the verbal evidence so gathered. In contrast, little doubt exists about the reliability of physical evidence obtained in violation of the fourth amendment. Furthermore, confessions frequently only serve to cement a conviction. See *Interrogations in New Haven: The Impact of Miranda*, 76 YALE L.J. 1519, 1588 (1967). In contrast, judicial suppression of evidence obtained by illegal police searches actually may be determinative in some prosecutions, especially those dealing with consensual offenses. See Oaks, *supra* note 3, at 746.

45. See *Bivens v. Six Unknown Fed. Narcotics Agents*, 403 U.S. 388, 413-16 (1971) (Burger, C.J., dissenting); Wright, *Must the Criminal Go Free If The Constable Blunders?*, 50 TEX. L. REV. 736, 741 (1972). Fourth amendment decisions, unlike decisions concerning self-incrimination, are not directed predominately toward police treatment of the poor and ignorant. See *Miranda v. Arizona*, 384 U.S. 436, 505 (1966) (Harlan, J., dissenting). Rather, their greatest impact is upon police investigations of activities frequently controlled by organized crime. Physical evidence is most essential in proving the offenses of gambling, prostitution, counterfeiting, and narcotics distribution. See Kaplan, *supra* note 3, at 1028; Oaks, *supra* note 3, at 681-89. Furthermore, the consensual nature of many of these offenses, coupled with the fear organized crime can inflict on its members and victims, makes police fulfillment of the traditional demands of probable cause before search unusually difficult. Trickery and deception, often perceived by law enforcement officials as necessary to overcome evidentiary difficulties in dealing with sophisticated crime, are the very seeds from which fourth amendment problems are born. Former New York City Police Commissioner Vincent L. Broderick stated:

The narcotics agent—local or federal—who seeks evidence against organized criminals

ing function of the trial proceeding, exclusionary rule opponents argue, because the rule's application diverts the trial's focus from ascertaining the guilt or innocence of the defendant to examining police behavior.⁴⁶ Additionally, critics assail the suppression remedy as protecting only those against whom police discover incriminating evidence while providing no protection for innocent persons subjected to illegal but fruitless searches.⁴⁷ Its opponents, therefore, insist that each application of the exclusionary rule exacts a substantial societal cost for the vindication of fourth amendment rights.⁴⁸

The exclusionary rule's indifference to the seriousness of the defendant's alleged offense or the severity of the police officer's infraction accentuates the appearance of excessive cost. Critics attack as "contrary to the idea of proportionality that is essential to the concept of justice"⁴⁹ the universal "capital punishment" that the rule inflicts on all improperly acquired evidence regardless of apparent disparities which at times can exist between the officer's error and the degree of a defendant's likely guilt. These critics, fur-

necessarily operates undercover, on the street, masquerading as someone he is not. Similarly the Secret Service agent attempting to capture counterfeiters assumes the identity of an underworld purchaser of counterfeit bills. Laws against consensual crimes, such as gambling and vice, can be enforced effectively only if police officers in plainclothes conceal their true identities and adopt other ones. Deception and trickery? Certainly. How else can we cope with sophisticated crime?

Broderick, Book Review, 53 CORNELL L. REV. 737, 741 (1968); see also *A Forum on the Interrogation of the Accused*, 49 CORNELL L. REV. 382, 394 (1964). Clearly, then, exclusion of illegally seized physical evidence can impede law enforcement efforts to convict the guilty more severely than does exclusion of illegally obtained confessions.

46. Oaks, *supra* note 3, at 755.

47. See *Irvin v. California*, 347 U.S. 128, 136 (1954); Oaks, *supra* note 3, at 755.

48. *Rakas v. Illinois*, 439 U.S. 128, 137 (1978). See *United States v. Payner*, 447 U.S. 727, 734 (1980). The cost of losing inherently trustworthy evidence may tempt courts either to water down the standards for probable cause, Barrett, *Personal Rights, Property Rights, and the Fourth Amendment*, 1960 SUP. CT. REV. 46, 66, or to draw unrealistic distinctions in favor of law enforcement, Ashdown, *supra* note 3, at 1315, in order to avoid the exclusionary effects of a potential fourth amendment violation. Justice White, in his *Gates* concurrence, suggested the monetary cost of exclusion also should be considered. Citing the Comptroller General's study of the impact of the exclusionary rule on federal prosecutions, White noted that one-third of federal defendants going to trial file fourth amendment suppression motions; 70% to 90% of such motions included the expense of formal hearings. 103 S. Ct. at 2343 (White, J., concurring) (citing COMPTROLLER GENERAL OF THE UNITED STATES, *IMPACT OF THE EXCLUSIONARY RULE ON FEDERAL CRIMINAL PROSECUTIONS* 10 (1979)).

49. *Stone v. Powell*, 428 U.S. 465, 490 (1976). "[S]ociety has . . . [a] right to expect rationally graded responses from judges in place of the universal 'capital punishment' we inflict on all evidence when police error is shown in its acquisition." *Bivens v. Six Unknown Fed. Narcotics Agents*, 403 U.S. 388, 419 (1971) (Burger, C.J., dissenting). The sense of the criminal defendant's windfall, of course, is greatest when the officer allegedly made the error in good faith.

thermore, herald the rejection by other nations of a comparable rule excluding evidence illegally obtained⁵⁰ as proof that the doctrine is both unwise and unnecessary. Thus, although the exclusionary rule's supporters defend it as a deterrent to unlawful police activity in part through the nurturing of respect for fourth amendment values,⁵¹ others fear that the public perception of a judicial system easy on criminals and unresponsive to the cries for greater security does considerable damage to the reputation of the criminal justice system and generates disrespect for the law generally.⁵²

If the exclusionary rule is nothing but a judicially created device chosen from among other devices to enforce fourth amendment guarantees⁵³ and if this device truly has led to widespread public disfavor with the criminal justice system,⁵⁴ then judicial

50. England, Canada, Germany, and Israel all have refused to exclude evidence obtained through illegal police searches and seizures. See *Bivens*, 403 U.S. at 415 (Burger, C.J., dissenting); Wilkey, *supra* note 3, at 217.

51. The *Mapp* court, for example, justified the exclusionary rule partially upon the need for legal institutions to educate the public through example. "If the government becomes a law breaker, it breeds contempt for law; it invites every man to become a law unto himself; it invites anarchy." 367 U.S. at 659.

52. See *Stone v. Powell*, 428 U.S. 465, 491 (1976); Paulsen, *The Exclusionary Rule and Misconduct by the Police*, 52 J. CRIM. L. CRIMINOLOGY & POLICE SCI. 255, 256 (1961); Wilkey, *supra* note 3, at 223; Wright, *supra* note 45, at 737.

Not only the public may react negatively to such a situation, but some jurists may as well, see *Bivens v. Six Unknown Fed. Narcotics Agents*, 403 U.S. 388, 414 (1971) (Burger, C.J., dissenting), and undoubtedly law enforcement officers also lose respect for a system they perceive as forcing the government to play the game fairly despite the criminal opponent's ruthlessness.

53. The fourth amendment, unlike the fifth, does not explicitly mandate exclusion of evidence as the proper response for violation of its guarantees. This textual distinction between the amendments, coupled with court decisions justifying fourth amendment exclusion on the basis of its deterrent value, has led many critics to view the fourth amendment's exclusionary rule as a judicially created device rather than as an integral aspect of the constitutional right. See *Stone v. Powell*, 428 U.S. 465, 482 (1976); Wilkey, *supra* note 3, at 216; Crovitz, *Exclusionary Rule Lives in a World Apart*, Wall St. J., June 30, 1983, at 28, col. 3. Clearly, this was not the *Weeks* Court's perspective of the importance of the rule in guaranteeing fourth amendment rights. See *supra* text accompanying notes 21-24.

54. The extent and depth of public disapproval in some quarters is clearly apparent in newspaper editorials. See, e.g., *Houston Post*, Editorial, Nov. 16, 1977, at 2E (The results of the exclusionary rule lead to "an absurdity that would likely be sensibly ordered in a more primitive society."); *Washington Star*, Editorial, July 7, 1975, at A16 (The exclusionary rule "is sometimes an outrage to common sense."); *Wall St. J.*, Editorial, July 12, 1971, at 8, col. 1 ("Given the . . . need to bolster public confidence that courts do dispense justice, it's scarcely unreasonable to ask that [the fourth amendment's exclusionary rule] be reexamined.") See also Crovitz, *supra* note 53, at 28, col. 6 ("The present U.S. exclusionary rule does substantive injustice.")

Professor John Kaplan reports that public opinion polls show a high rate of disapproval for courts seen by the public as coddling criminals. He identifies the exclusionary rule as the prototype judicial doctrine causing this perception. Kaplan, *supra* note 3, at 1035-36.

concern over the rule is justified.⁵⁵ The legitimacy of law exists on a foundation of faith. The public must perceive its views of "good" and "right" as prevailing rather than being circumvented by institutional machinations or it may lose its belief in the ultimate moral nature of law. The loss of such a belief can lead to spiritually unstable institutions and a general weakening of governmental legitimacy.⁵⁶

Societal faith thus is essential for effective functioning of any legal system. The criminal justice system, in addition to controlling unacceptable deviance, also serves to control the response by societal enforcers to such deviance.⁵⁷ Although vendetta and mob rule could serve to control deviance, only a criminal justice system that also controls the controllers can enhance human dignity. A justice system that limits the enforcers, however, will be able to promote human dignity only if the system fulfills public expectations of societal protection sufficiently to remain legitimate. When legal institutions so limit the enforcers that the community believes its needs for safety and justice are unfulfilled, however, the system may lose its legitimacy and thus its power to control societal responses to deviance. Excessive judicial concern for individual rights in the face of a public outcry for greater safety, consequently, may lead to a popular backlash decreasing rather than enhancing respect and conformity to enforcement limitations.⁵⁸

If utilization of the exclusionary rule leads to lost system legitimacy, the cost is high indeed.⁵⁹ But detractors of the exclusionary rule do not rest their position on costs alone; they also stress the

55. As Chief Justice (then Judge) Warren Burger remarked, "If a majority—or even a substantial minority—of the people in any given community . . . come to believe that law enforcement is being frustrated by what laymen call 'technicalities,' there develops a sour and bitter feeling that is . . . sociologically unhealthy." Burger, *Who will Watch the Watchman?*, 14 AM. U.L. REV. 1, 22 (1964).

56. For a fuller discussion of the legitimation of law, see Ingber, *The Interface of Myth and Practice in Law*, 34 VAND. L. REV. 309, 338-46 (1981).

57. This dual purpose of criminal law is developed in Ingber, *A Dialectic: The Fulfillment and Decrease of Passion in Criminal Law*, 28 RUTG. L. REV. 861, 863-68 (1975).

58. Fear of such a backlash was the basis of Justice White's dissent in *Miranda v. Arizona*, 384 U.S. 436 (1966). He expressed concern that the call for human dignity in the name of the defendant had gone too far. If governmental authority was unable to protect the public from crime because of judicially imposed limitations on law enforcement, White feared that "those who rely on the public authority for protection" might revert to the savage techniques generally associated with self-help. *Id.* at 542 (White, J., dissenting). He further asserted that "[w]ithout the reasonably effective performance of the task of preventing private violence and retaliation, it is idle to talk about human dignity and civilized values." *Id.* at 539.

59. *But see infra* text accompanying notes 309-16.

limited benefits of the rule. As developed earlier,⁶⁰ the main articulated and recognized benefit of the exclusionary rule is deterrence of illegal police behavior. Scholars⁶¹ and Supreme Court Justices,⁶² however, have expressed grave doubts on whether the exclusionary rule actually influences police behavior. Statistics do not support the rule's efficiency in deterring police misbehavior since criminal defendant's make many motions to suppress illegally obtained evidence each year. The abundance of such motions, however, also does not support the conclusion that the rule is ineffective. No way exists to determine whether many more fourth amendment violations would occur without the rule.⁶³ The statistical battle, therefore, often turns upon who has the burden of proof: the rule's proponents to verify its deterrent value or its opponents to verify its lack of such value.⁶⁴ Despite the lack of statistical proof either for or against the rule's deterrent value, at least three factors justify some intuitive doubt as to the rule's ability to influence police behavior: ideology of the police subculture, police goals other than criminal convictions, and lack of meaningful communications between the courts and police.

Police gain their identity from their image as crime fighters.⁶⁵

60. See *supra* text accompanying notes 34-37.

61. See, e.g., J. SKOLNICK, JUSTICE WITHOUT TRIAL 211-19 (1966); Dworkin, *Fact Style Adjudication and the Fourth Amendment: The Limits of Lawyering*, 48 IND. L.J. 329 (1973); Oaks, *supra* note 3; Wilkey, *supra* note 3.

62. See, e.g., *United States v. Calandra*, 414 U.S. 338, 348 n.5 (1974) (Powell, J.) (doubt as to effectiveness of the exclusionary rule); *Bivens v. Six Unknown Fed. Narcotics Agents*, 403 U.S. 388, 415-16 (1971) (Burger, C.J., dissenting) ("the hope that [the objective of control of official activity] could be accomplished by the exclusion of reliable evidence from criminal trials was hardly more than a wistful dream . . . [and] the history of the suppression doctrine demonstrates that it is both conceptually sterile and practically ineffective in accomplishing its stated objective").

63. See Allen, *Federalism and the Fourth Amendment: A Requiem for Wolf*, 1961 SUP. CT. REV. 1, 34.

64. The burden of proof issue, always implicit in the exclusionary rule deterrence controversy, recently has become an explicit focus of the conflict. Compare Canon, *The Exclusionary Rule: Have Critics Proven that It Doesn't Deter Police?*, 62 JUDICATURE 398 (1979) with Schlesinger, *supra* note 3.

65. The popular view of the police officer in this society, at least of the urban police officer that predominately has influenced the constitutional developments in criminal investigatory procedures, is one of a committed crime fighter protecting society. This image, largely based on the false belief that police devote most of their time to crime prevention, influences both who chooses to join the police force and the self-image of those on the force. See Goldstein, *Police Policy Formulation: A Proposal for Improving Police Performance*, 65 MICH. L. REV. 1123, 1124-25 (1967). Thus, although police spend most of their time dealing with family squabbles, noisy disturbances, and requests for various forms of services and information, law enforcement officials gear police agencies primarily to deal with crime, see Goldstein, *Police Response to Urban Crisis*, 28 PUBL. AD. REV. 417, 418 (1968), and police

The police force trains its officers to perceive themselves as better equipped for crime fighting work than anyone else. This perceived expertise is a source of status from which police hope to gain societal recognition for their "professional" position.⁶⁶ Police, therefore, tend to resent any interference by outsiders that reflects upon their professionalism or inhibits performance of their professional responsibility to control crime. Consequently, police officers are not predisposed to accept and accommodate the court imposed exclusionary rule.⁶⁷

Police officers' increasing isolation from the rest of the community accentuates their lack of appreciation for the rule. While police seek public admiration and support, their working personality which emphasizes danger and authority estranges them from the conventional citizenry⁶⁸ as well as from those who represent potential threats.⁶⁹ Many civilians are uncomfortable about including police officers among close acquaintances.⁷⁰ The extreme moralistic nature of the criminal law that police must enforce only exacerbates this civilian concern about socializing with police officers. Police, in turn, perceive themselves as representing authority.

see crime fighting as the "real" police work, see J. SKOLNICK, *supra* note 61, at 220.

66. Professionalization brings a sense of individual worth through identification with a group publicly recognized as better than others at performing a specific job.

67. "[The police officer's] perception of himself as a crime-fighting craftsman," notes Professor Albert Quick, "is outwardly manifested by a general hostility toward concepts of procedural due process and those institutions that are identified with securing individual rights—the courts and the liberal elements of society." Quick, *Attitudinal Aspects of Police Compliance with Procedural Due Process*, 6 AM. J. CRIM. L. 25, 26 (1978).

68. See J. SKOLNICK, *supra* note 61, at 44.

69. According to Professor Jerome Skolnick,

The element of danger seems to make the policeman especially attentive to signs indicating a potential for violence and lawbreaking. As a result, the policeman is generally a "suspicious" person.

. . . .

. . . [B]ecause his work requires him to be occupied continually with potential violence, [he] develops a perceptual shorthand to identify certain kinds of people as symbolic assailants, that is, as persons who use gesture, language, and attire that the policeman has come to recognize as a prelude to violence.

Id. at 44-45. See also Piliavin & Briar, *Police Encounters with Juveniles*, 70 AM. J. SOC. 206 (1964). Often having to make instantaneous decisions, a police officer is likely to form and to use stereotypes to facilitate his decisionmaking. Police may watch blacks, the poor, and those acting in any way out of the ordinary more warily, not necessarily out of any prejudice, but because these groups represent higher statistical risks of danger to society. See Balch, *The Police Personality: Fact or Fiction?*, 63 J. CRIM. L. CRIMINOLOGY & POLICE SCI. 106, 111 (1972).

70. Civilians' knowledge that police officially are always on duty and frequently carry guns generally inhibits civilians' speech and actions when in police presence. See Hahn, *A Profile of Urban Police*, 36 LAW & CONTEMP. PROBS. 449, 453 (1971).

Thus, they feel unable to shed this pose and act naturally except among other officers and their families. A tendency, therefore, exists for police officers to nearly totally restrict both their professional and social relationships to other members of the police force.⁷¹ This alienation from society and dependence upon other officers for needed support and approval results in the creation of a police subculture, which provides a basis for self-respect somewhat independent of civilian norms.⁷²

The prevalent attitudes of this police subculture, consequently, far more than any abstract judicial articulation of right and wrong, determine the pressure police experience to abide by court imposed restrictions on their crime fighting ability.⁷³ Officers, "engaged in the often competitive enterprise of ferreting out crime,"⁷⁴ may gain peer support or even reward for behavior that the courts would not consider legal.⁷⁵ The need to respond situationally either to a perceived danger to others or to a threat against his own authority may provoke an officer into action that courts do not approve.⁷⁶ As long as the officer strongly identifies with the police subculture, which approves of his behavior, he has immediate and certain reinforcement from his primary peer group. The mere possibility that sometime in the future, at trial or on appeal, some aversive contingency may follow is significantly less effective in controlling his future behavior.⁷⁷ In fact, given the ex-

71. See M. BANTON, *THE POLICEMAN IN THE COMMUNITY* 188-219 (1964).

72. Lipset, *Why Cops Hate Liberals—and Vice Versa*, *THE ATLANTIC*, Mar. 1969, at 80.

73. The existence of a powerful police subculture results in police officers potentially lacking awareness of the offsetting value positions possessed by individuals who identify with numerous groups. Strong group cohesion among officers reinforces their attitudes and limits needed appreciation and status to that given by the subculture. See J. SKOLNICK, *supra* note 61, at 219. Society's evaluation becomes far less important to a police officer than his peer group evaluation. Consequently, the impact of nonpolice criticism is minimal in affecting police esteem and, thus, behavior.

74. *Johnson v. United States*, 333 U.S. 10, 14 (1948).

75. For example, valued goals of the police subculture include return of stolen goods or confiscation of contraband. See *infra* text accompanying notes 82-83. Furthermore, judges are rule oriented; police are situation oriented. The police officer "approaches incidents that threaten order not in terms of enforcing the law but in terms of 'handling the situation.'" J. WILSON, *VARIETIES OF POLICE BEHAVIOR* 31 (1968) (emphasis omitted). Consequently, police view many strategies that would be illegal from a judicial standpoint as totally justifiable.

76. Chief Justice Warren recognized the inability of the exclusionary rule to deter police behavior in circumstances that the police officer finds threatening. See *Terry v. Ohio*, 392 U.S. 1, 13-14 (1968).

77. Chief Justice Burger, in a vigorous dissent assailing the exclusionary rule, noted that "the long time lapse—often several years—between the original police action and its

clusionary rule's focus on trial proceedings, it arguably has more effect on the prosecution than on the offending police officer.⁷⁸ Whether the prosecutor's inability to gain a conviction will influence the police depends on whether police interests encompass concerns other than successful prosecution.

A review of institutional incentives and pressures affecting police, however, shows that numerous purposes other than obtaining convictions at trial⁷⁹ may induce officers to use searches and arrest.⁸⁰ For example, law enforcement agencies often rely upon the percentage of reported cases "cleared by arrest" as a measure of police alertness and professionalism.⁸¹ Since an arrest need not result in a conviction to affect favorably the "clearance rate," this institutionalized sign of police success subtly encourages illegal searches and arrests.

The public itself sometimes encourages illegal searches and seizures by vehemently demanding more aggressive law enforcement.⁸² In an attempt to gain broader societal acceptance while still fulfilling their subculture's primary objective of fighting crime,

final judicial evaluation" dilutes any "deterrent impact" on police of evidence suppressions. *Bivens v. Six Unknown Fed. Narcotics Agents*, 403 U.S. 388, 417-18 (1971).

78. The Court may be incapable of using evidentiary decisions as a means of controlling pretrial police behavior. See *Amsterdam*, *supra* note 3, at 350-51.

79. Because a knowing and intelligent guilty plea does not require the defendant's awareness of evidentiary defenses, *Edwards v. United States*, 256 F.2d 707, 710 (D.C. Cir.), *cert. denied*, 358 U.S. 847 (1958), and because all prior police behavior may be invulnerable to attack once defendant enters such a plea, see *Tollett v. Henderson*, 411 U.S. 258 (1973); *McMann v. Richardson*, 397 U.S. 759 (1970), the incriminating evidence illegally gained may sufficiently impress a defendant that he will plead guilty and avoid trial. Illegal searches, consequently, still may aid in gaining convictions.

80. See *Dworkin*, *supra* note 61, at 330; *Oaks* *supra* note 3, at 721-22; see also *Terry v. Ohio*, 392 U.S. 1, 14 (1968) (exclusionary rule is powerless when police motive is other than successful prosecution).

81. See J. GRIFFIN, *STATISTICS ESSENTIAL FOR POLICE EFFICIENCY* 69 (1958); J. SKOLNICK, *supra* note 61, at 167. Bureaucratic pressures can significantly affect police attitudes. The factors used to evaluate merit and the hierarchy of roles all act as a system of incentives for the officer intent on advancing within the force. The police bureaucracy places value on obedience and efficiency. "Real" police work—work that gains both peer and official acknowledgement—involves catching criminals. An officer rarely gets credit for *not* arresting, searching, or questioning. In addition, "it is relatively easier to measure output reflected by the number of people arrested and the number of crimes solved than it is to measure, for example, the number of people whose rights were protected. In a sense measurement determines goals." Milner, *Supreme Court Effectiveness and the Police Organization*, 36 *LAW & CONTEMP. PROBS.* 467, 472 (1971) (footnote omitted). See also Newman, *The Effect of Accommodations in Justice Administration on Criminal Statistics*, 46 *Soc. & Soc. RESEARCH* 144 (1962).

82. In times of public danger there rarely exists any community demand for increased police self-control.

police frequently respond to such public pressure by making arrests as efficiently and quickly as possible. While such arrests may not gain judicial acceptance, they certainly improve the police's public image. Community dissatisfaction, consequently, shifts from what the public could have perceived as ineffective law enforcement to what now it perceives as an overly technical court. Public support and peer approval also can encourage illegal police behavior if the officer, through an illegal search or seizure, can perform a socially valuable service such as rescuing a kidnap victim, returning stolen goods to their rightful owner, or removing narcotics from the streets. Often police find lost convictions preferable to lost opportunities to fulfill these goals.⁸³

Police also believe that substantially increasing the expenses attendant to criminal activity aids in their crime fighting responsibility. Confiscation of criminal paraphernalia, the necessity for repeated bail payments, and the inconvenience and time loss that the criminal experiences in arrest procedures all serve to cripple criminal activity whether or not the corresponding searches and seizures ultimately result in conviction. Incarceration prior to trial can remove a danger⁸⁴ or public nuisance⁸⁵ from the streets, sometimes dissipating the problem despite the lack of a permanent conviction. Police may view arrest of a delinquent juvenile or a petty hoodlum as necessary to reestablish official authority and reinforce respect for the law. Additionally, arrest and search intimidation greatly facilitates recruitment of informers, allegedly so vital in fighting organized crime, regardless of any police intent to press for prosecution.

Accordingly, the police have many alternative goals to conviction that, at times, serve to encourage illegal behavior.⁸⁶ When one or more of these goals is present in a situation, the exclusionary

83. As the importance of the "pinch" becomes greater, the deterrence value of the exclusionary rule correspondingly increases. Ironically, the public most questions the exclusionary rule precisely when the crime is serious and the rule most likely will deter illegal police behavior. Professor Kaplan, for example, has recommended exempting police misbehavior relating to serious crimes from the rule's coverage. Kaplan, *supra* note 3, at 1046-49.

84. Fearing flare-ups of teenage gang rivalries, for example, police may incarcerate gang leaders until passions have calmed.

85. For example, police often pick up prostitutes and drunks simply to remove them from the streets.

86. Some critics even have suggested that the exclusionary rule allows corrupt officers to appear to attempt zealous enforcement of the law while, through use of illegal searches, immunizing from effective prosecution those criminals willing to pay the price. Dash, *Cracks in the Foundation of Criminal Justice*, 45 ILL. L. REV. 385, 391-92 (1951); Wilkey, *supra* note 3, at 226.

rule's deterrent value likely is minimal. Furthermore, if the police view the fourth amendment standards as impossible to fulfill, these alternative goals become of primary rather than secondary importance. Difficulties in the court's communications with the police make the amendment's standards at least difficult to understand, if not impossible to fulfill.

Even assuming police desire to gain a conviction in a given case, they must still know and understand what the court requires of them before the exclusionary rule can deter their illegal behavior. The case-by-case method of constitutional adjudication in search and seizure matters often creates decisions no broader in application than the case at hand.⁸⁷ Additionally, court opinions, striving to gain and retain a court majority, are rarely models of clarity.⁸⁸ Issues that badly divide the Supreme Court, fostering strong dissents, further confuse police understanding of what the Court requires of them.⁸⁹ Police, therefore, may receive little guidance for their behavior in the diverse situations that may confront them.⁹⁰ Consequently, the deterrent benefit of the rule once again is put in question.⁹¹

87. *E.g.*, *United States v. Place*, 103 S. Ct. 2637 (1983); *Florida v. Royer*, 103 S. Ct. 1319 (1983); *United States v. Mendenhall*, 446 U.S. 544 (1980); *Schneekloth v. Bustamonte*, 412 U.S. 218 (1973).

88. Often a dissenting opinion in a Supreme Court decision surpasses the Court's opinion in logic and clarity. This occurs because the dissenter does not need to appease his colleagues to retain a majority or worry about how his opinion actually will affect the law's practical functioning. The majority opinion, on the other hand, must satisfy all the joining Justices and must consider carefully whether the decision will obtain the public support required for legitimacy and peaceful compliance.

89. Prosecutors trying to educate police about court cases face these same obstacles.

90. One scholar argues that the fact orientation of appellate court decisions is the greatest impediment to control of police. Dworkin, *supra* note 61, at 334. Another author recommends that courts draw "bright lines" formulating a clear set of rules that, in most instances, will allow police to quickly reach a correct determination of whether a situation constitutionally justifies an invasion of privacy. LaFave, *supra* note 3, at 320-24. The Supreme Court, in recent years, has made some steps in this direction. *See United States v. Ross*, 456 U.S. 798 (1982); *New York v. Belton*, 453 U.S. 454 (1981).

91. Attempting to limit the costs of the exclusionary rule, the Supreme Court has created the exceptions discussed *supra* in text accompanying notes 38-42. Ironically, these attempts to limit the rule's costs likely have the corresponding impact of reducing its deterrent benefit. For example, because motions to exclude illegally obtained evidence only may be made by a person with the requisite standing, *Alderman v. United States*, 394 U.S. 165, 174 (1969), such evidence may have great trial utility. The requisite standing exists only when police violate that particular defendant's legitimate expectations of privacy. *Rawlings v. Kentucky*, 448 U.S. 98 (1980); *Rakas v. Illinois*, 439 U.S. 128 (1978). This standing requirement has encouraged illegal searches, especially illegal wiretapping and electronic eavesdropping. These devices are most useful in attacking organized crime. For example, the information obtained by police through an illegal wiretap may not be used against the owner

The above arguments suggest that from a utilitarian perspective the exclusionary rule has excessive costs and limited benefits. Opponents, thus, advocate total abandonment of the rule,⁹² or at least, through judicial recognition of a good faith exception, further restriction of the rule's effects in situations in which its benefits are least likely. The proponents of a good faith exception base their arguments on reason and precedent. If courts exclude evidence in order to deter unlawful police behavior, exclusion arguably has no value when the officer is honestly and reasonably unaware that his behavior is illegal. An officer acting in a manner he believes consistent with fourth amendment mandates has no reason to fear exclusion and, thus, no motivation to alter his conduct.⁹³ In such circumstances, good faith exception proponents note, exclusion retains its high cost without even the semblance of any deterrent benefit gains. Furthermore, releasing a hardened criminal because of insufficient evidence while "punishing" a diligent police officer for making a reasonable good faith error in a complex area of law seems manifestly unjust.

The Fifth Circuit has extended this logical position. In addition to having no deterrent benefit, argued the court, exclusion of evidence obtained by a good faith error affects police behavior in an undesirable manner. The court insisted that an officer, punished for behavior he honestly believed was legal, would eventually become reluctant to act when his "proper and reasonable instinct tells him that the activity he observes is criminal."⁹⁴ The above utilitarian perspective, thus, provides proponents of a good faith

of the phone or against those participating in the overheard conversation. *Alderman v. United States*, 394 U.S. 165, 178-79 (1969). For police willing to trade the lower echelon of the criminal world for the higher-ups, however, illegal wiretaps can produce evidence admissible against anyone other than those owning the phone or participating in the discussions because they lack standing to object. *Id.* at 171-72, 174. See also *United States v. Payner*, 447 U.S. 727 (1980). Since fighting organized crime is an important police goal, the standing requirement depletes the rule of much of its deterrent value. While the exclusionary rule's critics approve of the standing requirement because it reduces the rule's costs, they fail to note that it is a primary cause of their second objection to the rule, the limited deterrent value of exclusion.

92. Even staunch critics of the exclusionary rule have sometimes expressed concern over the negative symbolic effect abandoning the rule prior to the development of some meaningful substitute might have upon fourth amendment interests. *Bivins v. Six Unknown Fed. Narcotics Agents*, 403 U.S. 388, 415 (1971) (Burger, C.J., dissenting); Kaplan, *supra* note 3, at 1055; Oaks, *supra* note 3, at 756.

93. Kaplan, *supra* note 3, at 1044.

94. *United States v. Williams*, 622 F.2d 830, 842 (5th Cir. 1980) (en banc) (quoting *United States v. Williams*, 594 F.2d 86, 97-98 (5th Cir. 1979)), *cert. denied*, 449 U.S. 1127 (1981).

exception with a logical and initially appealing argument.

In addition to this utilitarian logic, exception proponents offer precedent to demonstrate that a good faith exception is not inconsistent with fourth amendment doctrine but rather is simply the fruition of a developing trend. The two cases most frequently mentioned by those advocating the exception are *United States v. Peltier*⁹⁵ and *Michigan v. DeFillippo*.⁹⁶ In *Peltier* the Supreme Court refused to retroactively apply its holding in *Almeida-Sanchez v. United States*.⁹⁷ In *Almeida-Sanchez* a roving border patrol seeking illegal aliens discovered contraband during a vehicle search. Neither probable cause nor consent existed for the search; the car was simply in the general vicinity of the Mexican border. The Court found the evidence seized inadmissible. The search in *Peltier* predated the *Almeida-Sanchez* opinion. In finding for the government the Court used language the meaning of which, some insist, extends far beyond the parameters of a retroactivity controversy:

If the purpose of the exclusionary rule is to deter unlawful police conduct[,] then evidence obtained from a search should be suppressed only if it can be said that the law enforcement officer had *knowledge, or may properly be charged with knowledge*, that the search was unconstitutional under the Fourth Amendment.⁹⁸

Such language, claim those supporting an exception, signals the Court's disapproval of applying the exclusionary rule when police act both reasonably and in good faith.

Michigan v. DeFillippo concerns a somewhat more complex factual situation than *Peltier*. Police arrested defendant for violating a Detroit ordinance requiring a person lawfully stopped by police to produce evidence of his identity. During the search incident to this arrest police found drugs on defendant's person. The Michigan Court of Appeals, finding the ordinance under which the arrest occurred unconstitutionally vague, believed that both the related arrest and search, thus, were invalid. The state court, therefore, overturned the defendant's drug offense conviction. The Supreme Court reversed.

The Court accepted the state court finding that the ordinance under which the police made the initial arrest was unconstitutional. The Court, however, held constitutional the police officer's

95. 422 U.S. 531 (1975).

96. 443 U.S. 31 (1979).

97. 413 U.S. 266 (1973).

98. 422 U.S. at 542 (emphasis added).

behavior and the resulting drug conviction. In upholding the search, the Court insisted that it should not require a prudent police officer to anticipate that a court subsequently will find invalid the ordinance under which he is making an arrest and incidental search.⁹⁹

Exception proponents use *DeFillippo* as precedential support for a good faith exception under the following reasoning: If courts cannot require a prudent officer to anticipate the ultimate constitutionality of an ordinance before a search can be valid, how can the courts require such an officer to anticipate a change in the Supreme Court's interpretation of fourth amendment guarantees? Exception proponents argue it is equally, if not more, inappropriate to punish an officer for resolving a situational crisis in a manner consistent with that recommended by many of the most deliberative legal minds merely because five of the nine sitting Supreme Court Justices reject his actions. The proponents, thus, argue that a general reasonable good faith exception is consistent with the spirit and logic of such precedent.

This Article has now described the historical development of the exclusionary rule, presented the arguments opposing the rule, and presented the logic and precedent supporting a good faith exception. The Article now shall prove that recognition of a good faith exception would be at best inadvisable; at worst, disastrous.

III. THE CONCEPTUAL FLAWS OF THE GOOD FAITH EXCEPTION

A good faith exception to the exclusionary rule has a number of serious conceptual flaws. The exception would thwart any vindication of recognized fourth amendment rights, and thereby deprecate the constitutional provision itself. Furthermore, the concept of deterrence, which detractors use to attack the exclusionary rule and thus to support the good faith exception, is improperly narrow. Most significantly, adoption of the exception would affect detrimentally the substance and development both of the fourth amendment and of the sixth amendment right to counsel.

A. *The Vindication of Right*

1. The Concept of Right

At the most conceptual level, the utilitarian perspective surrounding and justifying the exclusionary rule generally and the

99. 443 U.S. at 37-38.

good faith exception specifically is misdirected. For the concept of right in this constitutional system to have any significance, it must embody more than an arithmetic computation of costs and benefits. In a democratic system of government, dominant viewpoints of established groups need no protection. Electoral accountability ensures that public officials who regularly dissociate themselves from or interfere with popularly accepted perceptions of public welfare cannot obtain or hold their positions. Only speech, religions, and claims of individual prerogative that the public does not recognize in their situational context¹⁰⁰ as generally advancing societal interests need protection under the auspices of "right." "The institution of rights against the Government," insists Professor Ronald Dworkin, "is not a gift of God, or an ancient ritual or a national sport. It is a complex and troublesome practice that makes the Government's job of securing the general benefit more difficult and more expensive"¹⁰¹ If rights were cost free they would need no constitutional articulation. The Bill of Rights is inherently anti-governmental. It exists precisely to prohibit government from using certain specified means, whether or not such means are efficient or necessary to reach what the public believes to be legitimate and laudable societal goals.¹⁰²

A utilitarian perspective on rights developed as an outgrowth of the writings of the legal realists who perceived law's authority as stemming from its ability to affect society beneficially.¹⁰³ While the end/means utilitarian justification for law advanced by the realists appears highly appropriate in legislative and administrative policymaking forums and in the development of common law doctrine, the justification denigrates the unique status of claims of constitutional rights. Such constitutional claims take on an ethical and normative aura that normally must transcend a cost benefit analy-

100. While the public rhetorically may embrace the value of free speech, for example, everyone is not likely to defend the rights of Nazi's to march and demonstrate in the predominantly Jewish community of Skokie, Illinois. See generally Goldberger, *Skokie: The First Amendment Under Attack by its Friends*, 29 MERCER L. REV. 761 (1978).

101. R. DWORKIN, *TAKING RIGHTS SERIOUSLY* 198 (1977).

102. Amsterdam, *supra* note 3, at 353. In a democratic system, the written guaranty of individual rights that judicial institutions are to enforce either may be (1) an act of extreme elitism because the guaranty abrogates policies supported by popularly accountable institutions or, (2) an act of self-imposed paternalism accepted by a polity not always confident of its situational wisdom.

103. See, e.g., R. POUND, *AN INTRODUCTION TO THE PHILOSOPHY OF LAW* (rev. ed. 1954). According to Professor Roscoe Pound, a renowned spokesman of the utilitarian perspective, law is concerned with "social engineering" and the lawyer and jurist, at least when fulfilling their most creative roles, are "social engineers." *Id.* at 47.

sis within the specific situational contexts in which the claims arise. The rights' acceptance in light of their cost demands a sense of their transcendental value, their mysticism, rather than merely their instrumental justification.¹⁰⁴ Justice Brennan recognized the uniqueness of claims of right when he wrote, "the efficacy of the Bill of Rights as the bulwark of our national liberty depends precisely upon public appreciation of the special character of constitutional prescriptions."¹⁰⁵

Certainly, to argue that resource availability and the impact on important societal goals are irrelevant to the recognition and acceptance of a claim of constitutional right would be simplistic and constitutionally naive. When the Justices claim that such considerations are improper,¹⁰⁶ they most often merely are acknowl-

104. Courts considered and supported the non-pragmatic but symbolic role of law during the Watergate debates over the amenability of the President to judicial process. Judge Wilkey of the District of Columbia Circuit writing in dissent asked the realist's question: "[I]f the court . . . has no physical power to enforce its subpoena should the President refuse to comply . . . [then] what purpose is served by determining whether the President is 'immune' from process?" *Nixon v. Sirica*, 487 F.2d 700, 792 (D.C. Cir. 1973) (Wilkey, J., dissenting). The majority of the Court, however, disagreed with the implication of Judge Wilkey's rhetorical question. Instead, they upheld the ruling of the trial judge, Judge Sirica, who insisted, "That the Court has not the physical power to enforce its orders to the President is immaterial to . . . resolution of the issues. . . . [T]he Court has a duty to issue appropriate orders. . . . [I]t would tarnish the Court's reputation to fail to do what it could in pursuit of justice." *In re Grand Jury Subpoena Duces Tecum Issued to Richard M. Nixon*, 360 F. Supp. 1, 9 (D.D.C. 1973). Impact was not the issue; constitutional responsibility was.

105. *United States v. Havens*, 446 U.S. 620, 634 (1980). Only on rare occasions in recent years has the Court recognized the special character of constitutional rights that forbids balancing them against the interest of other societal goals. For example, in *New Jersey v. Portash*, 440 U.S. 450 (1979), the Court held that testimony given under a grant of legislative immunity could not be used to impeach the witness at his later trial for misconduct in office. Regardless of the societal interest in keeping a defendant from benefiting from a possible perjury, the Court concluded that: "Here . . . we deal with the constitutional privilege against compulsory self-incrimination in its most pristine form. Balancing, therefore, is not simply unnecessary. It is impermissible." *Id.* at 459.

106. In *Mayer v. Chicago*, 404 U.S. 189 (1971), the Supreme Court upheld an indigent's right to a free trial record for appeal purposes even though the crime was punishable only by fine. Justice Brennan, writing for the Court, insisted that

Griffin [v. Illinois] does not represent a balance between the needs of the accused and the interests of society; its principle is a flat prohibition against pricing indigent defendants out of as effective an appeal as would be available to others able to pay their own way. . . . The state's fiscal interest is, therefore, irrelevant.

Id. at 196-97. In a society more and more conscious that resources are finite, however, decisionmakers (judges and others) are more likely to recognize rights when resources are readily available to effectuate such rights than when recognition would bankrupt other societal goals. Few judges have been as candid as Judge Clement Haynesworth who, in *Moffitt v. Ross*, 483 F.2d 650 (4th Cir. 1973), *rev'd*, 417 U.S. 600 (1974), supported a right to appointed counsel for permissive appellate review by arguing that "[w]hat is requisite today

edging the myths of constitutional continuity and stability, and are using them to bolster their own position even though they are aware that the myths are unattainable. The proper decisional process for courts when determining whether to recognize a right possessing a scope broad enough to encompass the claim of its violation in a specific case, however, is fundamentally different from the decisional process that courts should use once they accept the right and find a violation. Interpretation of the open textual clauses of the Bill of Rights may require an evaluation of whether the interpreted right proffered is worth the societal cost that its acceptance would entail. If courts can define the right in such a way to allow most of its potential benefit while reducing its cost, so much the better.¹⁰⁷ But once courts recognize a right, they no longer should be able to find its enforcement too costly in specific cases. Furthermore, if the concept of right is to have any meaning, once the court acknowledges a violation it must provide a remedy that allows the injured individual to gain vindication for his claim. Courts, therefore, justifiably may consider costs and benefits when determining the meaning and scope of rights. But consideration of such factors is improper when the query is whether courts should grant a remedy to an individual whose constitutional rights acknowledgedly were violated by the state.

Admittedly, the fourth amendment is not cost free. Public safety, indeed, may be more difficult to provide or expensive to obtain when the fourth amendment restricts the types of behavior government can utilize in fighting crime. The existence of the amendment, however, exemplifies a societal belief that, although community safety is a legitimate public concern, the price for such

may not have been constitutionally requisite ten years ago, or even a few years ago. As our legal resources grow, there is correlative growth in our ability to implement basic notions of fairness." *Id.* at 655.

107. This perspective explains those cases that only permit prospective application of newly articulated criminal process rights. *See, e.g.,* *Desist v. United States*, 394 U.S. 244 (1969); *Tehan v. United States ex rel. Shott*, 382 U.S. 406 (1966); *Linkletter v. Walker*, 381 U.S. 618 (1965). Requiring retroactive application of all such rights might require the release or retrial of numerous convicted criminals. Jurists, thus, may find it difficult to accept that the societal gain from the new interpretations is worth the societal loss that extremely disrupting the criminal justice system would entail. Prospective only application allows for evolution of new interpretations of process rights as ethical perspectives change without requiring the system first to absorb the cost of former misinterpretations. Consequently, nonretroactivity cases such as *Peltier*, *see supra* text accompanying notes 95-98, encourage constitutional development in the fourth amendment field. This Article later notes that a general good faith exception retards precisely such a development. *See infra* text accompanying notes 200-15. *Peltier*, therefore, is distinguishable and good faith proponents cannot justifiably view it as precedential support for a broad good faith exception.

safety, seen in terms of lost individual privacy, can become too great.¹⁰⁸ The Constitution accomplishes this accommodation of the public's interest in safety and the individual's interest in privacy through a guarantee of individual rights that is not absolute: the amendment's language clearly indicates that government rightfully can invade individual privacy when it acts reasonably, often necessitating the existence of probable cause and the possession of a warrant.¹⁰⁹ An analysis of costs and benefits is built into the amendment itself. What courts must decide in the individual case is whether the government acted reasonably in invading the individual's privacy, not whether protection of such privacy is too costly or of too little benefit to support vindication of the right once they find a violation.

A legal system that attempts simultaneously to control unacceptable deviance and the behavior of the laws' enforcers¹¹⁰ is understandably complex in its application to individual situations. A result in a specific case may seem inefficient and improper.¹¹¹ But

108. The development and utilization of technology capable of keeping all people under constant surveillance could drastically reduce the threat to societal order. But, in spite of the imminence of the year 1984, the prices of such an Orwellian state are too high.

109. Even under such a qualified right, the Constitution does inhibit somewhat the criminal justice system's quest for truth. The searches prohibited by the fourth amendment, if permitted, likely would expose much evidence of criminality that the courts, in turn, could use to convict the truly guilty and acquit the truly innocent. No question exists that the quest for truth is a fundamental goal of our criminal justice system, but it is not an exclusive goal nor even one superior to all others. The quest for truth rather is an objective that the criminal justice system may obtain only by means consonant with other objectives, most particularly those established in the Constitution. Justice Brennan articulated this view, insisting that,

The procedural safeguards mandated in the Framers' Constitution are not admonitions to be tolerated only to the extent they serve functional purposes that ensure that the "guilty" are punished and the "innocent" freed; rather, every guarantee enshrined in the Constitution, our basic charter and the guarantor of our most precious liberties, is by it endowed with an independent vitality and value, and this Court is not free to curtail those constitutional guarantees even to punish the most obviously guilty.

Stone v. Powell, 428 U.S. 465, 524 (1976) (Brennan, J., dissenting).

The community's concern about the effective prosecution of crime is understandable. Society, however, must be wary for "[t]he history of liberty has largely been the history of observance of procedural safeguards." *McNabb v. United States*, 318 U.S. 332, 347 (1943). Furthermore, the relaxation of those safeguards historically almost always has taken place in the face of plausible governmental claims of threats to good social order. Amsterdam, *supra* note 3, at 354.

110. See *supra* text accompanying notes 57-58.

111. *E.g.*, *Ybarra v. Illinois*, 444 U.S. 85 (1979). Police operating under a warrant to search a bar for narcotics also searched the several patrons of the bar then present. The state courts held the heroin discovered on one of the customers admissible because a state statute authorized the police when executing a warrant to search all those present. The statute's purpose was to protect the officers from attack and to prevent the disposal or con-

such inefficiency and impropriety may be the cost of having and enforcing controls upon the controllers—the fourth amendment, for example—in the first place.

2. Remedy as Vindication

For constitutionally guaranteed rights to represent something beyond simple platitudes, the remedy provided for their violation must have some measurable consequence that vindicates the right in a manner which “invokes and magnifies the moral and educative force of the law.”¹¹² The exclusionary rule, of course, does provide this needed confirmation of the right of privacy of one’s person and domain.¹¹³ During the evolution of exclusionary rule doctrine, however, courts increasingly have ignored the rule’s efficacy in vindicating the violated right. Judicial perception of the rule’s purpose has become misdirected. The relevant concern has switched from whether the rule vindicates the right violated to whether it serves to deter illegal police behavior by “punishing” a police wrongdoer.¹¹⁴ The legitimate purpose of exclusion is not to place sanctions upon the specific officer, but to impose them upon the legal system in whose name he is functioning. Exclusion not only vindicates the right violated, but also eliminates the appearance of a system that encourages practically what it condemns rhetorically.

Weeks clearly defended exclusion upon this basis.¹¹⁵ But courts have interpreted *Mapp* as recognizing the suppression doctrine’s utilitarian deterrent function. *Mapp*, however, does not dis-

cealment of those things described in the warrant. The Supreme Court reversed the convictions concluding that an individual’s mere presence at the site of a warranted search does not subject him to a search of his person. Permitting such a search, however, certainly would provide protection of the officer and prevent defeat of the warrant’s goal. Despite these efficiency concerns, however, the Court recognized that probable cause to search must be more particularized than the statute dictated if the fourth amendment was to remain relevant to individuals functioning in a mobile society. Rights and symbols took precedence over efficiency.

112. Professor Dallin Oaks, often viewed as an exclusionary rule detractor, recognized that the rule functions in this manner. Oaks, *supra* note 3, at 756.

113. If a court admitted illegally obtained evidence at trial, the propriety of the illegal conduct would appear confirmed rather than the violated right vindicated. See *Terry v. Ohio*, 392 U.S. 1, 13 (1968).

114. This focus on punishing police is very strong among the exclusionary rule’s critics. Wilkey, *supra* note 3, at 220. The good faith exception further confuses the issue of the fulfillment of a right with that of the culpability of an officer. Certainly, the law should not punish a nonculpable officer. See *infra* text accompanying notes 142-43, & 175. But a police officer’s good faith error should not make the violation of a citizen’s reasonable expectations of privacy a matter of inconsequence and system indifference.

115. See *supra* note 24 and text accompanying notes 23-24.

cuss punishment of culpable officials, but rather in using the language of *Elkins*, it insists that exclusion is necessary to deter by "compell[ing] respect for the constitutional guaranty . . . by removing the incentive to disregard it."¹¹⁶ Courts, thus, were to achieve deterrence not by punishing the erring officer, but by nurturing respect for the right through creation of a system that in no way supported or encouraged the right's violation. The Court recognized that the exclusionary rule would not end all illegal searches or seizures, but it would assure that the judicial system would not approve violations, thus ensuring the right's vindication. *Mapp's* concept of deterrence was inherent in, rather than separate from, its interest in judicial integrity and in law playing a role as educator. The Court did not consider the exclusionary rule to be based upon a shakey utilitarian foundation.¹¹⁷

Largely due to this confusion surrounding *Mapp*, Supreme Court opinions considering fourth amendment violations and their appropriate remedies, at times, have appeared to accept inconsistent or contradictory doctrines.¹¹⁸ Standing cases have required that defendants demonstrate distinctly personal rights or interests in the evidentiary items they seek to suppress before they can press for an exclusionary remedy because "Fourth Amendment rights are personal rights which may not be vicariously asserted."¹¹⁹ Thus, prosecutors may use illegally obtained evidence at trial against all but those whose particular expectation of privacy the police invaded. At other times, however, the Court has insisted that although the defendant can prove a personal violation, exclusion may not be necessary. These cases view the exclusionary rule as a "judicially created remedy designed to safeguard Fourth Amendment rights generally through its deterrent effect, rather than a personal constitutional right of the party aggrieved."¹²⁰ Under this doctrine the vindication of individual rights becomes

116. *Mapp*, 367 U.S. at 656 (emphasis added).

117. Ironically, the utilitarian analysis of situationally balancing costs and benefits, apparently accepted by a majority of the Court, may encourage law enforcement agents to do similar balancing when considering whether to search or arrest. Given the police officer's likely bias for ferreting out crime, see *supra* text accompanying notes 65-66, the utilitarian perspective may diminish the exclusionary rule's ability to deter a specific officer from acting illegally.

118. See generally Burkoff, *The Court that Devoured the Fourth Amendment: The Triumph of an Inconsistent Exclusionary Doctrine*, 58 OR. L. REV. 151 (1979).

119. *Alderman v. United States*, 394 U.S. 164, 174 (1969).

120. *United States v. Calandra*, 414 U.S. 338, 348 (1974). See also *Stone v. Powell*, 428 U.S. 465, 482 (1976); *United States v. Peltier*, 422 U.S. 531, 538-39 (1975).

secondary. The primary purpose becomes regulating governmental conduct by assuring *adequate* deterrence of illegality. Thus, a prosecutor may not use unlawfully acquired evidence at trial for his case in chief but may use the evidence for impeachment purposes¹²¹ or before a grand jury.¹²²

Professor Anthony Amsterdam suggests that these two lines of cases represent divergent views of the fourth amendment. Under one view the amendment is a "collection of protections of atomistic spheres of interest of individual citizens . . ." Under the other view the amendment is establishing a regulatory principle requiring police to act in a certain way in order to protect people generally.¹²³ While the atomistic language used in the standing line of cases does appear inconsistent with the regulation language in the other, two explanations are possible. The first explanation accepts that the rhetoric is inconsistent, but points out that both lines of cases functionally support limitations upon the scope and effect of the exclusionary rule. Under this explanation the Court is merely focused on the unitary goal of reducing the scope and effect of the rule and is willing to be inconsistent in its rhetoric if such inconsistency furthers its goal.

The second, and more palatable explanation, argues that the two perspectives are not inconsistent, rather, they represent two different stages in fourth amendment analysis. The first stage concerns a determination of who may demand vindication of a violated right. Only those who can demonstrate deprivation of personal rights will have the court address their claims at stage two. The second stage concerns choosing an appropriate remedy. Not all defendants whose fourth amendment rights the police have violated may insist upon the same remedy. At this stage, the court may consider the public utility of various remedies. From this perspective, a court legitimately may restrict application of the exclusionary rule "to those areas where its remedial [deterrent] objectives are thought most efficaciously served."¹²⁴ The remedial deterrent objective may not be as readily achievable when a police infraction is due to a reasonable good faith error. Thus, under this explanation the good faith exception indeed seems attractive.

An individual whose claim has reached the second stage, however, already has proved violation of his fourth amendment rights.

121. *United States v. Havens*, 446 U.S. 620 (1980).

122. *United States v. Calandra*, 414 U.S. 338 (1974).

123. Amsterdam, *supra* note 3, at 367.

124. *Calandra*, 414 U.S. at 348.

The only question remaining is what remedy device a court should choose to vindicate such rights. A decision against applying the exclusionary rule, therefore, must assume that an alternative remedy exists and is preferable. Otherwise a court, although acknowledging the right's violation, would be treating it as a matter of indifference, and the claim of right would go unvindicated. This argument is essentially the same as that the Court made in *Mapp* itself: "without the [exclusionary] rule the assurance against unreasonable . . . searches and seizures would be 'a form of words,' valueless and undeserving of mention in a perpetual charter of inestimable human liberties . . ."¹²⁵ An appropriate retort is that all this Article has proved is that a court must vindicate an individual whose fourth amendment rights it recognizes as violated by imposing a remedy of some measurable consequences, but not necessarily with a remedy as costly to the criminal justice system as the exclusionary rule. Chief Justice Burger has championed this contention observing that "[r]easonable and effective substitutes [for the exclusionary rule] can be formulated . . ."¹²⁶ Contrary to the Chief Justice's assertion, however, exclusion provides the only effective remedy for fourth amendment violations, especially those made in good faith.

Although alternative legal remedies for illegal searches and seizures have superficial appeal, all have serious defects.¹²⁷ A court likely would not issue injunctions prohibiting illegal police behavior: although they could be useful against violations sanctioned by promulgated state policy,¹²⁸ such injunctions likely would be inef-

125. *Mapp*, 367 U.S. at 655.

126. *Bivens v. Six Unknown Fed. Narcotics Agents*, 403 U.S. 388, 421 (1971) (Burger, C.J., dissenting).

127. *Oaks*, *supra* note 3, at 673.

128. Even such a limited use of injunctions is unlikely. In a recent suit, petitioner, claiming to have been choked illegally by police who were acting consistent with police department policy, sought an injunction against such police behavior. In *City of Los Angeles v. Lyons*, 103 S. Ct. 1660 (1983), the Supreme Court ruled that federal courts could not grant such injunctions. The Court held that while plaintiff conceivably should receive damages for injuries already suffered, an injunction applies to future behavior. Since the plaintiff could not demonstrate that in the future police again would stop and illegally choke him, his need for an equitable remedy was too speculative. *Id.* at 1669. The plaintiff was unable to show a "personal stake in the outcome" of the case, *id.* at 1665, and thus the Court dismissed plaintiff's appeal for lack of an established case or controversy. Justice Marshall noted in dissent that the ruling virtually removed an entire class of constitutional violations from the courts' equitable powers. "It immunizes from prospective equitable relief," he insisted, "any policy that authorizes persistent deprivations of constitutional rights as long as no individual can establish with substantial certainty that he will be injured, or injured again, in the future." *Id.* at 1683.

fective against situational violations by individual police officers.¹²⁹ Criminal prosecution of police officers for improper searches and seizures also would be ineffective. Police are reluctant to arrest, prosecutors are reluctant to prosecute, and juries are reluctant to convict officers who make illegal searches in an effort to protect society from crime.¹³⁰

Additionally, although improper police behavior is a common-law tort, juries are not likely to find significant damages¹³¹ against police, especially in favor of a plaintiff accused or convicted of crime.¹³² Besides, damages may not be collectible if the officer lacks recoverable resources and sovereign immunity protects the officer's employing governmental unit.¹³³ Although *Monell v. Department of Social Services*¹³⁴ held that injured parties could sue cities for violating constitutional rights under 42 U.S.C. § 1983, the Supreme Court specifically cautioned that liability could not exist unless the city's agent was acting in furtherance of official policy.¹³⁵ Consequently, most illegal searches and seizures, involving situational use of discretion by police, can not lead to governmental liability. Thus, tort liability also proves an ineffective remedy for these fourth amendment violations. Even assuming such liability were available, however, such a remedy might give the impression that government has the authority to simply buy out of any responsibility it has to comply with constitutional commands. Damages from illegal searches then would become just one more added

129. The Court often has demonstrated its lack of enthusiasm for granting injunctions that would require active and continuous monitoring by the courts in the situational enterprise of police work. See *Rizzo v. Goode*, 423 U.S. 362 (1976); *O'Shea v. Littleton*, 414 U.S. 488 (1974).

130. Edwards, *Criminal Liability for Unreasonable Searches and Seizures*, 41 VA. L. REV. 621, 628-29(1955); see *Mapp v. Ohio*, 367 U.S. at 651-52.

131. Even if juries did find damages, common law torts, such as trespass, focus upon a different kind of damage to the plaintiff than a fourth amendment violation. "It is not the breaking of his doors, and the rummaging of his drawers, that constitutes the essence of the offense," recognized Justice Clark in *Mapp*, "but it is the invasion of his inalienable right of personal security, personal liberty and private property. . . ." 367 U.S. at 641 (quoting *Boyd v. United States*, 116 U.S. 616, 630 (1886)). Treating a fourth amendment violation as simply another property tort insufficiently vindicates the claim of right involved.

132. Foote, *Tort Remedies for Police Violations of Individual Rights*, 39 MINN. L. REV. 493 (1955); see generally *Symposium—Police Tort Liability*, 16 CLEV.-MAR. L. REV. 397 (1967).

133. See Dakin, *Municipal Immunity in Police Torts*, 16 CLEV.-MAR. L. REV. 448 (1967). Of course, the sovereign may waive its immunity.

134. 436 U.S. 658 (1978).

135. *Id.* at 694. *Monell*, therefore, discourages police departments from prescribing rules for police conduct. Such rules would increase the risk of departmental liability.

cost of doing business.¹³⁶ Such a response would not vindicate the constitutional right but rather would diminish its luster.¹³⁷

Furthermore, more informal remedies also are impractical. Police solidarity and their intense commitment to values consistent with their view of job requirements and professional skills¹³⁸ make internal administrative review and officer discipline of questionable value.¹³⁹ Internal review also may exacerbate the concern segments of the community have over the bias they believe is evident when police investigate complaints of police abuse. The appearance of partiality may cause possible complainants to lose all trust in a police-controlled system. The police's intensely negative reaction to outsider control has made civilian review boards with any power of decision and enforcement difficult to establish and maintain.¹⁴⁰

All alternative remedies to exclusion are even more clearly ineffective when the officer's violation was due to a reasonable good faith error. If the police subculture tolerates and sometimes encourages intentional fourth amendment violations,¹⁴¹ then good faith violations could not conceivably result in intradepartmental discipline. Criminal conviction for good faith errors is even less conceivable or desirable. Furthermore, the requisite *mens rea* would be extremely difficult, if not impossible, to establish. Impos-

136. Dellinger, *Of Rights and Remedies: The Constitution as a Sword*, 85 HARV. L. REV. 1532, 1562 (1972).

137. Of course this danger assumes police consciously would decide to violate fourth amendment rights if the benefits thus gained justify in their minds the damages that would ensue. Such decisions would not be a simple good faith error. Other problems arise, however, in using tort remedies when dealing with good faith mistakes. See *infra* text accompanying note 142.

138. See *supra* text accompanying notes 65-78.

139. Occupational pressures frequently require police to make instantaneous decisions. The low visibility of the frequent discretionary actions of police limits the ability even of police officials to direct the behavior of their rank and file. See generally Goldstein, *Police Discretion Not to Invoke the Criminal Process: Low-Visibility Decisions in the Administration of Justice*, 69 YALE L.J. 543 (1960). Police, in fact, represent one of the rare bureaucracies in which discretion increases as one goes down the hierarchy. This inability to supervise the police officer on the streets constitutes an unresolved problem for those who recommend an increased effort at rulemaking by police officials. See, e.g., Amsterdam, *supra* note 3, at 416-39; Caplan, *The Case for Rulemaking by Law Enforcement Agencies*, 36 LAW & CONTEMP. PROBS. 500 (1971); Kaplan, *supra* note 3, at 1050-55. Rulemaking efforts may succeed only in alienating police rank and file from their upper echelon.

140. Barton, *Civilian Review Boards and the Handling of Complaints Against the Police*, 20 U. TORONTO L.J. 448 (1970); Hudson, *Police Review Boards and Police Accountability*, 36 LAW & CONTEMP. PROBS. 515 (1971). Civilian review boards in both New York City and Philadelphia, for example, had exceedingly short lives. Hudson, *supra* at 522-28.

141. See *supra* text accompanying notes 79-86.

sition of tort liability, in addition to all the difficulties heretofore mentioned, is virtually impossible since courts clearly recognize that law enforcement officers possess a common-law immunity to civil liability for good faith errors.¹⁴²

Protecting the individual offending officer from "punishment" for good faith errors is justifiable since, although he is a wrongdoer, he is without culpability. But inherent in the concept of a good faith exception is the admission that the police have violated an individual's right and the courts must vindicate his claim of right. Once a court acknowledges the individual's claim as justified it may choose among remedies. But as Professor Oaks conceded "[i]t would be intolerable if the guarantee against unreasonable search and seizure could be violated without practical consequence."¹⁴³ Since all alternative remedies to exclusion are *least* convincing when dealing with good faith police error, the suppression doctrine is the only conceivable device for vindication of the admittedly violated right. In the context of a good faith error Justice Murphy's dissent in *Wolf* is correct: "there is but one alternative to the rule of exclusion. That is no sanction at all."¹⁴⁴ Courts must not treat constitutional violation as a matter of indifference or the public soon will perceive the right itself similarly.

B. *The Meaning of Deterrence*

The preceding discussion illustrates that the proper justification for the exclusionary rule is not its deterrent value but rather its ability to appropriately remedy fourth amendment violations. Since exclusion is truly the only remedy available when the offending officer has made a good faith error, focusing on the rule's inability to deter such an error ignores that a violated right goes unremedied. Even if, however, the rule's deterrent value were a valid consideration in deciding whether exclusion is a proper remedy for a good faith error, the specific deterrence concept, which the exception's proponents use as their primary justification for the good faith exception, is unnecessarily narrow.

Specific deterrence emphasizes the effect of evidence exclusion upon the specific offending officer. The good faith exception's proponents focus on the exclusionary rule's inability to deter an of-

142. The Supreme Court discussed and extended this common law immunity to § 1983 suits in *Pierson v. Ray*, 386 U.S. 547 (1967).

143. Oaks, *supra* note 3, at 756. Professor Oak's reputation as one of the most respected critics of the exclusionary rule makes this admission quite significant.

144. 338 U.S. at 41.

ficer's behavior when he, in good faith, is unaware of his transgression.¹⁴⁵ Certainly, an officer who reasonably and in good faith believes his behavior legal is not likely to avoid that behavior for fear of exclusion. To argue that the system should not punish such a nonculpable officer by exclusion, however, confuses the concepts of deterrence and punishment and ignores the deterrent potential of exclusion with respect to the officer's future behavior.

Admittedly a nonculpable officer does not deserve punishment. The appropriateness of punishment, however, is a separate and distinct concern from the rule's deterrent value. Exclusion of the evidence seized under a reasonable good faith error undoubtedly has the potential to deter the officer from committing the same error in the future.¹⁴⁶ Consequently, while punishing a nonculpable officer may not seem fair, such a concern should not obscure the exclusionary rule's future deterrent potential.

Those urging a good faith exception also argue that the fear of exclusion will not prevent an officer who has had evidence excluded because of a reasonable good faith error from thereafter making a second good faith error in a different context. They, therefore, argue that excluding evidence seized in reasonable good faith has no specific deterrent effect. But this argument proves too much. If the inability of the exclusionary rule to deter specifically justifies a reasonable good faith exception it also justifies an unreasonable good faith exception. Deleting the objective reasonableness component of the good faith exception has no effect on this proposed nondeterrence justification. An officer making an unreasonable, but honestly held, good faith error is no more likely to question his own behavior and is no more likely to have his behavior deterred by the fear of exclusion than is the officer making a reasonable good faith error. As Judge Rubin of the Fifth Circuit aptly observed, "A policeman who is in complete subjective good faith is

145. Specific deterrence is also the focus of much of the conflict about the exclusionary rule generally. Insisting that the suppression doctrine does not deter illegal police behavior, *see supra* text accompanying notes 60-92, critics argue for its total abolition. The same critics, however, also argue that the exclusionary rule hamstringing police efforts at law enforcement. This argument is paradoxical: how can the rule both fail to specifically deter and yet hamstring the officer? For an attempt to resolve this paradox, see Ingber, *Procedure, Ceremony and Rhetoric: The Minimization of Ideological Conflict in Deviance Control*, 56 B.U.L. REV. 266, 305-07 (1976).

146. Arguably each new situation may contribute different complicating factors that allow for new good faith mistakes. See *infra* text accompanying notes 150-53 for a discussion of a method to avoid this danger.

unlikely to stop to ask himself, 'Am I also reasonable?' ”¹⁴⁷

An unreasonable good faith exception is, of course, unacceptable because of the premium it would place on police ignorance.¹⁴⁸ Consequently, the foregoing rationale makes patently obvious that the rule's inability specifically to deter good faith errors by offending officers cannot be the only justification for a good faith exception. If a lack of the rule's deterrent value is, indeed, a justification for the proposed exception, then its proponents must expand their concept of deterrence. Upon such expansion, however, the deterrent value of excluding evidence obtained through a good faith police error becomes obvious.

Judicial suppression of evidence also may deter generally by influencing other officers not to duplicate the offending officer's error. An identification of the conduct causing a constitutional violation (whether committed culpably or not) will communicate to others that a trial court will exclude the fruits of their search if they act similarly. Hopefully such communication will prevent such unconstitutional searches in the future. Viewed in this manner, the suppression remedy does not protect only those against whom police find incriminating evidence. The exclusion of such evidence will have an "overflow effect" upon all individuals by discouraging unwarranted searches whether or not they would be fruitful. Courts then, by removing the incentive to search, would protect the rights of both the innocent and the guilty.¹⁴⁹

Good faith exception proponents claim that such general deterrence is both unrealistic and often undesirable. They assert, for example, that court decisions are too complex, too vague, and too situational to give police the guidance necessary for general deterrence to function. The inability of police to understand judicially articulated fourth amendment standards, however, is not related to the issue of remedy, but is a product of the fact-orientation of search and seizure litigation and court opinions. Unintelligible rules of conduct will diminish the deterrent potential of any remedy chosen to vindicate fourth amendment rights.

147. *United States v. Williams*, 622 F.2d 830, 850 n.4 (5th Cir. 1980) (en banc), cert. denied, 449 U.S. 1127 (1981) (Rubin, J., concurring).

148. Kaplan, *supra* note 3, at 1044.

149. Expanding the deterrence beyond a specific infractor exposes one danger of the good faith exception. If an officer knows that an honest and reasonable error excuses his behavior, he willfully may feign a good faith mistake hoping that the court will find his error reasonable without discovering the secret of his willfulness. If the good faith exception has such an effect it would encourage both illegal searches and police perjury.

The negative effect that judicial imprecision has had on the exclusionary rule's general deterrent value does not now justify recognition of a good faith exception. The deterrence problem lies not with the rule, but rather with its judicial administration. Admittedly, the Court has made difficult its responsibility of finding a remedy for fourth amendment violations by erroneously focusing on the rule's deterrent value rather than on its ability to remedy the constitutional violation, and then by negating that deterrent value through unclear opinions. The solution, however, cannot be the abrogation of the Court's solemn responsibility to provide a remedy for a violated fourth amendment right.¹⁵⁰

The manner in which to avoid decreasing the rule's deterrent potential through court opinions unintelligible to police is not to treat the constitutional violations with indifference, but rather to clarify fourth amendment standards through opinions drawing simple and understandable "bright lines" for police conduct.¹⁵¹ "Fourth Amendment doctrine, given force and effect by the exclusionary rule," Professor LaFave observes,

is primarily intended to regulate the police in their day-to-day activities and thus ought to be expressed in terms that are readily applicable by the police in the context of the law enforcement activities in which they are necessarily engaged. A highly sophisticated set of rules, qualified by all sorts of ifs, ands, and buts and requiring the drawing of subtle nuances and hairline distinctions, may be the sort of heady stuff upon which the facile minds of lawyers and judges eagerly feed, but they may be "literally impossible of application by the officer in the field."¹⁵²

Lack of judicial clarity should not lead to a device for excusing fourth amendment violations, but to a method for simplifying the task of officers truly endeavoring to respect constitutional prerogatives.¹⁵³ Consequently, those urging adoption of a good faith excep-

150. The good faith exception, as already demonstrated, leaves the violated right unremedied. See *supra* text accompanying notes 141-44.

151. In recent years courts have made some attempts to formulate such bright lines. See, e.g., *United States v. Ross*, 456 U.S. 798 (1982); *New York v. Belton*, 453 U.S. 454 (1981).

152. LaFave, "Case-by-Case Adjudication" Versus "Standardized Procedures": The Robinson Dilemma, 1974 SUP. CT. REV. 127, 141 (footnotes omitted). Of course, while opinions setting out "bright lines" more clearly would communicate search and seizure standards for police, not all bright lines are equally defensible. See LaFave, *supra* note 3, at 325-33. Conflicts over where the courts should draw the line—what they should define the standard of conduct to be—would continue. See *id.*

153. Ironically, a court that often has rejected the precision of per se rules for the flexibility (and uncertainty) of "totality of circumstances" analysis, see *Manson v. Brathwaite*, 432 U.S. 98 (1977); *supra* note 87 and accompanying text, also bemoans the inability of law enforcement officials to deduce appropriate constitutional standards. When

tion should not justify the exception on the grounds that lack of judicial clarity creates general deterrence problems.

Exception proponents, however, also argue that general deterrence in a good faith context actually may be undesirable.¹⁵⁴ Justice White, for example, one of the most vociferous advocates of a good faith exception,¹⁵⁵ defended his position by asserting:

The officers [making a reasonable good faith error], if they do their duty, will act in similar fashion in similar circumstances in the future; and the only consequence of the rule as presently administered is that unimpeachable and probative evidence is kept from the trier of fact and the truth-finding function of the proceedings is substantially impaired or a trial totally aborted.¹⁵⁶

Consequently, he concludes, that "[m]aking the arrest in such circumstance is precisely what the community expects the police officer to do."¹⁵⁷ According to Justice White, therefore, deterring such behavior is undesirable.

This position reveals the hidden agenda behind the good faith exception. If the police officer truly acted, under the circumstances, in a manner that courts should not deter, why is his behavior anything but reasonable? If, then, the search and arrest is reasonable, it is also constitutional under the fourth amendment.¹⁵⁸ Rather than merely advocating a good faith exception to the exclusionary rule, Justice White is surreptitiously, consciously or unconsciously, disagreeing with the judicial determination that the police action violates fourth amendment standards. If a violation actually did take place, even assuming the officer was only reasonably confused and chose to act erroneously, the community still should have preferred him to have acted otherwise and not to have violated the constitutional mandate. White's argument lies not with the remedy but with the judicial determination that a constitutional violation had taken place. He would have chosen to find that the search was constitutional. Other exception proponents who argue that the general deterrence of the exclusionary rule is undesir-

the Court then considers correcting the problem by disregarding legitimate claims of right under the auspices of a good faith exception the court magnifies this irony.

154. The Fifth Circuit, for example, asserted that general deterrence in a good faith context is undesirable. *United States v. Williams*, 622 F.2d 830, 849 (5th Cir. 1980) (en banc), cert. denied, 449 U.S. 1127 (1981). See *supra* text accompanying note 94.

155. Justice White presents his view most completely in *Illinois v. Gates*, 103 S. Ct. 2317, 2340-47 (1983) (White, J., concurring).

156. *Stone v. Powell*, 428 U.S. 465, 540 (1976) (White, J., dissenting).

157. *Id.* at 539.

158. Fourth amendment privacy is not an absolute right, but only a qualified one outweighed when governmental intrusion is reasonable. See *supra* text accompanying note 109.

able with respect to good faith errors also actually are saying that the scope of the fourth amendment should be sufficiently narrow as not to prohibit the police behavior at issue.

Further proof of this observation becomes evident by tracing the application of Justice White's position in the decisions of other courts such as *Colorado v. Quintero*.¹⁵⁹ In *Quintero* the police arrested defendant for theft and possession of a stolen television set. At the time of the arrest, defendant claimed to have no identification and to have purchased the television set from someone in the neighborhood. The police knew that defendant was a stranger to the community and had seen him attempt to cover the set with his shirt. At the time of the arrest, however, no complaints or evidence established that defendant had committed a crime.¹⁶⁰ On the basis of these facts, the Supreme Court of Colorado found that the officer had acted without probable cause in arresting and searching the defendant. According to the court, probable cause to arrest exists only "when the facts and circumstances within an officer's knowledge are sufficient to support a reasonable belief that a crime has been committed by the person arrested."¹⁶¹

Justice Rovira dissented. He initially argued that the officer's actions were reasonable and, therefore, constitutional. Only subsequently did he assert that a good faith error should not activate the exclusionary rule, quoting for support Justice White's comment that the officer had acted precisely as the community would expect.¹⁶² Clearly both arguments are the same. According to Justice Rovira, the court's interpretation of the fourth amendment standard is too stringent. If no constitutional violation occurred, then no remedy vindicating the right would be necessary. Justice Rovira is using the good faith exception covertly to undermine the claim of right acknowledged by the majority. The opinions of both Justices White and Rovira illustrate that exception proponents who argue that the general deterrence of the exclusionary rule is undesirable with respect to good faith errors actually are saying that the fourth amendment itself should not prohibit the implicated police behavior. Perhaps the fourth amendment should not. The issue, however, is one of substantive standards that needs de-

159. 657 P.2d 948, *cert. granted*, 103 S. Ct. 3535 (1983), *cert. dismissed*, 52 U.S.L.W. 3460 (Dec. 13, 1983).

160. The police, in fact, did not learn who owned the television set until more than five hours after the arrest. *Id.* at 950.

161. *Id.*

162. *Id.* at 952.

bate and resolution on that level rather than obfuscation and camouflage in the garb of the seemingly innocuous issue of remedy.

The foregoing discussion sets the stage for a reconsideration of whether *DeFillippo* is good precedent for a general good faith exception.¹⁶³ The officer in *DeFillippo* arrested and searched defendant for violating a statute later found unconstitutionally vague. Chief Justice Burger, ruling the seized contraband admissible, properly asserted that “[s]ociety would be ill-served if its police officers took it upon themselves to determine which laws are and which are not constitutionally entitled to enforcement.”¹⁶⁴ The officer in *DeFillippo* acted precisely as the community would wish. He limited his concern about constitutional rights to the sphere of police law enforcement. The fourth amendment, however, asserts that an officer acting unreasonably, without probable cause and a warrant when necessary, clearly is not acting as the Constitution requires. Had such an officer acted otherwise, in accordance with his responsibility under the clause, his choice would be constitutionally preferable and worthy of praise. In *DeFillippo*, the constitutional error was the legislature’s not the officer’s, and was not of fourth amendment origin. The exclusionary rule’s justification is to protect an individual’s privacy from unwarranted intrusion; it is not a remedy for all constitutional violations.¹⁶⁵ The Court’s holding in *DeFillippo* is distinguishable from one adopting a general good faith exception on the basis of differences in the desirability of general deterrence directed toward police behavior when the constitutional error is that of the legislature rather than of the officer. *DeFillippo*’s precedential value for a good faith exception, thus, is questionable.

The foregoing analysis of general and specific deterrence reveals that some advocates of the good faith exception utilize a deterrence rationale which is conceptually flawed. A consideration of systemic deterrence¹⁶⁶ reveals additional shortcomings in the advocates’ concept of deterrence. In addition to its specific and general deterrent value, the exclusionary rule also may deter illegal

163. See *supra* text accompanying notes 96-99.

164. *DeFillippo*, 443 U.S. at 38.

165. When a statute forbids leafletting on public streets the police officer’s responsibility is to enforce that law. A court need not exclude the narcotics found on an arrested leafletter because the statute violates first amendment rights. Other remedies are truly available for this claim of right violation such as declaring the statute unconstitutional and reversing any conviction for the statute’s breach.

166. “Systemic deterrence” is a construct borrowed from the writings of Mertens and Wasserstrom, *supra* note 3, at 399-401.

police behavior by increasing police officials' awareness of constitutional requirements and their responsibility to abide by them. If police officials realize the importance of complying with fourth amendment dictates, they can encourage individual officers to comply through training and developing departmental guidelines for searches, seizures, and arrests.¹⁶⁷ While individual officers initially may be hostile toward fourth amendment demands, departmental insistence on respect for such demands likely would reduce resistance to restrictions that police officers otherwise may view as illegitimately imposed by authorities external to the police subculture. The increased use of search warrants, the stepped up development of police educational programs, and the creation of ongoing working relationships between police and prosecutors already reflects the exclusionary rule's potential beneficial effects.¹⁶⁸

The exclusionary rule, however, also has had a less obvious, yet more important, systemic impact. Fourth amendment considerations only became part of the national consciousness when courts, due to the rule, had occasion and incentive to articulate them. The rhetoric of these decisions slowly acclimated the public to the fourth amendment values that the decisions defended.¹⁶⁹ Police departments, of course, recruit their officers from the general public. The police today, therefore, are more likely to be sympathetic to

167. *Id.* at 399. The low visibility of the discretion exercised by lower echelon police officers may limit the effectiveness of departmental rulemaking. *See supra* note 139.

168. LaFave, *supra* note 3, at 319. One Florida prosecutor noted that within his jurisdiction police are forbidden to seek a warrant before a prosecutor reviews their requesting papers.

169. Rhetoric can be a powerful instrument for change. For years ethnologists studying the relation of language to culture have insisted that any change in language constitutes an influence on both perception and conception. *See, e.g.,* R. BROWN, I. COPI, D. D. DULANEY, W. FRANKENA, P. HENLE & C. STEVENSON, *LANGUAGE, THOUGHT & CULTURE* 1-24 (1958); E. Sapir, *Language*, in *SELECTED WRITINGS OF EDWARD SAPIR IN LANGUAGE, CULTURE AND PERSONALITY* 7-32 (D. Mandelbaum ed. 1949); B. Whorf, *The Relation of Habitual Thought and Behavior to Language*, in *LANGUAGE, CULTURE AND PERSONALITY: ESSAYS IN MEMORY OF EDWARD SAPIR* 75-93 (L. Spier, A. Hallowell & S. Newman eds. 1941). Edward Sapir, an early leader in ethnology wrote:

The relation between language and experience is often misunderstood. Language is not merely a more or less systematic inventory of the various items of experience which seem relevant to the individual, as is so often naively assumed, but is also a self-contained, creative symbolic organization, which not only refers to experience largely acquired without its help but actually defines experience for us by reasons of its formal completeness and because of our unconscious projection of its implicit expectation into the field of experience.

E. Sapir, *Conceptual Categories in Primitive Languages*, in *LANGUAGE IN CULTURE AND SOCIETY* 128 (D. Hymes ed. 1964). Rhetoric thus may grease the wheels of ideological change. *See generally* Ingber, *supra* note 145, at 268-73.

fourth amendment values then were their pre-exclusionary rule predecessors.

Some critics of the exclusionary rule, however, claim that the lack of a good faith exception reduces rather than enhances the potential of systemic deterrence. They argue that departments which attempt to educate their officers but still find evidence excluded due to further development or change in judicial interpretation may adopt an attitude that "the courts cannot be satisfied, that the rules are hopelessly complicated and subject to change, and that the suppression of evidence is [thus] the courts' problem and not the departments'."¹⁷⁰ Such frustration may discourage a department from even trying to educate its officers to avoid fourth amendment violations.

This assertion is similar to the objection of tort defendants that the law does not give sufficient direction as how to avoid liability.¹⁷¹ A jury determination that the defendant failed to act in a reasonable, prudent manner, indeed, provides little guidance as to what behavior would have been sufficient. Even a finding of no liability leaves the defendant and those in similar situations wondering whether he was more cautious than required.

This vagueness in tort law, however, is socially valuable. While violation of a known or knowable standard of conduct makes moral reproach and punishment more justified,¹⁷² tort law does not focus on individual moral responsibility.¹⁷³ Tort law functions not to earmark personal culpability, but to continuously press societal members to make choices that will lower the risk of injury and pain.¹⁷⁴ Legally fixing the duty of an individual in a particular con-

170. Kaplan, *supra* note 3, at 1050. See J. SKOLNICK, *supra* note 61, at 212-29.

171. Justice Holmes enthusiastically supported the movement to define more stably the responsibility of both tort defendants and plaintiffs. *Baltimore & O. R.R. Co. v. Goodman*, 275 U.S. 66 (1927); *Lorenzo v. Wirth*, 170 Mass. 596, 49 N.E. 1010 (1897).

172. Consequently, the state must have given the defendant in a criminal trial fair notice at the time of his offense of the standards of criminal liability. Criminal law equates individual moral culpability and accountability. The law justifies punishment only if the wrongdoer deserves some measure of societal approbrium. As the text suggests, violation of a known or knowable standard of conduct makes such a societal response more appropriate.

173. The widespread use of liability insurance, which associates risks with activities rather than with individuals, clearly indicates a concern other than individual accountability. Furthermore, courts hold mentally limited individuals to a standard of care even when they can demonstrate their personal inability to function at the required level. *Vaughan v. Menlove*, 3 Bing. 468 (N.C. Ct. Comm. Pleas 1837). Clearly, moral responsibility is not the crucial issue.

174. The use of tort law to reduce the costs of accidents is discussed in G. CALABRESI, *THE COSTS OF ACCIDENTS* 26-28 (1970). See also Ingber, *Defamation: A Conflict Between Reason and Decency*, 65 VA. L. REV. 785, 812-33 (1979).

text and steadfastly refusing to change that duty removes the need to strive to develop better ways of avoiding injuries. Legally fixed duties, thus, would retard the valuable evolution of tort law.

The exclusionary rule's application to fourth amendment violations serves the same type of socially valuable function as tort law. The rule's purpose is not to punish culpable officers, but to structure a criminal justice system with incentives for police to strive continuously to respect more fully, rather than to violate, constitutional norms.¹⁷⁵ Yet the desire of police departments that courts fix and do not change fourth amendment standards represents the improper view that the exclusionary rule's purpose is to punish the culpable officer. If the legal system encourages police departments always to improve and more fully to respect constitutional values, then systemic deterrence more likely will result. The articulation by courts of fixed and unchanging constitutional standards, given the pressure of the police role, likely will convert constitutional minimums into institutional maximums. Likewise, a good faith exception will encourage departments not to counsel or advise their officers to give any more consideration for fourth amendment rights than existing case law absolutely mandates.¹⁷⁶ The good faith exception, therefore, is inconsistent with the goal of systemic deterrence.

While the exclusionary rule clearly has deterrent potential, measuring the general and systemic deterrent value of the rule is likely impossible. No statistics reveal the number of unlawful searches that the rule has deterred. But empirical evidence of the rule's deterrent effect is no weaker than evidence of the deterrent effect of laws, for example, against larceny.¹⁷⁷ As Professor Dwor-

175. Professor Amsterdam analogizes the exclusionary rule to police sponsored anti-theft programs that brand personal property. Amsterdam, *supra* note 3, at 431-32. Identification marks diminish the property's value to a prospective thief because he knows it will be more difficult to sell the goods at a worthwhile price. Although these identification programs do not influence those who steal for excitement, revenge, or personal use, diminishing the property's worth correspondingly decreases the motivation to steal it.

176. A tension, of course, exists between the need for "bright line" decisions offering police guidance, *see supra* text accompanying notes 151-53, and the value of ever changeable standards encouraging police departments to review continuously and to strive for improvement in their procedures. Bright lines demand court decisions that are not excessively fact bound. As argued earlier, however, not all bright lines are equally defensible. *See supra* note 152. Conflicts over where the courts should draw the line would continue to assure the flexibility of fourth amendment doctrine necessary to encourage continued systemic deterrence.

177. Some writers suggest the case supporting the exclusionary rule's deterrent effect is the stronger of the two. *E.g.*, Amsterdam, *supra* note 3, at 431.

kin observed: "Deterrence is partly a matter of logic and psychology, largely a matter of faith. The question is never whether laws do deter, but rather whether conduct ought to be deterred. . . ." ¹⁷⁸ Certainly, then, the deterrent value of the exclusionary rule deserves as much respect as does the deterrent value of the criminal law generally.

C. *The Good Faith Exception's Effect on the Scope of the Fourth Amendment*

Thus far this Article has shown how the exclusionary rule both offers appropriate vindication for fourth amendment violations and has deterrent value. Its critics, therefore, only can argue that the rule adversely affects the police's ability to fight crime. Accordingly, this section discusses these alleged adverse effects while developing insights into the good faith exception's effect on the scope of the fourth amendment.

A portion of society perceives the exclusionary rule as hamstringing police and allowing guilty defendants to go unpunished. Statistical studies, however, suggest this perception is faulty. First while the exclusionary rule only functions to suppress evidence at trial, our criminal justice system focuses more on obtaining guilty pleas than on securing convictions at trial. Seventy to ninety percent of prosecuted cases result in guilty pleas. ¹⁷⁹ The criminal justice system probably will not notice any police misconduct in these cases. ¹⁸⁰ One plausible interpretation of such statistics is that they show only that prosecutors screen out and never prosecute a significant number of cases that contain illegal searches and seizures. Another study, however, refutes this interpretation. After monitoring felony case processing in five cities, the study found that due process violations, including but not limited to fourth amendment violations, accounted for one (1) to nine (9) percent of the cases rejected through prosecutor screening. ¹⁸¹ "While it may be that po-

178. Dworkin, *supra* note 61, at 333.

179. Y. KAMISAR, W. LAFAVE & J. ISRAEL, *MODERN CRIMINAL PROCEDURE* 22 (5th ed. 1980); Allen, *Central Problems of American Criminal Justice*, 75 *MICH. L. REV.* 813, 819 (1977).

180. Police misconduct often may go unnoticed even by defense counsel who may spend only enough time with his client to ask whether or not the client is guilty. One study in which public defenders represented defendants, for example, indicated that 30% of the defendants reported their attorney spent less than 10 minutes with them; 32% stated 10 to 29 minutes; 27% stated one-half hour to 3 hours; and only 14% stated more than 3 hours. J. CASPER, *CRIMINAL COURTS: THE DEFENDANTS PERSPECTIVE IV* (1978) (abstract).

181. Brosi, *A Cross-City Comparison of Felony Case Processing* 18-19 (1979). Groups

lice do not arrest some suspects because of search and seizure limitations," the report concluded, "these percentages seem to counter the conventional wisdom that Supreme Court decisions cause many arrests to fail because of technicalities In felony cases other than drugs, less than 2% of the rejections in each city involved abrogation of due process."¹⁸² While the exclusionary rule's effect on the government's decisions to prosecute appears deminimus, its effect upon trial outcomes is even smaller. The same Comptroller General's study that Justice White cited in *Gates* to support a good faith exception¹⁸³ noted that of the 2,804 defendants charged in thirty-eight United States attorney offices during a two month period only 30% of the cases included a search or seizure and only 11% of the defendants filed motions to suppress. Courts denied these motions in the overwhelming majority of cases, so that in only 1.3% of the sampled criminal cases did the courts actually exclude evidence.¹⁸⁴ Furthermore, over half of the cases in which the court granted the motion to suppress still resulted in conviction.¹⁸⁵

Even though these statistics belie the popular impression that the exclusionary rule severely undermines law enforcement, the rule's supporters cannot disregard the existence of popular misperception. Institutional legitimacy is a matter of perception whether or not the perception is correct. If the exclusionary rule's interference with police work even in a small number of cases significantly undermines public confidence in the criminal justice system, the rule is a fitting subject of concern. Consequently, the actual relationship between the exclusionary rule and limitations on police behavior needs exploration.

Soon after *Mapp* extended the exclusionary rule to state criminal proceedings, New York City Police Commissioner Murphy wrote an article complaining:

I can think of no decision in recent times in the field of law enforcement which had such a dramatic and traumatic effect as this I was immediately caught up in the entire program of reevaluating our procedures . . .

such as the U.S. Department of Justice and the Law Enforcement Assistance Administration commissioned the study.

182. *Id.* at 19. Prosecutors encounter search and seizure problems most commonly in crimes, such as possession of narcotics, that lack traditional complainants. *See supra* note 45.

183. *See supra* note 48.

184. COMPTROLLER GENERAL, IMPACT OF THE EXCLUSIONARY RULE ON FEDERAL CRIMINAL PROSECUTION II (Apr. 19, 1979).

185. *Id.* at 13.

modifying, amending, and creating new polices and new instructions for the implementation of *Mapp* . . . Retraining sessions had to be held from the very top administrators down to each of the thousands of foot patrolmen. . . .¹⁸⁶

Obviously *Mapp* caused New York police to consider dramatic behavior changes. But why such dramatic changes proved necessary is unclear. Fourth amendment standards should have limited New York police since the *Wolf* decision, which preceded *Mapp* by twelve years.¹⁸⁷ *Mapp* merely changed the remedy for violations; the decision should not have changed the substantive standards of the fourth amendment with which police already were required to comply. Police Commissioner Murphy's statement demonstrates that, prior to the implementation of the exclusionary rule, police departments were indifferent to fourth amendment restrictions on police conduct and that the suppression doctrine does have systemic deterrent value.

The contention that the exclusionary rule "handcuffs" police is simply wrong.¹⁸⁸ Rather than the remedy chosen to vindicate violated fourth amendment rights, the interpreted scope of the fourth amendment itself mandates actual limitations on police conduct.¹⁸⁹ Consequently, the public's perception that the exclusionary rule hamstringing police in fulfilling their crime fighting re-

186. Murphy, *Judicial Review of Police Methods in Law Enforcement: The Problem of Compliance by Police Departments*, 44 *TEX. L. REV.* 939, 941 (1966).

187. See *supra* text accompanying note 26.

188. The assertion that the exclusionary rule allows the criminal to go free because the constable has blundered is misleading. Such an assertion implies that had the constable been more careful the court would have convicted the criminal. Mertens & Wasserstrom, *supra* note 3, at 366 n.5. Often, however, this is not the case. If the officer, lacking probable cause to arrest or search, abided by fourth amendment dictates, the criminal still would have gone free because no arrest or search would have taken place. In such contexts the constable *must* have blundered for any prosecution to occur at all.

189. As long ago as 1938 Senator Robert Wagner of New York noted before his state's constitutional convention that the right, rather than the remedy, mandates police restraint: All the argument [that the exclusionary rule will handicap law enforcement] seems to me to be properly directed not against the exclusionary rule but against the substantive guarantee itself. . . . It is the [law of search and seizure], not the sanction, which imposes limits on the operation of the police. If the [law of search and seizure] is obeyed as it should be, and as we declare it should be, there will be no illegally obtained evidence to be excluded by the operation of the sanction.

. . . It seems to me inconsistent to challenge the exclusionary rule on the ground that it will hamper the police, while making no challenge to the fundamental rules to which the police are required to conform.

1 *New York Constitutional Convention, Revised Record* 560 (1938), reprinted in J. MICHAEL & H. WECHSLER, *CRIMINAL LAW AND ITS ADMINISTRATION* 1191-92 (1940). See Oaks, *supra* note 3, at 754.

sponsibility is not only statistically but also conceptually incorrect.¹⁹⁰

Furthermore, an internal analytical inconsistency exists in condemning the exclusionary rule on the one hand for not deterring police misconduct and on the other for hamstringing police in their crime fighting efforts. If the courts chose an alternative fourth amendment remedy that more effectively deterred illegal searches and seizures, then the same or even a greater number of guilty individuals would go free, because without such arrests and searches police would discover even less evidence upon which prosecutors could base conviction¹⁹¹ or gain a guilty plea. Presently, for example, a defendant's lack of standing to protest the admission of incriminating evidence obtained in violation of someone else's privacy right allows some use of illegally seized evidence at trial.¹⁹² If a different remedy more effectively had deterred the offending officer, he would not have committed the illegal search, he would not have found the incriminating evidence, and the defendant more likely would have avoided conviction. Consequently, a remedy more efficient at deterring illegal police behavior than the exclusionary rule as presently applied also would restrict more severely the police's ability to gain convictions.¹⁹³

In addition to the loss of evidence that contributes to convictions and guilty pleas, the police no longer would make many searches that presently serve to recover stolen goods and confiscate contraband. Nor could they respond as easily to community demands for increased enforcement. Consequently, replacement of the exclusionary rule with a more effective police deterrent greatly would impede some practices law enforcement officers presently find beneficial in controlling crime.¹⁹⁴

190. One critic of the exclusionary rule credits it with the defeat of effective gun control law. See Wilkey, *supra* note 3, at 224. But the fourth amendment itself, not the exclusionary rule, prevents police from searching anyone with a bulging pocket. Furthermore, no reason exists to believe that the community's criminal element refrained from carrying guns until the Court decided *Mapp* in 1961.

191. This argument assumes that police could not find alternative means to obtain evidence. If legal alternatives did exist, however, and the police still acted illegally, release of criminals would not be a product of overly restrictive constitutional standards but of police ineptness.

192. See *supra* note 91.

193. Evidence also would not be available for impeachment purposes or for submission to a grand jury. See *supra* text accompanying notes 38-39.

194. Society, in fact, may saddle police with performing tasks that it wants accomplished but not condoned. Professor Paul Chevigny suggests that:

For legislators and judges the police are a godsend, because all the acts of oppres-

The ramifications of adopting a more effective remedy therefore would be extensive.¹⁹⁵ Because police would make no unlawful searches or arrests, many crimes would appear totally unsolved. Although courts would free fewer defendants on what the public perceives as legal technicalities, a reduction would occur in the overt demonstration of police capability, for example, to solve crime, to protect the public, and to retrieve stolen property. The more effective remedy could result in increased public satisfaction with the visible functioning of the courts and decreased approval of the visible functioning of the police. The courts would appear to be dealing effectively with those prosecuted; by contrast, the police would appear as unequal to the task of apprehending criminals deserving prosecution.

Obviously, public dissatisfaction with the exclusionary rule occurs because the rule functions after an unlawful arrest or search already publicly has unveiled a wrongdoer. The rule, thus, openly flaunts the price of the fourth amendment;¹⁹⁶ in fact, it "rubs our noses in it."¹⁹⁷ Critics of the exclusionary rule, thus, in reality are critics of the fourth amendment itself. They apparently believe that sufficient discredit of the rule could lead to judicial interpretations either further limiting the rule's scope or replacing it with a less, rather than a more, effective remedy.¹⁹⁸ A forthright demand for a less effective remedy clearly would focus the resulting conflict not on the remedy for illegal intrusions of privacy by police, but on whether the law should deem illegal and deter such intrusions at all. This demand thereby would entail a frontal attack on the val-

sion that must be performed in this society to keep it running smoothly are pushed upon the police. The police get the blame, and the officials stay free of the stigma of approving their highhanded acts. The police have become the repository of all the illiberal impulses in this liberal society; they are under heavy fire because most of us no longer admit so readily to our illiberal impulses as we once did.

P. CHEVIGNY, *POLICE POWER* 280 (1969).

195. Imagine, for example, an effective tort remedy that required police officers violating the fourth amendment to assume significant individual damage liability. If, as exclusionary rule critics argue, fourth amendment doctrine is intolerably obscure, officers wishing to protect themselves and their families from financial hardship would be fearful of making arrests and searches that even arguably may be improper. Courts, however, upon formal review might find proper many of such chilled searches. An effective tort remedy, consequently, may chill the legally justified searches now occurring under the exclusionary rule.

196. J. KAPLAN, *CRIMINAL JUSTICE* 215-16 (1978).

197. Kaplan, *supra* note 3, at 1037.

198. Even Professor Oaks recognized that the argument concerning the exclusionary rule's "handcuffing" police put the rule's critics in the "untenable position of urging that the sanction be abolished so that [police] can continue to violate [constitutional] rules with impunity." Oaks, *supra* note 3, at 754.

ues embodied in the fourth amendment itself.

A frontal attack on the fourth amendment, however, is difficult because the Bill of Rights has held an almost mystical position in American heritage. Accepting fourth amendment values in the abstract, while whittling away their importance by developing exceptions to the remedy, avoids the blasphemy of an overt challenge to such values. Avoiding such blasphemy, in turn, increases the chances of societal acceptance of reductions in the fourth amendment's scope and significance.¹⁹⁹

Avoiding a blatant attack on the fourth amendment while reducing the significance of the rights it embodies would be the effect, if not the purpose, of a good faith exception. The attack is upon the citadel of the fourth amendment rather than simply upon the exclusionary rule. Consequently, when courts entertain adopting a good faith exception, they should consider how the exception actually retards the development and restricts the scope of the fourth amendment rather than the more comfortable issue of whether the rule has any deterrent value in specific situations. In anticipation of this evaluation, this Article now will demonstrate how the proposed exception is a subtle and, consequently, highly dangerous attack on the viability of the fourth amendment.

1. Retarding Fourth Amendment Development

A good faith exception will either slow or freeze fourth amendment development "dead in its tracks."²⁰⁰ Under such an exception evidence obtained by a police officer when acting upon a reasonable good faith belief that his conduct was lawful is admissible whether or not his behavior was, or was not, constitutional. Consequently, the preliminary issue of judicial concern will be whether the officer acted under such a reasonable good faith belief. When a court finds the existence of such a belief, the actual constitutionality of the officer's behavior, thus, becomes irrelevant. Courts using

199. A literary example of the procedure's use as a tool of conflict settlement that conceals disagreements of substantive values is Shirley Jackson's short story, *The Lottery*, in *THE LOTTERY* 291 (1949). The society Jackson describes in her story has no predilection against individual sacrifice for collective goals. In fact, the society prefers such an arrangement. The story describes a communal ceremony wherein lots are drawn to determine whom the community will stone to death for some unspecified community need. Although the eventual winner of the lottery objects, she couches her objection in terms of procedure—that the lots were drawn too quickly—rather than in terms of the substance of the activity. *Id.* at 299.

200. *United States v. Peltier*, 422 U.S. 531, 554 (1975) (Brennan, J., dissenting). See LaFave, *supra* note 3, at 354-55.

the good faith exception, therefore, will have no need to further develop or refine fourth amendment doctrine.²⁰¹ No articulation of the appropriate rule of conduct that police and courts should follow in future situations will be forthcoming; thus, no general or systemic deterrence will ensue. If, in fact, the officer's behavior was unconstitutional, nothing will prevent that officer or others from continuing to make the same good faith error in the future.

Justice White, in his *Gates* concurrence, rejected the inevitability of this outcome if the Court adopted the good faith exception. He insisted that if a fourth amendment case presented a novel situation in which a court had to decide the constitutionality of an officer's behavior in order to provide guidance for his and others' future actions, the court would have sufficient reason to decide whether a violation had occurred before turning to the good faith question.²⁰²

Both constitutional precedent and prudential considerations, however, suggest that this advisory opinion procedure advocated by Justice White is not likely to occur.²⁰³ Consider, for example, the Supreme Court's response to the Ninth Circuit's decision in *Bowen v. United States*.²⁰⁴ While refusing to retroactively apply the standards articulated in *Almeida-Sanchez* for roving border patrol searches,²⁰⁵ the Ninth Circuit first found the challenged search illegal under those standards.²⁰⁶ The Supreme Court agreed that retroactive application was not necessary but then warned,

This Court consistently has declined to address unsettled questions regarding the scope of decisions establishing new constitutional doctrine in cases in which it holds those decisions nonretroactive This practice is rooted in our reluctance to decide constitutional questions unnecessarily *Because this reluctance is in turn grounded in the constitutional role of the federal courts* . . . the district courts and courts of appeal should follow our practice, when issues of both retroactivity and application of constitutional doctrine

201. Justice Brennan observed that if, under a general good faith exception, evidence is admissible unless clear precedent, not subject to reasonable misinterpretation, declares the search in question unconstitutional, "the first duty of a court will be to deny the accused's motion to suppress if he cannot cite a case invalidating a search or seizure on identical facts." *Peltier*, 422 U.S. at 554.

202. *Gates*, 103 S. Ct. at 2346.

203. Whether Justice White would wish such a procedure actually to develop is not clear. Within the same paragraph discussed in the text, he observes that: "It is not entirely clear to me that the law in this area has benefited from the constant pressure of fully-litigated suppression motions." *Id.*

204. 422 U.S. 916 (1975).

205. For a discussion of *Almeida-Sanchez*, see *supra* text accompanying notes 97-98.

206. *United States v. Bowen*, 500 F.2d 960 (9th Cir. 1974).

are raised, of deciding the retroactivity issue first.²⁰⁷

Even assuming the Court did not mean literally that Article III limitations prohibit a court from addressing such constitutional issues, prudential considerations alone probably would lead most courts first to decide the good faith issue and, thus, avoid unnecessary judicial interference in effective law enforcement.²⁰⁸

Even assuming Justice White's position ultimately proves correct, however, recognition of a good faith exception still would restrict severely fourth amendment development. According to the Justice's theory, courts could articulate new fourth amendment standards while refusing to exclude evidence in a particular case because the officer reasonably could not have anticipated the change. But if a good faith error does not lead to suppression of the illegally obtained evidence, a defense attorney would have little incentive to make the new and creative argument that might support recognition of a new standard. Criminal defense counsels try to make arguments that immediately benefit their clients rather than to urge abstract and conceptual points of law for the future benefit of others.²⁰⁹ Proposing radically new approaches or considerations in search and seizure cases would not further criminal defense strategy. Consequently, defense counsels are not likely to argue cases that raise a good faith exception at a level that would allow Justice White's theory to function.²¹⁰

207. *Peltier*, 422 U.S. at 920 (emphasis added).

208. Some jurists forthrightly have acknowledged this fact. "It is no sufficient objection that [a good faith exception] would require courts to make still another determination" writes Judge Henry Friendly. "[R]ather, the recognition of a penumbral zone where mistake will not call for the drastic remedy of exclusion would relieve them of exceedingly difficult decisions whether an officer overstepped the sometimes almost imperceptible line between a valid arrest or search and an invalid one." Friendly, *The Bill of Rights as a Code of Criminal Procedure*, 53 CALIF. L. REV. 929, 953 (1965).

209. Mertens & Wasserstrom, *supra* note 3, at 451. In response to this concern, Justice White proposed allowing application of a new fourth amendment standard to the party in whose case the rule is first announced. *Illinois v. Gates*, 103 S. Ct. at 2347 n.19. Depending upon the interpretation of the Justice's proposal, one of two flaws is obvious. If the Justice means that newly announced fourth amendment rules will not receive general retroactive application, this is already the case, see *United States v. Johnson*, 457 U.S. 537 (1982), and a good faith exception adds nothing. If he means to motivate attorneys to make new arguments by permitting exclusion of evidence for some but not all good faith errors, determining which errors should lead to suppression becomes a matter of judicial whim rather than fairness to law enforcement authorities. That very fairness to police, however, is the precise rationale allegedly justifying the exception. Thus, under this interpretation Justice White's proposal undercuts the exception's validity.

210. An alternative scenario is possible. A good faith exception might result in endless judicial debate over whether an officer should have anticipated a specific interpretation of the law. What constitutes a "reasonable mistake of law" is exceedingly difficult to deter-

In the past the exclusionary rule has functioned to assure that courts' continually reassess and refine fourth amendment protections.²¹¹ Some procedure must exist that allows courts to review claims of right and articulate constitutional doctrine.²¹² A good faith exception severely would reduce the need of courts to consider and adopt new fourth amendment standards. Stripped of this responsibility to develop new standards²¹³ or subtly refine those already recognized,²¹⁴ the courts no longer will offer the guidance necessary for general or systemic deterrence of behavior that violates the fourth amendment.²¹⁵ Without systematic deterrence, of course, the importance of and respect for the fourth amendment in this society gradually will diminish.

Previously, this Article presented *United States v. Peltier* as possible precedential support for a general good faith exception.²¹⁶ The above discussion, however, provides the analysis necessary to distinguish *Peltier*. The *Peltier* court simply refused to apply ret-

mine. Kaplan, *supra* note 3, at 1044. Cf. *Smith v. Lewis*, 13 Cal. 3d 349, 357, 530 P.2d 589, 593 (1975) (finding malpractice when attorney failed to perform adequate research to advise his client on an uncertain point of law); *Smith v. St. Paul Fire & Marine Ins. Co.*, 366 F. Supp. 1283 (M.D. La. 1973) (courts uneasy about finding that an attorney has malpracticed when failing to anticipate changes in the law). Instead of a forthright confrontation over what standards should exist, judicial attention would be focused upon the meaning of prior decisions articulating prior standards. This scenario, consequently, also would impede fourth amendment development.

211. See *Oaks*, *supra* note 3, at 756.

212. Although *Wolf* in 1949 declared that the fourth amendment bound the states, no court cases attempted to determine the contour of the states' responsibility under the Constitution until after *Mapp* imposed the exclusionary rule on the states in 1961.

213. If the fourth amendment standard proposed at trial differs dramatically from or fills a gap in prior fourth amendment doctrine, the officer reasonably could claim a good faith error. If evidence is then admissible, the court need not decide whether the officer's conduct is constitutionally improper. Under such an exception the Court would not have had a chance to decide the following landmark decisions, which dramatically deviated from prior case law: *Payton v. New York*, 445 U.S. 573 (1980); *Delaware v. Prouse*, 440 U.S. 648 (1979); *Chimel v. California*, 395 U.S. 752 (1969); *Katz v. United States*, 389 U.S. 347 (1967); *Berger v. New York*, 388 U.S. 41 (1967).

214. If the officer's behavior in a specific case is objectionable only if the court makes subtle distinctions from prior case law, the officer again may claim a reasonable good faith error. Doctrinal contours thus will remain unclear if courts have no opportunity or incentive to refine doctrine through the process of repeated application. Yet, scholars traditionally deem this type of development highly appropriate for judicial institutions. *Levi, An Introduction to Legal Reasoning*, reprinted in G. CHRISTIE, JURISPRUDENCE 964 (1963). *But see supra* note 176 and text accompanying notes 151-53.

215. Exposure to flesh and blood cases, which force courts to learn experientially, properly discourage courts from relying on rarefied abstractions. A good faith exception would diminish such valuable exposure. See O.W. HOLMES, THE COMMON LAW 5 (Howe ed. 1963) ("The life of the law has not been logic: it has been experience.").

216. See *supra* text accompanying notes 95-98.

roactively a fourth amendment standard newly articulated in *Almeida-Sanchez*. Cases such as *Peltier* that acknowledge the existence of newly articulated constitutional standards, but then refuse to apply them retroactively²¹⁷ do not impede the development of constitutional doctrine and do not destroy the general and systemic deterrence provided by the new standards. Rather, such cases encourage novel and creative approaches by reducing the societal cost of judicial acceptance of new constitutional standards. Consequently, the ramifications of non-retroactivity decisions and of a good faith exception are considerably different. Given such different effects, *Peltier* is unconvincing precedential support for a good faith exception.

2. Erosion of Fourth Amendment Doctrine

Besides retarding fourth amendment development and diminishing the general and systemic deterrence of fourth amendment standards, the good faith exception also would erode the fourth amendment's already articulated substantive doctrines. Consider, for example, its effect on the requirement of probable cause. Probable cause is the linchpin of fourth amendment protection for it marks the point at which the qualified individual right to privacy must succumb to the community's need for security and law enforcement. Probable cause to arrest exists when the facts as known would lead a reasonably cautious individual to believe that the person police are to arrest has or is committing an offense.²¹⁸ The law requires the same quantum of evidence before it authorizes a full search except the conclusions are somewhat different. Probable cause for a search requires a reasonable belief that the items sought relate to criminal activity and that police will find the items at a specific time in the place the police are to search.²¹⁹ A finding of probable cause for purposes of arrest or search, thus, does not require certainty but merely the reasonableness of the officer's or, in the case of a warrant, the magistrate's belief. The concept of probable cause, therefore, already subsumes the possibility of a reasonable error.

Consequently, for a court to find that a police officer or magis-

217. See, e.g., *Desist v. United States*, 394 U.S. 244 (1969); *Johnson v. New Jersey*, 384 U.S. 719 (1966); *Tehan v. United States ex rel Shott*, 382 U.S. 406 (1966); *Linkletter v. Walker*, 381 U.S. 618 (1965).

218. See *Draper v. United States*, 358 U.S. 307 (1959); *Brinegar v. United States*, 338 U.S. 160 (1949).

219. Y. KAMISAR, W. LAFAYE & J. ISRAEL, *supra* note 179, at 268.

trate lacked probable cause it must believe that the specific arrest or search was unreasonable. If a court applied the reasonable good faith exception under these circumstances, it necessarily would have to determine whether the offending official, who admittedly held an unreasonable belief that probable cause existed, had a reasonable good faith belief that probable cause did exist. The paradox is evident: how can one reasonably believe an unreasonable belief? Applying the reasonable good faith exception to the exclusionary rule in this context clearly requires a determination that the official acted reasonably, with probable cause, in the first place.²²⁰ Thus, courts using such an exception would be imposing covertly a far more permissive interpretation of probable cause. Once again, the exception is reducing the scope of the fourth amendment's already articulated standards rather than merely providing an innocuous exception to the remedy chosen to vindicate fourth amendment rights.

3. Interfering With the Sixth Amendment Right to Counsel

The good faith exception also would subtly and insidiously erode the importance and effectiveness of the sixth amendment right to counsel. Ever since the early 1960's, the Supreme Court has recognized that the aid of counsel in a criminal trial is essential to ensure both a fair trial²²¹ and proper protection of the individual liberty interests at stake.²²² The defense attorney, then, as well as the prosecution, is crucial to society's adversarial system of criminal justice. In a 1963 report by the Attorney General's Committee on Poverty and the Administration of Federal Criminal Justice²²³ the Committee recognized that,

The essence of the adversary system is challenge. The survival of our system of criminal justice and the values which it advances depend upon a constant, searching, and creative questioning of official decisions and assertions of authority at all stages of the process. The proper performance of the defense function is thus as vital to the health of the system as the performance of the

220. Consequently, the dissent in *People v. Quintero*, as argued previously, *see supra* text accompanying notes 159-62, is attacking the majority's determination that probable cause for a search did not exist.

221. *See Gideon v. Wainwright*, 372 U.S. 335 (1963).

222. *Scott v. Illinois*, 440 U.S. 367 (1979), held that a court only need grant counsel to a defendant on trial for a misdemeanor if, upon conviction, the court in fact sentenced him to a period of incarceration. The Supreme Court found that defendants only sentenced to fines have no such constitutional right. *See also Argersinger v. Hamlin*, 407 U.S. 25 (1972).

223. Commentators often refer to the report as *The Allen Report*, after the Chairman of the Committee, Professor Francis A. Allen.

prosecuting and adjudicatory functions.²²⁴

The defense attorney, thus, is valuable not merely because he ensures a fair trial and protects his client's liberty interests, but also because he provides the open criticism and analysis of governmental behavior necessary for a healthy criminal justice system. While defending his client, he also exposes governmental improprieties and, thus, aids in the control of the controllers.²²⁵

Seemingly, therefore, defendant's sixth amendment right to counsel concerns more than the presence or appointment of someone labeled "defense counsel"; the individual so labeled also must function effectively as counsel. Ineffective counsel may result either from lack of attorney diligence or from a system structured to prevent defendants from benefitting from attorneys' efforts.

Over twenty years ago Justice Douglas spoke of a criminal defendant's right to have and benefit from the efforts of an "imaginative lawyer."²²⁶ A good faith exception would diminish such a right. If an imaginative lawyer has not convinced any prior court that a new fourth amendment standard is necessary, the offending officer in a given case only need claim that he in good faith did not anticipate the proposed perspective in order to deprive defense counsel of any opportunity to bring a creative and imaginative argument to the court's attention.²²⁷ If the officer acted in good faith the evidence would remain admissible. No matter how capable and convincing is defense counsel's presentation, the good faith exception will deprive defendant of the benefit of his attorney's argument unless prior defendants had equally capable counsel.²²⁸ Defendants should not need to rely on anyone other than their own attorney for vindication of their constitutional rights. A system that re-

224. REPORT OF THE ATTORNEY GENERAL'S COMMITTEE ON POVERTY AND THE ADMINISTRATION OF FEDERAL CRIMINAL JUSTICE (1963), reprinted in Y. KAMISAR, W. LAFAYE & J. ISRAEL, *supra* note 179, at 53.

225. Some jurists and scholars reject this system regulation role of defense counsel and advise a more complete focus on the criminal justice system's truth-finding function. Frankel, *The Search for Truth: An Umpireal View*, 123 U. PA. L. REV. 1031, 1055 (1975). Criminal law, thus, would seek to control only the potential deviant. For a discussion of the law's need also to control the controllers, see *supra* text accompanying notes 57-58.

226. *Chewning v. Cunningham*, 368 U.S. 443, 447 (1962).

227. As previously asserted, defense counsel will have no incentive to develop or propose new and creative arguments that do not inure to his client's benefit. See *supra* notes 209-10 and accompanying text.

228. In fact a danger exists of an infinite recursion, in which an officer's claim of good faith error defeats defense counsel's arguments, which courts in turn similarly had dismissed on good faith grounds when raised in earlier cases. This process would prevent further development of the law.

quires otherwise interferes with the fundamental right to counsel.

A conceptual critique of the good faith exception, thus, reveals that the exception is premised erroneously on the inability of the exclusionary rule to deter illegal police behavior in good faith error circumstances and ignores that a constitutional violation goes unremedied. Further, as this Article has shown, the very premise of the exception is faulty because it neglects the general and systemic effects of the rule in good faith error contexts. More significantly, however, recognition of the exception severely would restrict the scope of the fourth amendment and potentially would freeze further development of fourth amendment standards. Additionally, while the exception's adverse effect on the sixth amendment right to effective counsel is subtle, the effect is substantial. All these factors make a strong conceptual case against recognition of a good faith exception. The practical problems associated with implementation of the exception, however, also are extensive and require exploration.

IV. THE DIFFICULTIES ASSOCIATED WITH IMPLEMENTATION OF A GOOD FAITH EXCEPTION

The practical problems associated with implementation of a good faith exception should be sufficient to discourage its adoption. The judicial process used to determine whether a court should apply the exception will diminish the importance and scope of the fourth amendment as well as encourage police to manipulate the criminal justice system.

A. Defining Reasonable Good Faith

The proposed good faith exception has both a subjective and objective component. The subjective component requires an officer honestly to believe his actions lawful. Since the primary justification for the exception is the exclusionary rule's inability to deter an officer who mistakenly regards his behavior as constitutional, the subjective aspect of "good faith" lies at the heart of the exception. The objective component, the requirement that the error be reasonable, attempts to avoid placing a premium on ignorance and poor training. Without the objective aspect, protection from governmental intrusion would be no greater than that observed by the least knowledgeable and least diligent member of the law enforce-

ment team.²²⁹ Both the subjective and objective components, however, pose extensive problems that require discussion.

The following hypothetical provides a basis for illustrating the difficulties associated with the good faith exception's subjective component. Assume the existence of an area of search and seizure law that courts have not defined clearly and that lacks authoritative determination by the Supreme Court.²³⁰ Given the law's uncertain condition, reasonable individuals could disagree whether certain police behavior is legal in a specific context. Assume that an officer confronts a situation within this nebulous area, and that he reasonably and in good faith believes a search would be legal, performs the search, and discovers illegal drugs. Another officer confronts the identical situation. He, in contrast to the first officer, reasonably believes a search would be illegal.²³¹ He, however, performs the search and also uncovers illegal drugs. A court, reviewing the second officer's behavior and recognizing the officer's bad faith, will eliminate any good faith assertion and find the officer's behavior unconstitutional.²³² Consequently, the court will exclude the drugs at the second defendant's trial. At the first defendant's trial, however, the same court, reviewing the first officer's behavior and recognizing that the officer reasonably believed his behavior was legal, will admit the drugs into evidence. Therefore, although both defendants have had their fourth amendment rights violated in precisely the same manner, under a system recognizing a good faith exception the same court will vindicate one defendant's violated right while allowing the other's to go unremedied.²³³

These incongruent results, mandated by the good faith exception, pose two immediate problems. First the inconsistent treat-

229. "A proper legal definition of 'good faith' involves not only a lack of malevolence but also a reasonable effort to comply with the law." *Commonwealth v. Sheppard*, 387 Mass. 488, _____, 441 N.E.2d 725, 738 (1982) (Liacos, J., concurring). See *Illinois v. Gates*, 103 S. Ct. 2317, 2340-47 (White, J., concurring).

230. For example, the extent to which police may use the "drug courier profile" to justify a forcible stop, search, and seizure is an area of law that similarly lacks authoritative determination by the Court. See *Florida v. Royer*, 103 S. Ct. 1319 (1983); *Reid v. Georgia*, 448 U.S. 438 (1980); *United States v. Mendenhall*, 446 U.S. 544 (1980).

231. Each officer's respective police department may have trained him: one to view such behavior as lawful, the other to view it as improper. Again, both departments would be interpreting the law reasonably although they disagree.

232. Of course, this assumes the court could overcome the evidentiary difficulty that finding subjective bad faith entails. See *infra* text accompanying notes 242-54.

233. Thus, the argument that the officers' different motivations justify different remedies is untenable. No alternative remedy exists for an individual whose rights police violated while acting on the basis of a good faith error. See *supra* text accompanying notes 141-44.

ment of two identically culpable defendants whose fundamental rights police violated²³⁴ in the same exact way is intuitively discomfiting. In such a situation the importance of fourth amendment norms seems unverified and the public's perception of courts acting as defenders of the faith, protectors of constitutional liberties, is severely shaken. Furthermore, a court's failure to follow the principle insisting that it treat like situations alike will violate the public's expectation of proper judicial functioning.²³⁵ Justice Harlan recognized the importance of this principle nearly fifteen years ago in his *Desist v. United States* dissent.²³⁶ In *Desist*, the Court refused to apply the new fourth amendment standard articulated in *Katz v. United States*,²³⁷ which required a warrant for telephone taps, to cases in which police misbehavior occurred prior to the *Katz* decision. The Court's rationale was that the *Katz* ruling conceivably could deter only subsequent searches, thus justifying exclusion only in such later situations. Justice Harlan dissented and insisted that *Katz* should apply to all cases not yet final regardless of the date of police misconduct. "If a 'new' constitutional doctrine is truly right," he asserted

we should not reverse lower courts which have accepted it; nor should we affirm those which have rejected the very arguments we have embraced. Anything else would belie the truism that it is the task of this Court, like that of any other, to do justice to each litigant on the merits of his own case. It is only if our decisions can be justified in terms of this fundamental premise that they may properly be considered the legitimate products of a court of law rather than the commands of a super-legislature.²³⁸

Although Justice Harlan was unable to convince a majority of the Court at the time, the Court belatedly accepted his dissent in the 1982 decision of *United States v. Johnson*.²³⁹ In *Johnson* the Court applied the newly recognized standard of *Payton v. New York*,²⁴⁰ which required police to obtain an arrest warrant before entering a suspect's home to make a routine felony arrest, to all cases pending on direct appeal at the time the Court decided *Payton* regardless of the date of police misconduct. Relying heavily on

234. See *Alderman v. United States*, 394 U.S. 165, 174 (1969) (fourth amendment rights are personal rights); *Wolf v. Colorado*, 338 U.S. 25, 33 (1949) (fourth amendment expresses a fundamental liberty of individuals binding upon the state through the operation of the due process clause of the fourteenth amendment).

235. See L. FULLER, *THE MORALITY OF LAW* 33-41 (1964).

236. 394 U.S. 244 (1969).

237. 389 U.S. 347 (1967).

238. 394 U.S. at 259 (Harlan, J., dissenting).

239. 457 U.S. 537 (1982).

240. 435 U.S. 573 (1980).

Justice Harlan's dissent in *Desist*, the Court in *Johnson* recognized the importance of the Court's behaving in a manner giving credence to itself and support to constitutional rights whose scope and violation the Court had acknowledged.²⁴¹ The inconsistent judicial reactions to violated fundamental rights generated by the good faith exception jeopardizes the same interests the Court intended *Johnson* to protect.

Conditioning the evidence's admissibility on an officer's nonculpable *subjective* state of mind existing at the time of his search reveals the second problem posed by the earlier hypothesized situation.²⁴² Under the good faith exception an officer truly must believe his conduct was lawful to avoid suppression of the discovered evidence.²⁴³ The trial court hearing the suppression motion, therefore, seemingly must determine the beliefs of the officer at the time he seized the evidence. If a law enforcement officer's averment of good faith seems reasonable, however, defense counsel will find the claim nearly impossible to refute.²⁴⁴ Consequently, hearings to ascertain the subjective intent and actual knowledge of the officer will be open invitations to perjury.²⁴⁵ The officer's adjustment of the facts in order to obtain conviction of a defendant he truly believes guilty is understandable given his engagement "in

241. If the fourth amendment protected citizens from only malevolent governmental intrusions upon privacy then only one of the two earlier hypothesized defendants, *see supra* text accompanying notes 230-33, would have a justified claim of right violation. Remedy in only that one case then would be proper. But motive is not relevant to the issue of whether police have infringed a citizen's fourth amendment rights. As suggested by Justice Marshall's dissent in *Mohile v. Bolden*, 446 U.S. 55 (1980), a focus on a tainted motive, while possibly justified when dealing with suspect classification and societal gratuities, is conceptually irrelevant when the issue concerns fundamental rights. *See id.* at 113-16 (Marshall, J., dissenting). Although closing a public swimming pool may be unlawful only if done with racial discriminatory intent, a ban on leafletting would be equally invalid under the first amendment whether desires to prevent littering or to suppress speech motivated the ban. *See Schneider v. State*, 308 U.S. 147 (1939).

242. *See supra* text accompanying notes 230-33.

243. If the officer had any doubt, the exclusionary rule's deterrent potential would justify the rule's application.

244. *See, e.g.,* Footo, *supra* note 132; Theis, "Good Faith" as a Defense to Suits for Police Deprivations of Individual Rights, 59 MINN. L. REV. 991 (1975).

245. Former Judge Irving Younger, a one-time federal prosecutor, frankly observed that "[e]very lawyer who practices in the criminal courts knows that police perjury is commonplace." Younger, *The Perjury Routine*, 204 THE NATION 596, 596 (1967). Ironically, opponents of the exclusionary rule generally criticize it for encouraging police distortion of facts. *See Rosenblatt & Rosenblatt, A Legal House of Cards*, HARPER'S, July 1977, at 18, 20 ("While intended to curb abuse of police power, the exclusionary rule has opened up a whole new field of police misconduct: perjury."). *See also* Kaplan, *supra* note 3, at 1032.

the often competitive enterprise of ferreting out crime."²⁴⁶ Police likely view allowing a criminal to escape "justice" as a much greater societal danger than a simple reconstruction of facts at a suppression hearing.²⁴⁷ While the possibility of perjury exists under the exclusionary rule, perjury is even easier under the subjective component of the good faith exception. Consequently, rather than alleviating this exclusionary rule weakness, the proposed exception exacerbates it. The exception, thus, merely adds another layer of potential police perjury to the suppression hearing: a layer most difficult to counteract because only the officer knows with certainty his subjective intent.

A presumption that all officers intentionally will commit perjury is not necessary, however, before the exception's critics can show that its subjective component is impractical. People selectively perceive, interpret, and recall their sensory impulses based upon their interests and experiences.²⁴⁸ They seldom want to hear, see, or remember that which is contrary to their needs. Furthermore, ideas, opinions, and positions that coincide with an individual's own interests or that appeal to his half-submerged prejudices are difficult for him to reject as untrue.²⁴⁹ Given the strength of an officer's identification with the interests and needs of the police subculture, if a particular subjective belief is necessary to allow use of the fruits of a search in a criminal trial, an officer will not have difficulty convincing himself of its existence.²⁵⁰ Whenever a good faith error reasonably explains police misconduct, an officer will seldom engage in such misconduct without simultaneously experiencing a strong, even if erroneous, belief in the lawfulness of his behavior.²⁵¹ Consequently, the good faith exception compounds the overt problem of perjury with the covert problem of selective perception.

The Supreme Court is not oblivious to the dangers of having fourth amendment exclusionary issues revolve around the subjective belief or intent of individual officers. As recently as 1978, when the Court was evaluating police conduct relating to wiretaps, it

246. *Johnson v. United States*, 333 U.S. 10, 14 (1948).

247. See J. SKOLNICK, *supra* note 61, at 215; LaFave, *supra* note 152, at 154; Younger, *supra* note 245, at 596.

248. See Baker, *Scope of the First Amendment Freedom of Speech*, 25 U.C.L.A. L. REV. 964, 974 (1978).

249. Wellington, *On Freedom of Expression*, 88 YALE L.J. 1105, 1130 (1979).

250. See LaFave, *supra* note 152, at 154.

251. Cf. Amsterdam, *supra* note 3, at 437.

stated that the judiciary should use a "standard of objective reasonableness without regard to the underlying intent or motivation of the officers involved."²⁵² Furthermore, ten years earlier, three Justices expressed their uneasiness with courts even attempting to determine police subjective intent and vehemently asserted that "sending state and federal courts on an expedition into the minds of police officers would produce a grave and fruitless misallocation of judicial resources."²⁵³ If most courts feel the same queasiness regarding the propriety of judicial determinations of subjective intent,²⁵⁴ judicial review of the subjective component of a claimed good faith error will be perfunctory at best.

Courts may avoid the aforementioned pitfalls of the good faith exception's subjective component if they exclusively stress the exception's objective component. If courts emphasize the exception's objective aspect, then the police behavior would be in good faith whenever it is consistent with a reasonable interpretation of prior doctrine. The officer's subjective intent or belief, therefore, becomes irrelevant. Such a purely objective good faith standard, however, also encourages police trickery and system manipulation. For example, if an officer doubts whether he has sufficient evidence to obtain a search warrant, he may proceed without the warrant in the hope that a trial judge will find his misbehavior reasonable in an effort to admit incriminating evidence.

Most unfortunately, whenever the law in an area is unsettled, an objective good faith standard will provide the incentive for officers to choose the perspective that most jeopardizes fourth amendment rights.²⁵⁵ Police would not protect individual privacy

252. *Scott v. United States*, 436 U.S. 128, 138 (1978).

253. Ironically, Justice White, presently the Court's strongest spokesman for a good faith exception, authored this assertion that an effort to determine police subjective intent would misallocate judicial resources. *Massachusetts v. Painten*, 389 U.S. 560, 565 (1968) (White, J., joined by Harlan and Stewart, J.J.) (dissenting from a dismissal of certiorari as improvidently granted).

254. Similar concerns over judicial determinations of subjective beliefs have led criminal courts frequently to reject mistake of law as a defense to criminal prosecution. Professor Lon Fuller discussed the danger of recognizing such a defense:

The required intent is so little susceptible of definite proof or disproof that the trier of fact is almost inevitably driven to asking, "Does he look like the kind who would stick by the rules or one who would cheat on them when he saw a chance?" This question, unfortunately, leads easily into another, "Does he look like my kind?"

L. FULLER, *supra* note 235, at 72-73. Courts may ask the same questions in suppression hearings with answers biased by the judges' affinity for the police officer's goal of crime control. *See infra* text accompanying notes 267-70.

255. *United States v. Peltier*, 422 U.S. 531, 559 (1975) (Brennan, J., dissenting); LaFave, *supra* note 3, at 352. Under a good faith exception, the system will encourage foot-

beyond the standards clearly mandated by case law because such protection would seem unwarranted and inconsistent with law enforcement goals.²⁵⁶ Consequently, judicially articulated constitutional minimums likely would become departmental maximums. A legal system and police culture that encouraged behavior directed predominantly toward crime fighting probably would influence even those officers who otherwise would not consciously decide to avoid doing more than they legally must.²⁵⁷ In contrast to the good faith exception, the exclusionary rule at least tempers the law enforcement system's tendency to discourage concern over fourth amendment rights by encouraging the police to err on the side of greater protection when the law is ambiguous.²⁵⁸ Precisely in such cases individuals most need the exclusionary rule.

Consequently, although the subjective component of the good faith exception is fraught with difficulties, under a purely objective standard every legal ambiguity or uncertainty constitutes an avenue for system manipulation and subterfuge by police. Consider, for example, the circumstances in *Brewer v. Williams*.²⁵⁹ Police were to transport defendant Williams to Des Moines, Iowa after

dragging as officials will have little incentive, in close cases, to err on the side of respect for constitutional liberties. See *United States v. Johnson*, 457 U.S. 537, 561 (1982).

The Court in *Johnson* rejected the government argument that courts never should apply decisions resolving unsettled fourth amendment questions to cases that involve police conduct which preceded the definitive decision. The Court's concerns, expressed in its decision, are also relevant to the effects of an objectively determined good faith exception: "Official awareness of the dubious constitutionality of a practice would be counterbalanced by official certainty that, so long as the Fourth Amendment law in the area remained unsettled, evidence obtained through the questionable practice would be [admissible until the rendering of an opinion] definitively resolving the unsettled question." *Id.* See *Desist v. United States*, 394 U.S. 244, 273 (1969) (Fortas, J., dissenting).

256. The exclusionary rule's existence provides an officer who believes present doctrine *insufficiently* protects fourth amendment rights with an opportunity to behave in accordance with his beliefs while justifying his greater concern to fellow officers in crime fighting terms. His behavior merely ensures that a future fickle court ruling will not reject his arrests or searches. If the court recognizes the good faith exception, however, such an officer no longer will be able to support fourth amendment interests in terms his colleagues will accept. Consequently, the proposed exception to the exclusionary rule not only creates an incentive to foot-drag and to interpret ambiguous areas of law so as to infringe upon fourth amendment rights, but also creates a disincentive to decide to act in a fashion supporting such rights by removing any and all legal insulation from peer pressure. See *Desist v. United States*, 394 U.S. 244, 276-77 (1969) (Fortas, J., dissenting) (rejecting court decisions that award "dunce caps" to those officers who try to anticipate future trends in fourth amendment doctrine).

257. This Article earlier discussed the role of selective perception and interpretation of sensory data. See *supra* text accompanying notes 248-51.

258. See LaFave, *supra* note 3, at 352.

259. 430 U.S. 387 (1977).

his arraignment for murder charges in Davenport. Police assured defendant's counsel that they would not interrogate defendant during the trip. Nevertheless the detective traveling with Williams made a "christian burial speech," suggesting that due to the worsening weather the murder victim would not receive a proper burial if police could not locate the body quickly. Williams, susceptible to such religious considerations, directed the police to the body. The Supreme Court, in a five to four opinion, found that the detectives had intentionally separated Williams from his attorney for the purpose of playing on his sensitivities and eliciting incriminating statements. Such conduct, concluded the Court, violated the defendant's sixth amendment right to counsel, and, thus, the trial court could not use the defendant's incriminating statements.

Upon retrial, defense counsel argued that the court also should not admit the evidence relating to the murder victim's body. The court should suppress such evidence, he insisted, as "fruit of the poisonous tree" since police discovered the body as the result of unconstitutional behavior. The state responded that the instituted police search of the area would have found the body eventually even without the defendant's assistance and, thus, the court properly had admitted the evidence under an "inevitable discovery doctrine."²⁶⁰ The Iowa Supreme Court agreed that such evidence was admissible but only if the state could prove that (1) the police did not act in bad faith for the purpose of hastening discovery of the evidence in question, and (2) that police would have discovered the evidence by lawful means.²⁶¹

For purposes of this Article, the court's evaluation of bad faith is highly informative. The court held,

While there can be no doubt that the method upon which the police embarked in order to gain William's assistance was both subtly coercive and purposeful, and that its purpose was to discover the victim's body, we cannot find that it was in bad faith. The issue of the propriety of the police conduct in this case . . . has caused the closest possible division of views in every appellate court which has considered the question. In light of the legitimate disagreement among individuals well versed in the law of criminal procedure who were given the opportunity for calm deliberation, it cannot be said that the actions of the police were taken in bad faith.²⁶²

The Iowa Court resolved the issue of good faith simply by noting that jurists, even Supreme Court Justices, reasonably had dis-

260. *State v. Williams*, 285 N.W.2d 248, 255-56 (Iowa 1979).

261. *Id.* at 260.

262. *Id.* at 260-61.

agreed on the constitutionality of the police behavior.²⁶³ That the detective clearly tried to manipulate the defendant while out of his counsel's presence and to trick him into incriminating himself was of no significance. Although the *Williams* case does not involve a fourth amendment issue it reveals the likely outcome if courts emphasize the objective component of the good faith exception to avoid the dangers of the exception's subjective component. *Williams* illustrates how the objective component threatens police abuse. In turn, courts can remove this threat of abuse only by determining police intent and once again evaluating police subjectivities.²⁶⁴ The implementation problems associated with a good faith exception, thus, are insoluble. Whether the courts focus on the exception's objective or subjective aspect, implementation of the good faith exception will result in diminished protection of fourth amendment rights. The institutional setting in which courts will determine the existence of good faith exacerbates this danger.

B. Increased Pro-Police Bias of Suppression Hearings

The good faith exception inevitably would convert suppression hearings into swearing contests. Police would assert that their beliefs are honest while the defense would question their integrity. Trial court judges, thus, would have to make suppression decisions on the basis of the least reliable or determinative kind of evidence. Additionally, the need to define a "reasonable" mistake of law would impose upon suppression judges the responsibility to make exceedingly difficult determinations on a regular basis. The trial court, therefore, would have tremendous discretion in ruling whether or not the police engaged in the objectionable behavior in

263. If a good faith exception extended beyond the fourth amendment to all motions for evidence suppression, the Iowa court's interpretation of good faith would have ended the need for the Supreme Court's decision in *Brewer* in the first place. The incriminating statements, obtained under a "good faith" error, would have been admissible. Additionally, if the Court decided the issue of good faith first, the Court would not have needed to determine whether the detective's behavior was proper. Sixth amendment doctrine in this area, consequently, would remain unclear and police could repeat identical mistakes in "good faith."

264. Following a second conviction, *Williams* sought federal habeus corpus relief. Judge Arnold of the Eighth Circuit rejected the perspective of the Iowa decision:

The question before us is not whether the Supreme Court's opinion in *Brewer v. Williams* is fairly debatable as a legal matter. Obviously it is. The question is rather what was in [the detective's] mind during [the] ride back to Des Moines *The relevant question—bad faith—is subjective* [Here the question is whether] the closeness of a later judicial decision necessarily establish[es] that conduct approved by a minority of judges is not in bad faith? We think not.

Williams v. Nix, 700 F.2d 1164, 1170 (8th Cir. 1983) (emphasis added).

good faith. Given that the lower judiciary "has hardly been very trustworthy in this area" of police credibility and misbehavior,²⁶⁵ this basically unfettered discretion likely will result, because of the fact finding propensities of state trial courts,²⁶⁶ in almost automatic admission of illegally obtained evidence whenever the officer claims good faith.

Even critics of the exclusionary rule acknowledge the anti-exclusionary rule bias of trial judges.²⁶⁷ The judges' continuing relationship with police and prosecutors, their discomfort with refusing to admit relevant and apparently reliable evidence,²⁶⁸ their interaction with crime victims, and their susceptibility to community pressure to punish the guilty affect their perception leaving them functionally and psychologically allied with police in the prosecution of crime.²⁶⁹ Since the above factors leave trial courts reluctant to interfere with police crime fighting efforts, under a good faith exception such courts likely will be eager to believe officers claiming an honestly held belief in the lawfulness of the search or seizure in question.²⁷⁰

Decisions that reflect this bias rarely are susceptible to appellate review because appellate courts are disinclined to review trial court findings of fact.²⁷¹ Furthermore, determinations of credibility, frequently based on factors such as witness appearance or atti-

265. Kaplan, *supra* note 3, at 1045.

266. Appellate courts arguably have constitutionalized whole areas of criminal procedure precisely because of the trial court's fact finding propensities. For example, Professor Amsterdam insists that the Supreme Court's fear of the trial court's pro-police bias shaped the confession cases. Amsterdam, *The Supreme Court and the Rights of Suspects in Criminal Cases*, 45 N.Y.U. L. REV. 785, 803-10 (1970). Lower courts consistently had resolved the "swearing contest" over what took place during police interrogation in favor of the police. Cf. Schaefer, *Federalism and State Criminal Procedure*, 70 HARV. L. REV. 1, 7 (1956).

267. *E.g.*, Wilkey, *A Call for Alternatives to the Exclusionary Rule: Let Congress and the Trial Courts Speak*, 62 JUDICATURE 351, 356 (1979). Some argue that the exclusionary rule pressures the courts to water down standards for proper searches and arrests in order to avoid the rule's effect of "freeing obviously guilty defendants." Barrett, *Personal Rights, Property Rights, and the Fourth Amendment*, 1960 SUP. CT. REV. 46, 66. See Ashdown, *supra* note 3, at 1294, 1315 (courts now use the privacy concept of *Katz*, originally meant to expand the scope of fourth amendment protection, to draw "unrealistic distinctions in favor of law enforcement"); Schlesinger, *supra* note 3, at 405.

268. See Schwartz, *Stop and Frisk (A Case Study in Judicial Control of the Police)*, 58 J. CRIM. L., CRIMINOLOGY & POLICE SCI. 433, 448-49 (1967); Wilkey, *supra* note 267, at 356.

269. Amsterdam, *supra* note 266, at 792.

270. See Kaplan, *supra* note 3, at 1038-39.

271. Appellate judges are better able to confront dispassionately fourth amendment exclusion issues than are trial judges since the drama of trial does not color the appellate deliberation.

tude which cannot be duplicated in a record for appeal, are virtually immune from review. Since the good faith exception's subjective component requires a trial court determination of police credibility, it gives trial judges, already opposed to the exclusionary rule, a virtually unreviewable means of avoiding the rule's impact.²⁷²

Trial courts' rulings on the objective component of the good faith exception also are likely to reflect an anti-suppression bias. The police error is reasonable if the facts and circumstances of the specific case as found by the trial court reasonably could justify the police action under the then existing case law. Yet reasonableness, arguably, has little more usefulness for analytical purposes than indicating consistency within a given system of values and belief.²⁷³ Given the role of selective perception and the similarity of police and trial judge attitudes about criminal law enforcement, judicial findings of fact that make an officer's error reasonable should not be surprising. If the error's reasonableness, *in the context of the facts found*, is initially for the trial court to determine, in practice appellate courts will defer to the factual findings of trial courts and trial courts will defer to the factual testimony of police.²⁷⁴

Under the good faith exception the traditional pro-police bias of suppression hearings not only would be insulated from appellate review, it likely would be aggravated. Under the present exclusionary rule trial judges determine whether an officer's judgment was wrong in a basically amoral, technical and mechanical sense. The good faith exception, however, changes the issue from whether an officer violated a fourth amendment right to whether he deserves moral reproachment. A court will be far more open to finding an officer "wrong" than to finding him "bad."²⁷⁵ Thus, a good faith exception to the exclusionary rule frames the exclusion issue in terms, and focuses its resolution at judicial levels, least conducive

272. See Mertens & Wasserstrom, *supra* note 3, at 449-50.

273. Weyrauch, *Taboo and Magic in Law*, 25 STAN. L. REV. 782, 800 (1973).

274. Amsterdam, *supra* note 3, at 394.

275. Courts and law enforcement agencies much more calmly accepted the Supreme Court decisions in *United States v. Wade*, 388 U.S. 218 (1967) (requiring presence of defense attorney at post-indictment lineup), and *Massiah v. United States*, 377 U.S. 201 (1964) (requiring presence of defense attorney at post-indictment interrogations of defendant), for example, than they did *Miranda v. Arizona*, 384 U.S. 436 (1966). *Wade* and *Massiah* seemed to revolve around the technical requirement of counsel for all post-indictment evidence producing contacts between government and defendant. *Miranda*, on the other hand, admonished police for trickery, perjury, and actions "destructive of human dignity," 384 U.S. at 457. *Miranda* was harder to accept precisely because it questioned the professionalism and integrity of the police. For further discussion, see Ingber, *supra* note 145, at 287-95.

to the sympathetic development of fourth amendment rights.

C. The Unconvincing Nature of the Strongest Case for a Good Faith Exception

Considering the good faith exception in its strongest context illustrates the practical problems of its implementation. Exclusion of evidence seems least fair when law enforcement officers reasonably have relied on a judicially-issued search or arrest warrant later found defective for lack of probable cause. When officers have tested the existence of probable cause by dutifully obtaining a search warrant from a judge or magistrate, and have executed the warrant according to its terms, exception proponents point out that exclusion of evidence can have no conceivable effect upon the officers' behavior.²⁷⁶ Particularly when the magistrate also made his determination of probable cause, although erroneous, reasonably and in good faith, little justification seems to exist for excluding the discovered evidence. Upon closer examination of this context, however, the argument for a good faith exception remains unconvincing.

The Supreme Court long has manifested a constitutional preference for arrests²⁷⁷ and searches²⁷⁸ made pursuant to a warrant. The Court has touted the ability of the warrant process to protect fourth amendment interest because the process ensures that a neutral and detached magistrate²⁷⁹ makes the probable cause determination in an informed and deliberative manner.²⁸⁰ Jurists perceive such a process of "judicial impartiality"²⁸¹ as clearly preferable to one in which police or prosecutors, engaged in the competitive enterprise of ferreting out crime,²⁸² reach such decisions hurriedly,²⁸³ and judicial officers review the decisions only after the fact and by hindsight judgment.

The actual existence of neutral and detached magistrates making impartial deliberative determinations, however, is highly

276. In *Gates*, in which Justice White made his strongest pro-exception argument, police officers, arguably, had so obtained and executed a search warrant. *Illinois v. Gates*, 103 S. Ct. 2317, 2344-47 (1983).

277. See, e.g., *Beck v. Ohio*, 379 U.S. 89, 91, 93-96 (1964).

278. See, e.g., *United States v. Ventresca*, 380 U.S. 102, 106-07 (1965).

279. *Johnson v. United States*, 333 U.S. 10, 14 (1948).

280. *Aguilar v. Texas*, 378 U.S. 108, 110 (1964) (quoting *United States v. Lefkowitz*, 285 U.S. 452, 464 (1932)).

281. *United States v. Jeffers*, 342 U.S. 48, 51 (1951).

282. See *Johnson v. United States*, 333 U.S. 10, 14 (1948).

283. See *Aguilar v. Texas*, 378 U.S. 108, 110 (1964).

doubtful. In most states the public elects magistrates to office.²⁸⁴ Thus, the magistrates are susceptible to public pressure to more effectively fight crime.²⁸⁵ In many states magistrates need not be, and many rural magistrates are not, lawyers.²⁸⁶ Such lay magistrates often rely on the advice of the legally educated prosecutor in determining whether to issue a warrant.

Urban magistrates, on the other hand, usually are trained lawyers and, thus, seemingly could avoid dependency on prosecutorial advice. They, however, have such extensive caseloads²⁸⁷ that normally they have little time to adequately review a warrant application. Consequently, such overworked urban magistrates, as their rural counterparts, often give warrants perfunctorily relying on the judgment of police and prosecutors to evaluate the need and propriety of the request.²⁸⁸

When a magistrate's psychological and perceptual allegiance to police crime fighting goals²⁸⁹ couples with his susceptibility to public pressure and his reliance on police and prosecutor probable cause evaluations, his neutrality and detachment in issuing war-

284. Y. KAMISAR, W. LAFAVE & J. ISRAEL, *supra* note 179, at 8.

285. In other states governors or local governmental officials appoint magistrates for a fixed term of office. *Id.* Political considerations make it unlikely that those seen as "easy on crime" will receive appointments.

286. *Id.* See also *Shadwick v. City of Tampa*, 407 U.S. 345 (1972) (upholding the constitutionality of lay magistrates). Federal magistrates, on the other hand, must be attorneys. Some federal magistrates, however, serve only part time while also practicing law. If a federal magistrate seeks clients in the community where he presides he may feel pressured to portray an image of an official protecting the community from crime rather than one overly sensitive to fourth amendment interests.

287. The responsibilities of these urban judges often include presiding over misdemeanor trials, minor civil cases, and preliminary felony proceedings. In a busy urban court the caseload per judge can reach 2000 per year. Y. KAMISAR, W. LAFAVE, & J. ISRAEL, *supra* note 179, at 9. The result is often "assembly line justice." "With their crowded facilities, lack of adequate staffing and routinized procedure, these magistrate courts create an atmosphere . . . that is more appropriate for a supermarket than a court of justice." *Id.*

288. "How can a magistrate be more than a 'rubber stamp' in signing warrants," questioned Professor Barrett,

unless he devotes at least some minutes in each case to reading the affidavits submitted to him in support of the request for a warrant, and inquiring into the background of the conclusions stated therein? And where is the judicial time going to be found to make such inquiries in the generality of cases? The Los Angeles Municipal Court with annual filings of about 130,000 (excluding parking and traffic) finds itself so pressed that in large areas of its caseload it averages but a minute per case in receiving pleas and imposing sentence. How could it cope with the added burden that would be involved in the issuance of warrants to govern the approximately 200,000 arrests [much less searches] made per year in Los Angeles for offenses other than traffic?

Barrett, *Criminal Justice: The Problem of Mass Production*, in *THE COURTS, THE PUBLIC, AND THE LAW EXPLOSION* 85, 117-18 (H.W. Jones ed. 1965).

289. See *Amsterdam*, *supra* note 266, at 792.

rants are highly questionable. Magistrates with needs and interests similar to law enforcement agents likely will perceive as reasonable those steps that such agents perceive as reasonable.²⁹⁰ Not surprisingly, therefore, in the context of an *ex parte* review of a warrant application, magistrates rarely fail to find probable cause to issue the warrant.²⁹¹

Even if a magistrate is doubtful whether probable cause exists, his reliance on police and prosecutor evaluations and his belief that defense counsel adequately can test the validity of his probable cause determination by a motion to suppress, make it likely that he will err on the side of greater rather than lesser police authority. Yet, once police obtain a warrant they tend to be less cautious. Furthermore, courts direct suppression hearing judges, whom magistrates rely upon to correct magistrate error, to defer to a magistrate's probable cause determination whenever possible.²⁹² Additionally, appellate judges are reluctant to overrule two prior judicial determinations.²⁹³ Responsibility for neutral and detached probable cause determinations, thus, passes from hand to hand. The warrant process, therefore, ultimately relies on the initial

290. See *supra* text accompanying notes 248-51.

291. Scholars widely have recognized the pervasive tendency for magistrates to grant warrant applications without adequate independent evaluation. See, e.g., Amsterdam, *supra* note 3, at 471-72 n.532; see also W. LAFAYE, ARREST—THE DECISION TO TAKE A SUSPECT INTO CUSTODY 30-36 (1965); T. TAYLOR, TWO STUDIES IN CONSTITUTIONAL INTERPRETATION 48 (1969); LaFave and Remington, *Controlling the Police: The Judge's Role in Making and Reviewing Law Enforcement Decisions*, 63 MICH. L. REV. 987, 991-95 (1965). In *Lo-Ji Sales, Inc. v. New York*, 442 U.S. 319 (1979), the Supreme Court castigated a magistrate for becoming a member of the police search party. While few magistrates would cooperate with police quite as overtly as the magistrate did in *Lo-Ji Sales*, institutional comradery may function covertly, but effectively, in merging the interest of police, prosecutor, and magistrate.

The problems associated with warrant grants do not suggest that warrants are of no value whatsoever. The pre-search affidavit in support of a warrant application ordinarily deprives the officer of the benefit of hindsight coloring the reasonableness of the search and seizure after it has produced evidence in support of itself. See *United States v. Ross*, 456 U.S. 798, 828-29 (1982) (Marshall, J., dissenting); see also W. LAFAYE, *supra* at 296-97; J. SKOLNICK, *supra* note 61, at 214-15.

292. "[T]he duty of a reviewing court is simply to ensure that the magistrate had a 'substantial basis for . . . conclud[ing]' that probable cause existed." *Illinois v. Gates*, 103 S. Ct. 2317, 2332 (1983).

293. A 1981 review of then recent decisions from the District of Columbia Court of Appeals and the United States Court of Appeals for the District of Columbia Circuit, for example, found no decision holding a warrant insufficient. Mertens & Wasserstrom, *supra* note 3, at 456. Judicial questioning of magistrate judgments is much more difficult than questioning police judgments. Magistrates are, of course, members of the same "trade union" as are judges. Judges are unlikely to question one of their own since they themselves are adverse to superior courts' questioning and overruling them.

probable cause determinations of police and prosecutors: the very judgments the legal system's architects designed the process to oversee.²⁹⁴

This practical rather than mythical perception of the warrant process belies its neutrality and detachment. Recognition of a good faith exception in this context merely would compound the problem. Under the exception whenever "well trained" officers and magistrates reasonably could differ on whether probable cause exists, evidence seized under the defective warrant remains admissible.²⁹⁵ Well trained officers, however, also are aware that a substantial disparity exists among magistrates as to the extent of evidence necessary before any given magistrate will issue a warrant.²⁹⁶ Thus, the good faith exception encourages an officer or prosecutor unsure whether probable cause for an arrest or search exists to "shop around" for a lenient magistrate likely to err only in the direction of effective law enforcement.²⁹⁷

Magistrate shopping, however, met with overt judicial disapproval in *United States v. Davis*.²⁹⁸ In *Davis*, a Treasury agent and an assistant U.S. attorney requested a federal magistrate to issue a warrant on the basis of an affidavit. The magistrate denied this request. The following day the same agent and assistant attorney presented the same affidavit to a second magistrate who issued the warrant. The federal district court found such magistrate shopping highly improper and concluded that the first magistrate's decision was binding and prevented a second magistrate from issuing a warrant on the exact same showing.²⁹⁹ A good faith exception covertly will encourage the magistrate shopping *Davis* overtly condemns. If the Court recognizes the exception, it merely will lead officers and prosecutors to approach initially those magistrates reputed to be more sympathetic to their needs, as was the second magistrate in *Davis*.³⁰⁰ Law enforcement officers in the jurisdiction need never repeat the agent's error in *Davis*; they merely can bypass the first

294. See Amsterdam, *supra* note 3, at 394.

295. See *Illinois v. Gates*, 103 S. Ct. at 2346 (White, J., concurring).

296. L. TIFFANY, D. MCINTYRE & D. ROTENBERG, *DETECTION OF CRIME* 204 (1967).

297. See *id.* at 120.

298. 346 F. Supp. 435 (S.D. Ill. 1972).

299. *Id.* at 442.

300. The good faith exception, thus, compounds the magistrate-shopping problem. At least in a *Davis*-type situation the second magistrate is aware of the proceeding before the first magistrate and, thus, likely will act cautiously. Under a good faith exception, however, if police or prosecutors merely choose strategically the first magistrate, that magistrate may remain unaware of the uncertainty of probable cause in the case at hand.

magistrate on questionable cases.

Ironically, although the defective warrant allegedly provides the strongest justification for a good faith exception, the exception is least necessary in this setting. Fourth amendment substantive law already recognizes that in doubtful or marginal cases a court should sustain a search with warrant when it would not without one.³⁰¹ Furthermore, the Supreme Court in *Gates* insisted that reviewing courts merely should ensure that there was a "substantial basis for . . . conclud[ing] that probable cause existed."³⁰² Courts, therefore, already are to disregard any reasonable deviation by a magistrate from some hypothetically knowable probable cause standard and sustain the resulting search and seizure as *constitutional*.³⁰³ To recognize additionally a good faith exception when a magistrate grants a warrant under circumstances in which even this strong judicial preference is insufficient to sustain it, is to condone and encourage a magistrate's gross insensitivity to fourth amendment values.³⁰⁴ Thus, even in the context its supporters most frequently use to justify a good faith exception, the exception again merely is a covert attack upon fourth amendment principles.

V. CONCLUSION: LEGITIMACY AND THE ROLE OF THE COURT

Judicial recognition of a good faith exception to the fourth amendment exclusionary rule would be more than a simple refine-

301. "[I]n a doubtful or marginal case a search under a warrant may be sustainable where without one it would fall." *United States v. Ventresca*, 380 U.S. 102, 106 (1965).

302. *Illinois v. Gates*, 103 S. Ct. 2317, 2332 (1983) (quoting *Jones v. United States*, 362 U.S. 257, 271 (1960)). "[W]e have repeatedly said that after-the-fact scrutiny by courts of the sufficiency of an affidavit should not take the form of *de novo* review. A magistrate's 'determination of probable cause should be paid great deference by reviewing courts.'" *Id.* at 2331 (quoting *Spinelli v. United States*, 393 U.S. 410, 419 (1969)).

303. *Gates* itself, consequently, may encourage magistrate shopping.

304. Some jurists have questioned whether exclusion ever is a proper remedy when a judge or magistrate rather than a police officer committed the constitutional error. *See, e.g., Illinois v. Gates*, 103 S. Ct. 2317, 2345 (1983) (White, J., concurring); *Commonwealth v. Sheppard*, 387 Mass. 488, —, 441 N.E.2d 725, 735 (1982). Yet, the system expects magistrates, even more than police, to understand court decisions and follow their directives. *See Rose v. Mitchell*, 443 U.S. 545, 563 (1979). Although the deterrent value of exclusion upon police confronted with hurried and situational choices may be questionable, *see supra* text accompanying notes 73-86, the potential effect upon magistrates, many of whom are trained in the law, is less open to doubt. The more direct effect of exclusion upon magistrates suggests they are more susceptible to the educational force of judicial actions under the exclusionary rule than are law enforcement agents. Without the possibility of exclusion in close cases, courts never may have the opportunity to identify a magistrate's error and articulate directives for magistrates to follow in future similar situations. Motions to exclude, thus, are more vital when the error is the magistrate's rather than the police officer's.

ment of remedy. The exception would serve as a well camouflaged and subtle attack upon the values and substantive meaning of the amendment itself. Courts rarely would articulate new search and seizure standards sensitive to fourth amendment interests because once a court found police good faith the constitutional determination would be irrelevant to the outcome of those very cases in which such articulation might take place.³⁰⁵ Insulated from the continuing pressure to decide new fourth amendment issues or dramatically review old ones, courts will lose the opportunity or incentive to set standards for future cases. Consequently, the disincentive to review new fourth amendment issues will reduce significantly the exclusionary rule's ability to generally and systemically deter illegal police behavior. Furthermore, the good faith exception greatly will dilute existing substantive fourth amendment protections, such as probable cause, that already factor in the possibility of reasonable error, and, thus, will leave the individual significantly more vulnerable to governmental intrusion. In addition to this clear, but often unrevealed, substantive effect upon the fourth amendment, the exception has the procedural ramification of substituting the trial for the appellate court as the critical institution for exclusion decisions. The good faith exception's focus on officer credibility combined with trial court's notorious pro-police bias likely would lead consistently to decisions sacrificing fourth amendment interests.

Of more subtle, but greater, significance than these erosions of the fourth amendment, would be the denigration of constitutional rights generally arising from a system that responds with indifference to violations of these rights.³⁰⁶ If constitutional rights were to lose their transcendental aura and the citizenry were to view them as deserving protection only when convenient or socially advantageous, the importance of a written bill of rights in an otherwise democratic society would be subject to grave doubts. The actual existence of the American Bill of Rights sufficiently sanctifies the right of privacy in this culture to justify a different approach for

305. As argued earlier, *see supra* text accompanying notes 226-28, this also may serve to inhibit a creative attorney's ability to defend his client. The good faith exception, thus, also may run afoul of the sixth amendment's right to counsel.

306. A legal system that treats the violation of one constitutional right with indifference may find it increasingly acceptable to treat the violation of other rights similarly. The fourth amendment is not the only constitutional right subject to the criticism of being excessively costly. Consequently, adopting a cost-benefit analytical mode for fourth amendment issues may lead to the use of such analysis to lessen the significance of other constitutional rights as well.

controlling law enforcement than that used by other nations.³⁰⁷

The above factors should be the Supreme Court's foremost concerns in evaluating the good faith exception during its October 1983 term.³⁰⁸ Yet, the Court has indicated its sensitivity to and

307. Whether the nations with which the United States is compared, *see supra* note 50 and accompanying text, have symbolic counterparts to the fourth amendment is unclear. Additionally, the exclusionary rule may be a valid response to a unique American experience. This country is more ethnically and racially heterogeneous than others. It also has a greater tradition of lawlessness, by citizen and official alike, and a more moralistic system of criminal law, which specifically pressures officials to disregard the privacy of citizens. Kaplan, *supra* note 3, at 1031. As suggested by Professor Phillip Johnson: "We can no more [expect to] import our solutions than we can export our problems." Johnson, Book Review, 87 *YALE L.J.* 406, 414 (1977).

308. Evaluation of the three cases originally scheduled for Supreme Court argument, *see supra* note 19 and accompanying text, is now appropriate:

(1) *Colorado v. Quintero*, ___ Colo. ___, 657 P.2d 948, *cert. granted*, 103 S. Ct. 3535 (1983), poses the issue whether a court should exclude evidence obtained by a good faith seizure of stolen property. If the seizure was reasonable in light of the officer's knowledge at the time, the seizure is based on probable cause and, thus, is constitutional. Consequently, for good faith even to be an issue in this appeal by the state prosecutor, the state courts must have found the seizure unreasonable. To allow a good faith exception in this context would constitute either a surreptitious attack upon and dilution of the reasonable ground/probable cause standard applied by the state court, *see supra* text accompanying notes 154-62, or an acceptance of a totally subjective belief standard of good faith. This Article already has described the dangers of such a subjective belief standard. *See supra* text accompanying note 234-54. Although the Court recently dismissed the writ of certiorari in this case, *see supra* note 19, if a similar situation presents itself in the future, the Court should reject both of these alternatives.

(2) *United States v. Leon*, No. 82-1093 (9th Cir. Jan. 19, 1983), *cert. granted*, 103 S. Ct. 3535 (1983), asks whether a court should exclude evidence seized in reasonable, good faith reliance on a search warrant subsequently held defective due to inadequate probable cause for its issuance. Not to exclude such evidence blatantly will encourage magistrate shopping with police and prosecutors presenting their warrant applications to magistrates reputed to be lenient or willing to rubber stamp warrant requests. *See supra* text accompanying notes 295-300. In addition, the presumptive constitutionality of searches and seizures executed under warrant already mandates courts to disregard possible marginal errors by a magistrate. *See supra* text accompanying notes 301-04. This presumption once compounded by a good faith exception again would dilute fourth amendment protections.

(3) *Commonwealth v. Sheppard*, 387 Mass. 488, 441 N.E.2d 725 (1982), *cert. granted sub nom. Massachusetts v. Sheppard*, 103 S. Ct. 3534 (1983), also pertains to seizure of evidence by police relying in good faith upon a warrant. But in this case the warrant is defective due to a magistrate error in failing to specify in the warrant itself items police were to seize. The police search, however, was no more extensive or detailed than a magistrate could have authorized under a properly drawn warrant. This case, therefore, is more equivocal.

The state supreme court held that the failure to describe the items police were to seize was a "serious omission of constitutional significance." 387 Mass. at ___, 441 N.E.2d at 733-34 n.17. The motion judge, in fact, had viewed the warrant as a "general warrant" akin to the colonial "writ of assistance" which led to the enactment of . . . the Fourth Amendment." *Id.* at 738 (Liacos, J., concurring). Although in the specific case the searching officer also was the affiant and, therefore, knew what items the warrant meant to allow police to retrieve, such is not always the case. Required itemization, thus, normally serves both to

empathy with the public's dissatisfaction with the seemingly high cost of the exclusionary rule.³⁰⁹ Concerns about its own legitimacy and that of the criminal justice system generally, likely will tempt the Court to recognize a good faith exception. Recognition of the exception, however, would not be the preferable response of the Court to the public outcry. The Court should acknowledge that the substance of fourth amendment standards rather than the exclusionary rule causes the alleged interference with police effectiveness. Perhaps as society becomes more complex and individuals and institutions become more interdependent, society must reduce privacy interests to assure sufficient public safety.³¹⁰ This nation actually may be so fearful of crime that its citizens are willing to compromise the sanctity of constitutional liberties. The rhetoric of remedy, however, should not mask the decision to make such a compromise.³¹¹ A frontal attack upon the fourth amendment at least would constitute a forthright admission that at issue is the extent to which the society is willing to protect individuals from governmental intrusion.³¹² The public and the Court must recognize overtly the values at stake in order to assure that the ensuing decisions reflects an accurate and preferred value choice.

Further, the proposition that the Court would enhance its role

give notice to an officer of the limits of his authority as well as to inform the party searched of the extent to which the law obliges him to defer to the officer's demands. Without these two ramifications, the warrant process is of little significance. Consequently, excluding such evidence fulfills general and systemic deterrence interests as well as vindication of a constitutional right.

Furthermore, in *Sheppard*, the magistrate's actions were *not* reasonable; they clearly were negligent. As Justice Liacos asserted in his separate concurrence: "The magistrate who utterly fails to describe the things to be seized under the search warrant has not made a reasonable effort to comply with the law." *Id.* Even if a good faith error is arguably appropriate when the exclusion of evidence is unlikely to have a deterrent effect on an innocent official, *Sheppard* is obviously not such a case.

309. See *supra* note 16 and accompanying text.

310. This author remains unconvinced by this allegation.

311. Legal institutions often have masked value choices through use of procedure, ceremony, or rhetoric. See J. NOONAN, *PERSONS AND MASKS OF THE LAW* (1976); Ingber, *supra* note 57, at 332-38; Weyrauch, *Law as Mask—Legal Ritual and Relevance*, 66 CALIF. L. REV. 699 (1978). While this device may avoid societal strife due to ideological conflict, see Ingber, *supra* note 145, at 266-73, it also relieves from both institutions and individuals the need to accept responsibility for the value choices in fact made. Ingber, *supra* note 57, at 325-29.

312. Some authors bemoan that the exclusionary rule has led to the narrowing of the fourth amendment's scope. See *supra* note 267. They suggest that abolishing or modifying the rule may lead to greater acceptance of a privacy right. Assuming they are correct, one must question the significance of a right accepted as an abstraction, but which upon violation results in no meaningful vindication of its claim. Narrowing of articulated constitutional protections in fact may be the only honest response to an unwillingness to continue to accept the costs that society must pay for meaningful individual liberties.

in our constitutional system by deferring to the public outcry itself is doubtful. All individual rights when placed into context are laden with societal cost and possible popular disapproval.³¹³ The heterogeneity of American society and the often irreconcilable tensions between honored values make unanimous approval of any decision impossible. Every decision, by necessity, will result in dissatisfaction by some. Consequently, no decisional institution can afford to concern itself only with increasing its popularity. The legitimacy that allows judicial institutions to be effective is significant only if, at appropriate times, those institutions use the law to educate and direct society rather than conform to and follow it. A court enforcing individual liberties must contribute something beyond what a system of popular fiat could accomplish.³¹⁴

A significant portion of the Supreme Court's legitimacy comes from fulfilling its role as protector of constitutional liberties. Whether situationally comfortable or not, the Court has played a role of moral leadership in our constitutional system.³¹⁵ The Court's adoption of a good faith exception because of public displeasure with the costs of fourth amendment rights would constitute an abdication of its leadership responsibilities. Such an abdication, in turn, would diminish the Court's importance and prestige.

Admittedly, the Court can lead society only if society is willing to follow. If, in order to protect individual liberties, the Court too frequently or too severely interferes with broad societal goals, public opinion may threaten substantially its legitimacy. But institutions such as the Court have the luxury of time at their disposal: they can lead gradually. Language and rhetoric are tools that the Court can use to slowly alter public perceptions in order to introduce new or protect old values. As Professor Anthony Amsterdam eloquently observed:

313. See *supra* text accompanying notes 100-02.

314. Many significant constitutional cases also are those that critics attacked most stridently. See, e.g., A. BICKEL, *THE LEAST DANGEROUS BRANCH* 195-98 (1962) (criticizing *Baker v. Carr*, 369 U.S. 186 (1962)); H. WECHSLER, *Toward Neutral Principles of Constitutional Law*, in *PRINCIPLES, POLITICS AND FUNDAMENTAL LAW* 36-48 (1961) (criticizing *Brown v. Board of Educ.*, 347 U.S. 438 (1954)); Baldwin & Nagan, *Board of Regents v. Bakke: The All-American Dilemma Revisited*, 30 U. FLA. L. REV. 843 (1978) (criticizing *Regents of the Univ. of Cal. v. Bakke*, 438 U.S. 265 (1978)); Ely, *The Wages of Crying Wolf: A Comment on Roe v. Wade*, 82 YALE L.J. 920 (1973) (criticizing *Roe v. Wade*, 410 U.S. 113 (1973)). Of course, at those times when law educates and leads society, it is most open to the accusation of elitism.

315. M. PERRY, *THE CONSTITUTION, THE COURTS, AND HUMAN RIGHTS* 99 (1982).

Like the Pythias cries, [Supreme Court decisions respecting suspects' and defendants' rights] have vast mystical significance. They state our aspirations. They give a few good priests something to work with. They give some of the faithful the courage to carry on and reasons to improve the priesthood instead of tearing down the temple.³¹⁶

Fourth amendment exclusionary cases may only "state our aspirations," aspirations frequently unfulfilled or unfulfillable in practice. But without such aspirations no transcendental importance to individual rights exists and their recognition slowly may decay into a matter of popular or official appreciation: a matter of costs and benefits.

Observers expect the Supreme Court's responses to the good faith proposal this term.³¹⁷ Hopefully, the Court will address the arguments presented in this Article and reject the exception. If not, virtually every state supreme court may refight the battle as defendants' attorneys ask them to interpret their state constitutions' search and seizure provisions. The citadel may yet be saved.

316. Amsterdam, *supra* note 266, at 793.

317. 52 U.S.L.W. 3201 (Sept. 27, 1983); *see supra* text accompanying notes 18-19.

VANDERBILT LAW REVIEW

VOLUME 36

NOVEMBER 1983

NUMBER 6

TIMOTHY R. DODSON
Editor in Chief

KENNETH B. COLE, JR.
Executive Editor

MICHAEL A. SULLIVAN
Senior Articles Editor

WILLIAM E. ANDERSEN
Senior Managing Editor

JOHN E. SULLIVAN
*Senior Student Writing
Editor*

DEBORAH McABEE BURNS
Articles Editor

DAVID MORGAN
Managing Editor

J. GRAHAM MATHERNE
Student Writing Editor

BRUCE KEVIN DALLAS
Articles Editor

ROBERT C. REYNOLDS, JR.
Managing Editor

PAUL C. NEY, JR.
Student Writing Editor

LINDA POTTER
Articles Editor

L. D. SIMMONS
Research Editor

Associate Editors

RANDALL D. BORUD
JOHN P. BURNS
ALFRED G. DUNCAN, III
JEFFREY S. GOODMAN

WILLIAM F. HAGERTY, IV
KEITH R. JACQUES
MARY KATHERINE KENNEDY
WILLIAM H. MATHIEU
STEVEN G. MCKNIGHT

STEVEN D. PARMAN
DREW T. PAROBK
KIMBERLY PAYNE
KIMBERLY FINCH WITHROW

Third Year Staff

G. EDWARD CASSADY, III
JEFFREY ALAN GREENE
PHILIP E. HOLLADAY, JR.

E. B. MARTIN
D. BLAINE SANDERS

Second Year Staff

ROBERT H. BEINFELD
DOUGLAS IVOR BRANDON
PATRICIA HYNES BRANDON
MELINDA LEE COOPER
KENNETH M. COX
CLARE H. DRAPER, IV
TODD S. FAIRCHILD
GRANT CHARLES GLASSFORD
JEREMY H. GRESHIN
ALVIN LOUIS HARRIS
ROBB S. HARVEY

JAMES M. HEAD, JR.
WILLIAM J. HOLLEY, II
GERALD F. MACE
GEORGE P. MCGINN, JR.
MICHAEL D. MCKIBBEN
CHARLES K. MCKNIGHT, JR.
LISA S. MESSING
JENNIE L. MICHAELS
ARTHUR E. MURPHY, III
LAURA A. NORMAN
SHEILA R. OETJEN

JEROME M. ORGAN
GREGORY R. SAMUEL
JOHN B. SCHERLING
LAURA L. SESSUMS
JERRY NEAL SMITH
RICHARD E. STABINSKI
BRIAN S. STEFFEY
MARK G. STRAUCH
STEPHANIE LEA STROMIRE
LEA WIGGINS
DAVID D. WILLOUGHBY

ROBERT N. COVINGTON
Faculty Advisor

JUDY SHARP NEAL
Executive Secretary

JEFFREY A. SCHOENBLUM
Faculty Advisor

NOTES

Human Leukocyte Antigen Testing: Technology Versus Policy in Cases of Disputed Parentage

I.	INTRODUCTION	1587
II.	HLA BLOOD TESTS	1589
	A. <i>History of Blood Tests in a Legal Setting</i> ...	1589
	B. <i>Use of Blood Tests as Evidence</i>	1592
III.	USE OF HLA BLOOD TESTS IN CASES OF DISPUTED PARENTAGE	1595
	A. <i>Paternity Determination of an Illegitimate Child</i>	1595
	1. The Necessity for Accurate and Objective Evidence	1595
	2. The Appropriate Evidentiary Weight for HLA Test Results	1598
	(a) <i>Reliability</i>	1598
	(b) <i>The Child's Best Interests</i>	1599
	(c) <i>The State's Fiscal Integrity</i>	1601
	B. <i>Paternity Disputes at Time of Divorce</i>	1602
	1. The Child's Best Interests	1605
	2. Preservation of Family Integrity	1606
	C. <i>Unwed Fathers Attempting to Establish Pa- ternity Rights</i>	1608
IV.	CONCLUSION	1611

I. INTRODUCTION

Many cases of disputed paternity in the past five years have developed an unusual focus. Instead of addressing the paternity issue in efforts to establish eligibility for welfare assistance,¹ an in-

1. The number of cases to establish an illegitimate child's paternity has increased since 1976 because of federal legislation requiring any mother of an illegitimate child who desires welfare assistance to cooperate with the state in establishing paternity. See 42 U.S.C. § 602 (1976 & Supp. V 1981). Congress' intent, therefore, is to alleviate some of the taxpay-

creasing number of recent cases have arisen in which a married father participating in a divorce proceeding claims that a child of the marriage is not his biological child² or an unwed father seeks to prove paternity in order to gain custody or visitation rights.³ The technology of blood testing has advanced dramatically with the increased number of disputed paternity cases and, although not completely accurate, this technology is better than ever before at positively identifying the father of a particular child.⁴ The Human Leukocyte Antigen (HLA) test presently can exclude a person from the list of potential fathers with ninety-five to ninety-nine percent certainty and can verify paternity with ninety to ninety-eight percent probability.⁵ Thus, the use of modern HLA testing can affect dramatically cases of disputed parentage.

State law controls the admission of HLA test results as evidence in paternity cases.⁶ Some states expressly allow admission of HLA test results through statutes,⁷ while other states permit the courts, in their discretion, to deal with this issue.⁸ In either situation, the courts enjoy a wide latitude in determining the weight that they will accord the scientific evidence.⁹ Policy considerations such as protection of the best interests of the child and preservation of family integrity often affect a court's decisions on the ad-

ers' burden by shifting the obligation of support to the illegitimate child's father.

2. See, e.g., *Balfour v. Balfour*, 413 So. 2d 1167 (Ala. Civ. App. 1982); *County of Fresno v. Superior Court*, 92 Cal. App. 3d 133, 154 Cal. Rptr. 660 (1979); *Rachal v. Rachal*, 412 A.2d 1202 (D.C. 1980); *B.S.H. v. J.J.H.*, 613 S.W.2d 453 (Mo. Ct. App. 1981).

3. See, e.g., *LaCroix v. Deyo*, 113 Misc. 2d 89, 447 N.Y.S.2d 864 (1981) (custody dispute between step-father and biological father of child after mother's death); *J.B. v. A.F.*, 92 Wis. 2d 696, 285 N.W.2d 880 (Ct. App. 1979) (custody dispute between maternal grandparents and biological father of child after mother's death).

4. Sussman & Gilja, *Blood Grouping Tests for Paternity and Nonpaternity*, 1981-1 N.Y. ST. J. MED. 343 (1981). See generally N. BRYANT, AN INTRODUCTION TO IMMUNOHEMATOLOGY 169-79 (1982) (discussion of HLA system and applications of HLA antigen typing).

5. Sussman & Gilja, *supra* note 4, at 343, 345. See also Terasaki, *Resolution by HLA Testing of 1000 Paternity Cases not Excluded by ABO Testing*, 16 J. FAM. L. 543 (1978).

6. H. KRAUSE, CHILD SUPPORT IN AMERICA, THE LEGAL PERSPECTIVE 218 (1981).

7. E.g., GA. CODE ANN. § 74-306 (1981); IND. CODE ANN. § 31-6-6.1-8 (Burns 1980).

8. E.g., *Cramer v. Morrison*, 88 Cal. App. 3d 873, 153 Cal. Rptr. 865 (1979); *Malvasi v. Malvasi*, 167 N.J. Super. 513, 401 A.2d 279 (1979); cf. *Pratt v. Victor B.*, 112 Misc. 487, 448 N.Y.S.2d 351 (Fam. Ct. 1982).

9. Imwinkelried, *A New Era in the Evolution of Scientific Evidence—A Primer on Evaluating the Weight of Scientific Evidence*, 23 WM. & MARY L. REV. 261, 270-72 (1981). See also *Berry v. Chaplin*, 74 Cal. App. 2d 652, 169 P.2d 442 (1946) (jury found actor Charlie Chaplin to be father even though blood tests excluded him); *State v. Camp*, 286 N.C. 148, 209 S.E.2d 754 (1974) (state supreme court upheld jury finding that man was natural father even though blood tests excluded him).

missibility of and the weight given to HLA test results.¹⁰

This Note, however, does not attempt to argue through evidentiary analysis that more courts should recognize statistical and mathematical probabilities of paternity that the HLA test furnishes. Instead, this Note assumes that the judicial trend of using HLA test results as affirmative evidence¹¹ will continue, and accepts the proposal that these results are scientifically reliable.¹² The Note will focus on the policy considerations and arguments that should affect the admissibility of the HLA blood test as affirmative evidence in various disputed parentage cases.

This Note first examines the use of HLA test results to determine the paternity of illegitimate children who do not have a legal father, and concludes that courts should admit the results unconditionally in these circumstances. Second, the Note analyzes the use of the HLA blood test to settle paternity disputes that arise in a divorce context. Because of the important policy considerations that exist, the Note recommends that courts admit the results subject to a legislatively developed statute of limitations of two or three years. Finally, the Note advocates that public policy factors indicate the need for a two-year statute of limitations to govern the admissibility of HLA test results in situations in which an unmarried father attempts to introduce them to assert paternity rights concerning a presumably legitimate child who has a legal father.

II. HLA BLOOD TESTS

A. *History of Blood Tests in a Legal Setting*

Dr. Karl Landsteiner discovered the major blood groupings at the University of Vienna in 1901.¹³ This discovery facilitated the first safe transfusions of human blood¹⁴ and resulted in increased

10. See J. GOLDSTEIN, A. FREUD & A. SOLNIT, *BEFORE THE BEST INTERESTS OF THE CHILD* 3-15 (1979); J. GOLDSTEIN, A. FREUD & A. SOLNIT, *BEYOND THE BEST INTERESTS OF THE CHILD* (1979). See generally Muench & Levy, *Psychological Parentage: A Natural Right*, 13 *FAM. L. Q.* 129 (1979).

11. *Pratt v. Victor B.*, 112 Misc. 2d at 489-90, 448 N.Y.S.2d at 353. See also *infra* notes 43-46 and accompanying text.

12. See *infra* note 25 and accompanying text.

13. R. RACE & R. SANGER, *BLOOD GROUPS IN MAN* 8-9 (1975). The major groupings are A, B, O, and AB. *Id.*

14. A. ERSKINE, *THE PRINCIPLES AND PRACTICE OF BLOOD GROUPING* 6-7 (1973). Dr. Landsteiner received the Nobel Prize in medicine in 1930 in recognition of his discovery of the major blood groups. *Id.*

knowledge of genetics¹⁵ and the use of blood tests as paternity determinants.¹⁶ The American Medical Association (AMA) fully endorsed and recommended the Landsteiner series of tests¹⁷ to the legal community in 1952.¹⁸ Because the tests were only fifty to sixty percent efficient in definitely proving non-paternity,¹⁹ however, states enacted statutes that allowed courts to admit results of the tests as evidence in paternity cases only if they excluded the accused father.²⁰

While such exclusionary statutes remained effective, scientific advancements in the discovery of identifying antigens in blood continued. Early blood tests had located antigens only on red

15. See generally M. LEVITAN & A. MONTAGU, *TEXTBOOK OF HUMAN GENETICS* (1971) [hereinafter cited as *HUMAN GENETICS*].

16. The discovery in 1927 by Dr. Landsteiner and his colleague, Dr. Phillip Levine, of the M-N antigen on the red blood cell proved more useful in court as a paternity test than in diminishing the risks that accompany blood transfusions. A blood group antigen is an inherited antibody on the surface of a red blood cell that determines a blood grouping reaction with specific antiserum. The genes that control development of blood group antigens vary in frequency in different racial and ethnic groups. *STEDMAN'S MEDICAL DICTIONARY* 88 (5th ed. 1982).

The M-N antigen test is a reliable genetic marker because of this high percentage of distinction in the genes of a random sample of the population. This test alone can establish nonpaternity with approximately 32% efficiency. *HUMAN GENETICS*, *supra* note 15, at 148-210. See also Page-Bright, *Proving Paternity—Human Leukocyte Antigen Test*, 27 *J. FORENSIC SCI.* 135, 139-40 (1982). See generally K. BOORMAN & B. DODD, *AN INTRODUCTION TO BLOOD GROUP SEROLOGY* 154-63, 322-23 (1970).

17. The Landsteiner series of tests refers to the ABO blood test joined with the later discovered M-N and Rh blood systems tests. The ABO test consists of identifying the blood type of parents (O, A, B, or AB) and comparing it to that of the child. Ross, *The Values of Blood Tests as Evidence in Paternity Cases*, 71 *HARV. L. REV.* 466 (1958). The M-N data proved useful in later studies of the Rh system, which Dr. Landsteiner and Dr. Alexander S. Wiener discovered in 1940. The Rh system showed that humans carry either an Rh-positive or Rh-negative antigen. In addition to explaining why persons who received a transfusion of their own blood group sometimes reacted adversely to a blood transfusion, the Rh system discovery helped explain and prevent *erthroblastosis fetalis*, a disease of the newborn. In addition, the Rh system provided another genetic marker helpful in improving the exclusion rate in a paternity test. A.G. ERSKINE, *supra* note 14, at 6. See also *HUMAN GENETICS*, *supra* note 15; J. QUEENAN, *MODERN MANAGEMENT OF THE RH PROBLEM* 1-55 (1977).

18. Davidsohn, Levine & Wiener, *Medicolegal Application of Blood Grouping Tests*, 149 *J. A.M.A.* 699 (1952).

19. K. BOORMAN & B. DODD, *supra* note 16, at 321 (citing 62% efficiency of tests); L. SUSSMAN, *BLOOD GROUPING TESTS, MEDICOLEGAL USES* 89 (1968) (citing 53% efficiency of tests).

20. *E.g.*, ALA. CODE § 26-12-5 (1975); CAL. EVID. CODE § 895 (West 1966); GA. CODE ANN. § 74-101 (1981); MD. ANN. CODE art. 16, § 66G (1981); OHIO REV. CODE ANN. § 3111.16 (Page 1980); TENN. CODE ANN. § 36-228 (1977). Georgia has amended its blood test statute and now permits the admission of HLA tests specifically as evidence to prove or disprove parentage. See authorities cited *supra* note 7.

blood cells.²¹ Tests for antigens on white blood cells, however, soon promised a higher probability of paternity exclusion because the antigen on the white cell—the human leukocyte antigen (HLA)—was more rare.²² Combining the HLA test with other blood tests achieved a ninety-one to ninety-three percent probability of paternity exclusion.²³ The more recent use of the HLA tests with antigen tests has improved the certainty of paternity exclusion to ninety-five percent, while adding the RBC enzyme and plasma protein tests increase certainty to 99.95%.²⁴ Additionally, the pooling of the results of the HLA test with those of the Landsteiner series and other blood antigen tests can lead to a significant positive statistical probability of paternity that approaches one hundred percent.²⁵

21. See STEDMAN'S MEDICAL DICTIONARY 1659-63 (5th ed. 1982). At least fifteen different red blood cell antigenic systems exist, including ABO, M-N, P, Rh, Lutheran, Kell, Lewis, Duffy, Kidd, Diego, Yt, I, Xg, Dombrock, and Colton. As of 1976, scientists had identified over 250 antigens on red blood cells. B. LEAVELL & O. THORUP, FUNDAMENTALS OF CLINICAL HEMATOLOGY 720 (1976).

22. N. BRYANT, *supra* note 4, at 172.

23. AMA & Family Law Section, ABA, *Joint AMA-ABA Guidelines: Present Status of Serologic Testing in Problems of Disputed Parentage*, 10 FAM. L. Q. 247, 257 (1976) [hereinafter cited as *AMA-ABA Guidelines*]. See also Comment, *Paternity Testing with the Human Leukocyte Antigen System: A Medicolegal Breakthrough*, 20 SANTA CLARA L. REV. 511 (1980).

24. Sussman & Gilja, *supra* note 4, at 345.

25. *Id.* at 343. In calculating the probability of paternity, the HLA test is more precise than other blood tests because it identifies one specific genetic marker, the HLA *haplotype*. Identification of this genetic marker allows more accurate determination of paternity because only approximately one person in one thousand has the same HLA *haplotype* as any given person. Terasaki, *supra* note 5, at 544. Every individual inherits two immutable *haplotypes*, one from each parent. N. BRYANT, *supra* note 4, at 171, 175-76. Four genes compose each *haplotype*, and two inherited *haplotypes* that share a location on a white blood cell are a *genotype*. Page-Bright, *supra* note 14, at 141. The genetic make-up of *genotypes* can vary, and the observable characteristics of these variations comprise an individual's genetic category of *phenotype*. STEDMAN'S MEDICAL DICTIONARY 1071 (5th ed. 1982). Since science has identified at least 6600 *phenotypes*, Lee, Lebeck & Wong, *Estimating Paternity Index from HLA-typing Results*, 74 AM. J. CLIN. PATH. 218, 218 (1980), and recognizes the theoretical probability that approximately 20 million different *phenotypes* exist, the possibility that a child and a putative father share a common *haplotype* but are unrelated is very unlikely. A. SVEJGAARD, M. HAUGE, C. JERSILD, P. PLATZ, L. RYDER, L. NIELSEN, & M. THOMSEN, THE HLA SYSTEM 8-12 (1975). See Note, *Blood Test Evidence in Disputed Paternity Cases: Unjustified Adherence to the Exclusionary Rule*, 59 WASH. U.L.Q. 977, 983-95 (1981).

In 1978, a study using HLA testing on one thousand putative fathers that the Landsteiner series test had not excluded determined the rate of error in HLA testing procedures to be 0.35%. Terasaki, *supra* note 5, at 548. The HLA test showed that 640 of the fathers whom the Landsteiner test had not excluded had at least a 90% probability of paternity, while it definitely excluded 250 of the men. The HLA test could not resolve 10% of the cases. *Id.* at 552. Thus, because of the low rate of error and the high rate of exclusion, courts

Courts also have addressed some of the test's disadvantages. The cost of the HLA test, for example, is quite high in comparison with other blood test costs.²⁶ In *Little v. Streater*,²⁷ however, the Supreme Court reduced the unequal effect of this cost by holding that an indigent male has a right to obtain blood tests, including the HLA test, if a statute setting forth the requirements for a child's receipt of welfare has caused the mother to name the indigent as a defendant in the case.²⁸ Other disadvantages of the HLA test involve difficulties with the testing itself. For example, since white blood cells generally are viable for only twenty-four to seventy-two hours after drawing the blood sample,²⁹ the HLA test must occur within that timespan. Additionally, mailing samples of blood to testing centers requires that the package arrive within this period in an insulated box that protects the sample from extreme temperatures.³⁰ These disadvantages, however, appear minor in comparison with the scientific proof and information that the HLA test provides.³¹

B. Use of Blood Tests as Evidence

Blood tests did not gain significant judicial acceptance in the United States until the 1940s,³² and even then the courts often

have begun to use the HLA test more widely in cases of disputed paternity. Two reports in which the HLA test results identified two different fathers for alleged twins further illustrate the power of the HLA test. See Majsky & Kout, *Another Case of Occurrence of Two Different Fathers of Twins by HLA Typing*, 20 TISSUE ANTIGENS 305 (1982); Terasaki, Gjertson, Bernoco, Perdue, Mickey, & Bond, *Twins with Two Different Fathers Identified by HLA*, 299 NEW ENG. J. MED. 590 (1978). HLA typing also has indicated that the husband of a married woman who had been raped was the father of the fetus with a 96% probability. This indication proved correct after the child's delivery. Pollack, Schafer, Barford, & Dupont, *Prenatal Identification of Paternity—HLA Typing Helpful After Rape*, 244 J. A.M.A. 1954 (1980).

26. Sussman & Gilja, *supra* note 4, at 344 (usual blood grouping tests cost approximately \$150, in comparison to \$600 for blood grouping tests that include the HLA test).

27. 452 U.S. 1 (1981). See *infra* section III.

28. *Streater*, 452 U.S. at 3. The Court found that requiring an indigent in such an instance to pay the high costs of the tests would violate due process since "[t]he effectiveness of the seven systems [including HLA] attests the probative value of blood test evidence in paternity cases." *Id.* at 8.

29. Keith, *Resolution of Paternity Disputes by Analysis of the Blood*, 8 FAM. L. REP. [BNA] No. 4, at 4005 n.25 (Nov. 24, 1981); N. BRYANT, *supra* note 4, at 235-36.

30. Keith, *supra* note 29, at 4005 n.41.

31. See *supra* text accompanying notes 22-25.

32. See, e.g., *Livermore v. Livermore*, 233 Iowa 1155, 11 N.W.2d 389 (1943). But cf. *Commonwealth v. Zammarelli*, 17 Pa. D. & C. 229 (1931) (first reported paternity case which used blood test results).

gave little evidentiary weight to the results.³³ Instead, the courts would adhere strictly to statutes that required blood tests to be held inconclusive on the issue of paternity. Thus, in the infamous 1946 paternity case *Berry v. Chaplin*,³⁴ the court held that Charlie Chaplin was the father of a child even though blood tests clearly excluded him as the biological father.³⁵ The result of the Charlie Chaplin case is by no means archaic; failure to grant conclusive effect to blood test evidence that definitely excluded the putative father occurred as recently as 1974.³⁶

Generally, however, standard blood tests have been important evidence in United States paternity suits since the 1940s.³⁷ Because of the Landsteiner series' low probability of successfully proving paternity,³⁸ legislatures fashioned statutes to permit the use of these tests only if they proved that a man was not the father.³⁹ The more precise HLA test, however, can serve as two types of evidence: as exclusionary evidence to show that the putative father could not be the biological father of the child,⁴⁰ and as inclusionary or affirmative evidence to indicate the high probability of paternity.⁴¹ Since most states originally fashioned blood test admissibility statutes to deal with exclusionary evidence that the Landsteiner series provided,⁴² the use of the HLA test to exclude the putative father has not created as much legal controversy as its affirmative use to prove paternity. A majority of the states which have such statutes mandate that blood tests are admissible only to

33. See *supra* note 9.

34. 74 Cal. App. 2d 652, 169 P.2d 442 (1946).

35. *Id.* at 664-65, 169 P.2d at 451. Because of the heated adverse publicity that the case attracted, the California Legislature adopted the Uniform Act on Blood Tests to Determine Paternity, CAL. EVID. CODE §§ 890-897 (West 1966). The Act provides in pertinent part: "If the Court finds that the conclusion of all the experts, as disclosed by the evidence based upon the tests, are [sic] that the alleged father is not the father of the child, the question of paternity shall be resolved accordingly." CAL. EVID. CODE § 895 (West 1966). See *infra* note 45.

36. *State v. Camp*, 286 N.C. 148, 209 S.E.2d 754 (1974). In 1975 the North Carolina Legislature amended the state statute to provide that blood test results which exclude the possibility of paternity be conclusive in effect. N.C. GEN. STAT. § 8-50.1 (1981). See also *supra* note 9.

37. See, e.g., *Jordan v. Mace*, 144 Me. 351, 69 A.2d 670 (1949) (Maine Supreme Court held that jury cannot disregard blood group tests which exclude one in paternity action); *Commonwealth v. D'Avella*, 339 Mass. 642, 162 N.E.2d 19 (1959) (if blood test excludes one as father in paternity suit, court must find in defendant's favor as matter of law).

38. See *supra* notes 16 & 21 and accompanying text.

39. See *supra* note 20 and authorities cited therein.

40. H. KRAUSE, *supra* note 6, at 218-19.

41. *Id.* at 219-42.

42. See *supra* text accompanying notes 19-21.

exclude paternity.⁴³ Eight states, however, have enacted statutes that adopt section 12 of the Uniform Parentage Act, thereby allowing admission of statistics which address the likelihood of paternity.⁴⁴ Ten other states have retained statutes from previously drafted uniform acts that permit introduction of blood tests as affirmative evidence if the court, within its discretion, deems the evidence relevant.⁴⁵ A few state courts have used judicial notice to admit evidence of statistical probability in paternity suits.⁴⁶

In those jurisdictions that have no controlling statute, courts can use the standard set forth in *Frye v. United States*⁴⁷ to admit blood test results as reliable scientific evidence. *Frye* allows a court to admit scientific evidence only if the evidence has "gained general acceptance in the particular field in which it belongs."⁴⁸ Courts additionally require stringent proof that scientific evidence is reliable because judges fear scientific test results unduly will impress and sway jurors,⁴⁹ and that the results or methods of scien-

43. See *supra* note 20 and authorities cited therein.

44. COLO. REV. STAT. § 19-6-113 (1978); HAWAII REV. STAT. § 584-12 (1976); MINN. STAT. ANN. § 257.63 (West 1982); MONT. CODE ANN. § 40-6-113 (1981); N.C. GEN. STAT. § 8-50.1 (1981); N.D. CENT. CODE § 14-17-11 (1981); WASH. REV. CODE ANN. § 26.26.110 (Cum. Supp. 1982); WYO. STAT. § 14-2-110(a)(iii) (1977). See also CAL. CIV. CODE §§ 7000 to 7021 (West 1983) (adopting Uniform Parentage Act, but not provision concerning statistical evidence of likelihood of paternity).

Section 12 of the Uniform Parentage Act provides that "[e]vidence relating to paternity may include . . . (3) blood test results, weighted in accordance with evidence, if available, of the statistical probability of the alleged father's paternity." UNIF. PARENTAGE ACT § 12, 9A U.L.A. 602 (1979). The National Conference of Commissioners on Uniform State Laws and the American Bar Association approved the Act in 1973.

45. CAL. EVID. CODE § 895 (Cum. Supp. 1981-82); ILL. ANN. STAT. ch.40 § 1401 (Smith-Hurd Cum. Supp. 1982); KY. REV. STAT. ANN. § 406.111 (Bobbs-Merrill 1970); LA. REV. STAT. ANN. § 9:397.2 (West Supp. 1982); ME. REV. STAT. tit. 19 § 280 (Cum. Supp. 1982); N.H. REV. STAT. ANN. § 522:4 (1974); OR. REV. STAT. § 109.258 (1981); 42 PA. CONS. STAT. ANN. § 6136 (Purdon 1982); R.I. GEN. LAWS § 15-8-14 (1981); UTAH CODE ANN. § 78-45a-10 (1977).

These ten states have adopted either § 4 of the Uniform Act on Blood Test to Determine Paternity (UBTA) or § 10 of the Uniform Act on Paternity (UPA). The two sections are nearly identical and § 10 of the UPA provides in pertinent part: "If the experts conclude that the blood tests show the possibility of the alleged father's paternity, admission of the evidence is within discretion of the court, depending upon the infrequency of the blood type." The National Conference of Commissioners on Uniform State Laws and the American Bar Association enacted the UBTA in 1952, and approved the UPA in 1960.

46. See, e.g., *Broadwater v. Broadwater*, 247 Md. 607, 233 A.2d 782 (1967); *Malvasi v. Malvasi*, 167 N.J. Super. 513, 401 A.2d 279 (1979).

47. 293 F. 1013 (D.C. Cir. 1923). See also *Imwinkelried*, *supra* note 9, at 262-73.

48. 293 F. at 1014.

49. E. IMWINKELRIED, *EVIDENTIARY FOUNDATIONS* 92 (1980). See also *Huntingdon v. Crowley*, 64 Cal. 2d 647, 414 P.2d 382, 51 Cal. Rptr. 254 (1966).

tific tests are faulty.⁵⁰ The HLA test, however, apparently meets the *Frye* standard because it has gained general acceptance and recognition in both legal and medical circles.⁵¹ Nevertheless controversy still surrounds the use of the HLA test as affirmative evidence in paternity suits because courts are hesitant to recognize a statistical probability of paternity. For example, in 1980 a Massachusetts court⁵² reiterated the concern of the Judicial Council of Massachusetts that litigants would raise "constitutional questions under the Fourth and Fifth Amendments . . . if . . . test results [were] used to establish the probability of paternity rather than being limited to proof of exclusion of paternity."⁵³ Other courts,⁵⁴ however, have viewed HLA test results as highly reliable admissible evidence even when state statutes did not expressly permit their use as affirmative proof of paternity. In *Pratt v. Victor B.*,⁵⁵ for instance, a New York Family Court judge wrote in dicta:

It is recommended that the current Legislature consider an amendment which would permit the use of . . . [the ABO red cell antigen series], even if it does not exclude the putative father, when offered in combination with the HLA test. Of course, there would still be no basis for using this test for proof of paternity except in combination with a nonexclusory HLA test.⁵⁶

This Note now focuses on the policy considerations that affect the admissibility of HLA test results in various disputed paternity cases.

III. USE OF HLA BLOOD TESTS IN CASES OF DISPUTED PARENTAGE

A. Paternity Determination of an Illegitimate Child

1. The Necessity for Accurate and Objective Evidence

Because a high percentage of the approximately three million illegitimate children born in the United States between 1972 and

50. E. IMWINKELRIED, *supra* note 49, at 92.

51. See *AMA-ABA Guidelines*, *supra* note 23 and accompanying text.

52. *Commonwealth v. Blazo*, 10 Mass. App. Ct. 324, 406 N.E.2d 1323 (1980).

53. *Id.* at 327 n.1, 406 N.E.2d at 1326 n.1 (paraphrasing Fifty-fifth Report of the Judicial Council of Massachusetts, PUB. DOC. NO. 144, 31, 38-40 (1979)).

54. See, e.g., *Pratt v. Victor B.*, 112 Misc. 2d 487, 448 N.Y.S. 2d 351 (1982); *J.H. v. M.H.*, 177 N.J. Super. 436, 426 A.2d 1073 (1980). See also, *Tice v. Richardson*, 7 Kan. App. 2d 509, 644 P.2d 490 (1982) (statute fails to specify type of blood test and court held that the certainty associated with HLA tests allows that evidence to be admitted to determine paternity).

55. 112 Misc. 2d 487, 448 N.Y.S.2d 351.

56. *Id.* at 490, 448 N.Y.S.2d at 353.

1979⁵⁷ received welfare assistance,⁵⁸ the United States Congress in 1975 enacted legislation that places the burden of supporting these children upon their parents, especially their fathers.⁵⁹ The Aid to Families with Dependent Children (AFDC) program requires mothers of illegitimate children to cooperate with the states in establishing their children's paternity in order to obtain support payments from fathers.⁶⁰ Unless a mother agrees to cooperate, the government can refuse to disperse funds to the family.⁶¹ Families can receive AFDC aid only if the government does not locate the child's putative father, or if a court finds that the man is not the father or is indigent.⁶² This forced cooperation requirement coupled with the large number of illegitimate children who attempt to qualify for AFDC funds⁶³ has precipitated the largest number of disputed parentage cases.⁶⁴

Recognition of the inadequacy of the other types of evidence that courts admit in disputed paternity cases demonstrates an imperative need for reliable affirmative scientific evidence in such litigation. Courts generally will admit the testimony of the mother and putative father concerning their alleged sexual relations.⁶⁵

57. BUREAU OF THE CENSUS, U.S. DEP'T OF COMMERCE, STATISTICAL ABSTRACT OF THE UNITED STATES 66 (1979).

58. For example, in Alabama, Arkansas, Florida, Georgia, Illinois, Kentucky, Louisiana, Maryland, Mississippi, Missouri, New Jersey, New York, North Carolina, Pennsylvania, Tennessee, Texas, and Virginia between 30-50% of the children receiving welfare funds were illegitimate. STAFF OF SENATE COMM. ON FINANCE, 94TH CONG., 1ST SESS., WAGE GARNISHMENT, ATTACHMENT AND ASSIGNMENT, AND ESTABLISHMENT OF PATERNITY 288 (Comm. Print 1975).

59. 42 U.S.C. § 602 (1976 & Supp. V 1981).

60. 42 U.S.C. § 602(26) (1976) provides in pertinent part that

as a condition of eligibility for aid, each applicant or recipient will be required-

. . . .

(B) to cooperate with the state (i) in establishing the paternity of a child born out of wedlock with respect to whom aid is claimed, and (ii) in obtaining support payments for such applicant and for a child with respect to whom such aid is claimed, or in obtaining any other payments or property due such applicant or such child, unless (in either case) such applicant or recipient is found to have good cause for refusing to cooperate as determined by the State agency in accordance with standards prescribed by the Secretary, which standards shall take into consideration the best interests of the child on whose behalf aid is claimed

61. *Id.*

62. 42 U.S.C. § 602 (1976 & Supp. V 1981).

63. *See supra* notes 57-58 and accompanying text.

64. *See supra* note 1 and accompanying text.

65. *See, e.g.,* *Gelinas v. Nelson*, 165 Conn. 33, 327 A.2d 565 (1973) (court admitted mother's testimony about intercourse with defendant and denial of sexual relations with other men near time of conception); *Pryor v. James*, 377 So. 2d 252 (Fla. Dist. Ct. App. 1979) (mother testified to several occasions of sexual intercourse with defendant who alleged

Such testimony leads to obviously inherent problems. The testimony not only will be self-serving, but a strong likelihood exists that parties and witnesses will commit perjury.⁶⁶ Other types of proof also invite perjury. For instance, putative fathers often will use the defense of multiple access—that the mother was sexually active with other men at or near the time of conception—to counter an allegation of paternity.⁶⁷ Other evidence that courts admit because of its “quasi-scientific” nature includes physical resemblance,⁶⁸ gestation period of the child,⁶⁹ and polygraph tests.⁷⁰ The reliability and accuracy of such evidence, however, is doubtful.

An obvious need, therefore, exists for accurate and objective evidence that does not depend on the veracity of witnesses or on “quasi-scientific” evidence. The recent Supreme Court decision in

mother had been promiscuous); *People ex rel. Raines v. Price*, 37 Ill. App. 3d 921, 347 N.E.2d 29 (1976) (parties and witnesses directly contradicted one another concerning sexual relationships of plaintiff-mother); *Collins v. Wise*, 156 Ind. App. 424, 296 N.E.2d 887 (1973) (mother testified to repeated acts of intercourse with defendant prior to and shortly after child's conception date); *Tice v. Richardson*, 7 Kan. App. 2d 509, 644 P.2d 490 (1982) (mother testified to numerous sexual relations with defendant but only one with present husband during three month period); *Pratt v. Victor B.*, 112 Misc. 2d 487, 448 N.Y.S.2d 351 (N.Y. Fam. Ct. 1982) (mother testified to having sexual relations with defendant who disputed number of times and chronological period of the alleged acts).

66. See Arther & Reid, *Utilizing the Lie Detector Technique to Determine the Truth in Disputed Paternity Cases*, 45 J. CRIM. L.C. & P.S. 213, 215 (1954).

67. *State ex rel. Leonard v. Hogan*, 32 Or. App. 89, 573 P.2d 328 (1978). See generally Sass, *The Defense of Multiple Access in Paternity Suits: A Comparative Analysis*, 51 TUL. L. REV. 463 (1977) (comparing the use of the multiple access defense in the United States, West Germany, France, Hungary, and the Soviet Union).

68. J. WIGMORE, EVIDENCE § 166, at 627 (3rd ed. 1940). Presenting “evidence” of physical resemblance usually entails examination of the child by the jury, which compares his features to those of the father. See, e.g., *McLemore v. Richardson*, 32 Conn. Supp. 533, 343 A.2d 229 (Conn. Super. Ct. 1975); *G.L. v. S.D.*, 382 A.2d 252 (Del. Super. Ct. 1977), *rev'd on other grounds*, 403 A.2d 1121 (Del. 1979); *Joseph v. State*, 149 Ga. App. 296, 254 S.E.2d 383 (1979); *Schigur v. Keck*, 93 Mich. App. 763, 286 N.W.2d 917 (1979); *Glascok v. Anderson*, 83 N.M. 725, 497 P.2d 727 (1972); *State v. Clay*, 236 S.E.2d 230 (W.Va. 1977). A minority of the states, however, do not allow the jury to examine the child. See *Esch v. Graue*, 72 Neb. 719, 101 N.W. 978 (1904); *Cook v. State*, 172 Tenn. 42, 109 S.W.2d 98 (1937); *Schlehein v. Duris*, 54 Wis. 2d 34, 194 N.W.2d 613 (1972). See generally Comment, *Evidence: Admissibility of Evidence of Resemblance, Where Paternity in Issue*, 11 CORNELL L. Q. 380 (1926) (discussion of older cases advocating use of all evidence of resemblance).

69. The alleged father usually presents evidence of gestation period to show that conception was not possible on admitted dates of sexual intercourse because the length of time between “conception” and the date of birth was beyond the normal 40 week duration of pregnancy. American courts have used evidence concerning the length of pregnancy for years. See *In re Estate of McNamara*, 181 Cal. 82, 183 P. 552 (1919) (claim of 304 days as duration of pregnancy held not normal period of gestation); *Karen K. v. Christopher D.*, 86 A.D.2d 633, 446 N.Y.S.2d 345 (N.Y. App. Div. 1982) (256 day gestation period required remanding for expert medical testimony).

70. See Arther & Reid, *supra* note 66, at 214.

*Mills v. Habluetzel*⁷¹ heightened this need. The Court held in *Mills* that a one year statute of limitations for filing a paternity action on behalf of an illegitimate child violated equal protection.⁷² Although the majority opinion did not suggest an appropriate statute of limitations for these actions, Justice O'Connor indicated in her concurring opinion that even Texas' four year statute of limitations was inappropriate.⁷³ Justice O'Connor supported her belief that statutes of limitation should not bar paternity actions involving illegitimate children who do not have a legal father during the child's minority by stating that the government has an interest not only in seeing that "justice is done"⁷⁴ but also in enhancing the state's fiscal integrity by reducing the welfare rolls.⁷⁵ Justice O'Connor also observed that the "State's concern about stale and fraudulent claims is substantially alleviated by recent scientific developments in blood testing dramatically reducing the possibility that a defendant will be falsely accused of being the illegitimate child's father."⁷⁶ Thus, if future litigants bring paternity actions five, ten, or even fifteen years after the birth of a child, the courts will need objective and reliable evidence such as the HLA test results, especially when the alternative types of evidence include possible perjury or hazy recollections.

2. The Appropriate Evidentiary Weight for HLA Test Results

(a) Reliability

Although a present and future need for trustworthy evidence such as HLA test results clearly exists in paternity suits, courts still are uncertain about the evidentiary weight they should accord affirmative, as opposed to exclusive, use of this evidence. Courts either can hold positive test results conclusive on the issue of paternity or can consider them in conjunction with other evidence.⁷⁷ A decision by the courts to deem affirmative evidence of the probability of paternity conclusive would foreclose the possibility of considering or testing another man as the possible father.⁷⁸

71. 102 S. Ct. 1549 (1982).

72. *Id.* at 1555-56.

73. *Id.* at 1556.

74. *Id.*

75. *Id.* at 1557.

76. *Id.* at 1557 n.2.

77. See *supra* note 9 and accompanying text.

78. Since the same HLA haplotype exists in only one person in one thousand, the probability that the test would reveal a 90-99% likelihood of paternity in the wrong man is

Courts, therefore, should not consider the HLA blood test conclusive until it can indicate with one hundred percent accuracy the fact, rather than the probability, of paternity. Currently, the more practical alternative is to allow the trier of fact to consider HLA blood test results in light of other evidence. Although the jurors could not base their decision strictly on the HLA evidence, the judge could instruct them to give it greater weight than the other evidence.⁷⁹ Recent cases⁸⁰ and medical reports⁸¹ that cite the accuracy of the HLA test in paternity disputes impliedly justify this type of jury instruction.

(b) *The Child's Best Interests*

In addition to the reliability of the tests, strong policy reasons also support a judicial determination to give HLA test results significant affirmative evidentiary weight in paternity actions. The two primary considerations are first, the best interests of the child⁸² and second, the fiscal integrity of the state.⁸³ A child's best interests in this context include the receipt of financial and emotional support from his father⁸⁴ and establishment of a legal relationship with his father that engenders a sense of identity and may facilitate inheritance⁸⁵ or "adoption" by the father or paternal grandparents if the mother dies during the child's minority.⁸⁶ Of course, a legal determination pursuant to HLA test results that a

highly doubtful. The possibility, however, is strong enough to create doubt in a juror's mind. See *supra* note 25.

79. See C. McCORMICK, *HANDBOOK OF THE LAW OF EVIDENCE* § 211, at 521-23 (2d ed. 1972).

80. See, e.g., *Reid v. White*, 112 Misc. 2d 294, 446 N.Y.S.2d 991 (N.Y. Fam. Ct. 1982); *State v. Meachem*, 93 Wash. 2d 735, 612 P.2d 795 (1980).

81. See *supra* note 25.

82. See *supra* note 10 and accompanying text.

83. See *infra* notes 91-96 and accompanying text.

84. See *supra* note 10. Goldstein, Freud, and Solnit emphasize the emotional needs of the child more than his financial needs.

85. See Stone, *Report of the Committee on the Law of Succession in Relating to Illegitimate Persons*, 30 *MOD. L. REV.* 552 (1967). See generally Note, *Uniform Probate Code—Illegitimacy—Inheritance and the Illegitimate: A Model for Probate Reform*, 69 *MICH. L. REV.* 112 (1970) (discussion advocating parity between legitimate and illegitimate children's inheritance rights).

86. See, e.g., *LaCroix v. Deyo*, 113 Misc. 2d 89, 447 N.Y.S.2d 864 (N.Y. Fam. Ct. 1981) (natural father and stepfather seek custody after death of minor child's mother); *Marticoarena v. Miller*, 597 P.2d 1349 (Utah 1979) (two men, each of whom had previously been married to deceased woman, sought status as biological father of child to facilitate custody); *J.B. v. A.F.*, 92 Wis. 2d 696, 285 N.W.2d 880 (Wis. Ct. App. 1979) (maternal grandparents and unwed male claiming to be natural father seek custody of deceased mother's minor child).

man must support a child financially because he is the biological father does not guarantee fulfillment of the social goal to provide the child with emotional support. The establishment of an emotional relationship between a father and child, however, is more likely to occur if the court forces a nonindigent father to build a financial tie to the child, than it is if the father never institutes any type of contact.

In a more practical vein, judicial determination of paternity can affect substantially the emotional and economic interests that accompany inheritance. Most recent Supreme Court cases have attempted to equalize the status of legitimate and illegitimate children⁸⁷ but the Court has allowed some forms of discrimination toward illegitimate children to continue in the area of inheritance law. In the 1978 case of *Lalli v. Lalli*,⁸⁸ the Supreme Court established a principle that would allow the denial of an illegitimate child's claim to any inheritance in his father's estate. In the five-to-four decision the Court specifically upheld a New York statute that required the adjudication of an illegitimate child's paternity during the father's lifetime for the child to have any claim to the father's estate via intestate succession.⁸⁹ The dissenting justices in *Lalli* argued forcefully that if a father dies intestate before judicial determination of a child's paternity then the state's intestacy laws may prohibit the illegitimate child from inheriting even though the community has common knowledge of the child's paternity or the father openly acknowledges it.⁹⁰

87. See, e.g., *New Jersey Welfare Rights Organization v. Cahill*, 411 U.S. 619 (1973) (Court invalidated on equal protection grounds a welfare statute that discriminated against illegitimate children); *Weher v. Aetna Cas. & Surety Co.*, 406 U.S. 164 (1972) (Court invalidated on equal protection grounds a workmen's compensation provision governing dishurserment of benefits because it discriminated against illegitimate children); *Levy v. Louisiana*, 391 U.S. 68 (1968) (Court struck down state wrongful death statute that discriminated against illegitimate children as violative of equal protection). *But see* *Califano v. Boles*, 443 U.S. 282 (1979) (Court upheld Social Security provision that denied insurance benefits to illegitimate children because it lacked sufficient impact to constitute discrimination); *Mathews v. Lucas*, 427 U.S. 495 (1976) (Court upheld Social Security provision denying benefits to illegitimate children not judicially recognized, as necessary for administrative convenience).

88. 439 U.S. 259 (1978).

89. *Id.* at 259 (Powell, J., Burger, C.J., and Stewart, J., announcing judgment); *id.* at 276 (Stewart, J., concurring); *id.* at 276 (Blackmun, J., concurring).

90. *Id.* at 277 (Brennan, J., White, J., Marshall, J., and Stevens, J., dissenting). The dissenting Justices pointed out that the father not only had acknowledged his illegitimate child openly, but also had provided financial support. The dissent also expressed doubt that a person who failed to write a will would have the cognizance to adjudicate his paternal relationship with an illegitimate child for purposes of insuring that the child would inherit

(c) *The State's Fiscal Integrity*

In addition to promoting children's best interests, the courts also can use HLA blood test results as affirmative evidence that can enhance a state's fiscal integrity.⁹¹ Under the AFDC program a mother cannot receive welfare assistance for an illegitimate child until she helps the state institute a paternity suit against the natural father.⁹² Since the incidence of illegitimacy is steadily increasing,⁹³ states will save money if the fathers, rather than the states, support these children.⁹⁴ One state supreme court expressed this policy as follows: "The State has a compelling interest in assuring that the primary obligation for support of illegitimate children falls on *both* natural parents rather than on the taxpayers of this state."⁹⁵ Other state courts have echoed this sentiment.⁹⁶ Another fiscal reason for judicial use of affirmative HLA blood test results is the possibility of avoiding unnecessary costly litigation if such use will expedite admissions of paternity by natural fathers and acceptance by them of financial responsibility.⁹⁷

Thus, strong public policy considerations favor use of the HLA blood test as affirmative evidence in paternity determinations of illegitimate children who do not have a legal father. No statute of limitations, however, except the child's reaching majority age

via intestate succession. *Id.* at 278.

91. *See, e.g., J.E.G. v. C.J.E.*, 172 Ind. App. 515, 360 N.E.2d 1030 (1977) (court acknowledged that legitimate interest in paternity suit is to protect the public interest by preventing the illegitimate child from becoming a ward of the state).

92. *See supra* notes 58-62 and accompanying text.

93. *See supra* note 57 and accompanying text.

94. *See supra* text accompanying note 73.

95. *State v. Wood*, 89 Wash. 2d 97, 102, 569 P.2d 1148, 1151 (1977).

96. *See, e.g., State ex rel. S.M.B. v. D.A.P.*, 284 S.E.2d 912 (W. Va. 1981). The West Virginia Supreme Court colorfully wrote in dicta:

In this regard we would note that the incidence of illegitimacy rises in this country every year; furthermore, a few women, perhaps following the example of T.S. Garp's mother, J. Irving, *The World According to Garp* (1978), deliberately choose to have children out of wedlock because they consciously decide that they want a child but not a husband. *See, e.g., Rivlin*, 'Choosing to Have a Baby on Your Own,' *Ms*, April 1979, p. 68. While we hardly find this either an intelligent or an appropriate approach to the sound upbringing of children, nonetheless, we must recognize the existence of new patterns of life. The difficulty, of course, with eccentric lifestyles is that when they fail to yield the results which were intended the ultimate burden of compensating for individuals' lack of foresight ultimately falls upon the inadequate resources of the West Virginia Department of Welfare.

Id. at 915-16 (footnote omitted).

97. *In re Karen K. v. Christopher D.*, 36 A.D.2d 633, 446 N.Y.S.2d 345 (N.Y. Fam. Ct. 1982).

should limit its use.⁹⁸ Although the Supreme Court had the opportunity in *Mills*⁹⁹ to specify an appropriate statute of limitations for paternity actions that mothers bring, it failed to do so.¹⁰⁰ Some state courts have commended the Court's reasoning in *Mills* that the statute in question was invalid since no similar statute of limitations governed a father's attempt to disprove paternity of children born to his marriage.¹⁰¹ State courts also have emphasized that the HLA blood test is highly reliable, objective, and always available.¹⁰² The Minnesota Supreme Court made the highest recommendation for the HLA blood test when it urged its state legislature to require the HLA test in all paternity cases.¹⁰³ The court attached as much importance to the test as to the right to counsel in paternity disputes.¹⁰⁴

B. Paternity Disputes at Time of Divorce

Many husbands who participate in divorces try to establish grounds of adultery¹⁰⁵ or avoid payment of child support¹⁰⁶ by alleging that they are not the father of a child born to the marriage. Generally, however, husbands must overcome a presumption that children of the marriage are legitimate.¹⁰⁷ Although some cases have traced the presumption of legitimacy to Roman law,¹⁰⁸ the English common law embraced it with the adoption of Lord Mansfield's Rule in the eighteenth century.¹⁰⁹ The Rule effectively made

98. See *supra* notes 72-74 and accompanying text.

99. See authority cited *supra* note 71.

100. See *supra* note 73 and accompanying text. The *Mills* court found a four year statute of limitations inappropriate. A later case cited *Mills* as precedent and found a six year statute of limitations inappropriate. *Hummel v. Smith*, 301 Pa. Super. 276, 447 A.2d 965 (1982) (Beck, J., concurring).

101. *In re Paternity of M.D.H.*, 437 N.E.2d 119, 125-26 (Ind. Ct. App. 1982); see *Lenoir ex rel. Cogdell v. Johnson*, 46 N.C. App. 182, 264 S.E.2d 816 (1980); *State ex rel. S.M.B. v. D.A.P.*, 284 S.E.2d 912 (W. Va. 1981).

102. *Hummel v. Smith*, 301 Pa. Super. 276, 285-86, 447 A.2d 965, 970 (1982) (Beck, J. concurring); see *Malvasi v. Malvasi*, 167 N.J. Super. 513, 401 A.2d 279 (Super. Ct. Ch. Div. 1979).

103. *Hepfel v. Bashaw*, 279 N.W.2d 342 (Minn. 1979).

104. *Id.* at 344.

105. See, e.g., *Rachel v. Rachel*, 412 A.2d 1202 (D.C. 1980) (husband seeking legal separation on ground of adultery claimed younger child was not his biological offspring). See also Note, *Divorce—Authority of Court to Order Mother to Submit Her Child to Blood-Grouping Tests to Determine Issue of Adultery*, 15 J. FAM. LAW 592 (1977).

106. See *L.A.J. v. C.T.J.*, 577 S.W.2d 151 (Mo. Ct. App. 1979) (father challenged child support decree on basis of his nonpaternity).

107. See *infra* notes 113-19 and accompanying text.

108. See, e.g., *Keunedy v. State*, 117 Ark. 113, 116, 173 S.W. 842, 843 (1915).

109. *Goodwright v. Moss*, 2 Cowp. 591, 98 Eng. Rep. 1257 (1777).

the presumption of legitimacy irrebuttable by preventing either spouse from testifying to nonaccess to each other at the time of conception.¹¹⁰ Several commentators have criticized the Rule,¹¹¹ however, and many American courts have abandoned it.¹¹²

Although a few states such as California¹¹³ and Nebraska¹¹⁴ still have a conclusive presumption of legitimacy, most states have adopted a rebuttable presumption through statutes¹¹⁵ or common law.¹¹⁶ Eight states have statutorily adopted¹¹⁷ a provision of the Uniform Parentage Act¹¹⁸ that establishes a rebuttable presumption of legitimacy.¹¹⁹ Thus, the majority of states allow evidence rebutting the presumption of legitimacy, including proof of the husband's nonaccess,¹²⁰ the wife's adultery,¹²¹ and the husband's

110. *Id.*

111. *E.g.* C. McCORMICK, HANDBOOK OF THE LAW OF EVIDENCE § 67 (2d ed. 1972); J. WIGMORE, EVIDENCE §§ 2063-64 (3d ed. 1940).

112. *See, e.g.*, *Serafin v. Serafin*, 401 Mich. 629, 258 N.W.2d 461 (1977); *Ventresco v. Bushey*, 159 Me. 241, 191 A.2d 104 (1963); *Moore v. Smith*, 178 Miss. 383, 172 So. 317 (1937).

113. CAL. EVID. CODE § 621 (West Cum. Supp. 1983). *See also* Comment, *The Irrebuttable Presumption of California Evidence Code Section 621*, 12 U.C.D. L. REV. 452 (1979).

114. *Hudson v. Hudson*, 151 Neb. 210, 36 N.W.2d 851 (1949).

115. *See, e.g.*, IOWA CODE §§ 675.7-.15 (1975); KAN. STAT. ANN. § 38-1101 (Supp. 1981); N.M. STAT. ANN. §§ 40-51 to -7 (1983); VT. STAT. ANN. tit. 15, § 331 (1974); WEST VA. CODE §§ 48-7-1 to 4 (1966 & Cum. Supp. 1975).

116. Missouri, Pennsylvania, South Carolina, and Virginia have adopted a common law rebuttable presumption of legitimacy. *See Note, Children Born of the Marriage—Res Judicata Effect on Later Support Proceedings*, 45 MO. L. REV. 307 n.5 (1980).

117. CAL. CIV. CODE § 7004 (West 1983); COLO. REV. STAT. § 19-6-105 (1973); HAWAII REV. STAT. § 584-4 (1976); MINN. STAT. ANN. § 257.55 (West 1982); MONT. CODE ANN. § 40-6-105 (1981); N.D. CENT. CODE § 14-17-04 (1981); WASH. REV. CODE ANN. § 26.040 (Supp. 1983); WYO. STAT. § 14-2-102 (1978). Montana's statute differs from the Uniform Parentage Act in that it allows rebuttal of the presumption of legitimacy by a preponderance of the evidence rather than by clear and convincing evidence. *See infra* note 137.

118. *See supra* note 44.

119. The Uniform Parentage Act provides:

(a) A man is presumed to be the natural father of a child if:

(1) he and the child's natural mother are or have been married to each other and the child is born during the marriage, or within 300 days after the marriage is terminated by death, annulment, declaration of invalidity, or divorce, or after a decree of separation is entered by a court;

....

(b) A presumption under this section may be rebutted in an appropriate action only by clear and convincing evidence. If two or more presumptions arise which conflict with each other, the presumption which on the facts is founded on the weightier considerations of policy and logic controls. The presumption is rebutted by a court decree establishing paternity of the child by another man.

UNIF. PARENTAGE ACT § 4(a)(1) & (b), 9A U.L.A. 590-91 (1979).

120. *See, e.g.*, *Mock v. Mock*, 411 So. 2d 1063 (La. 1982). *See supra* notes 128-30 and accompanying text.

impotence or sterility.¹²² State courts recently have admitted evidence from blood tests like the HLA.¹²³ Obviously, the older forms of "proof" are neither as reliable nor as objective as the HLA blood test¹²⁴ since they rely partly on past recollections and self-serving commentary.¹²⁵

The courts in those states with a rebuttable presumption of legitimacy generally will allow the results of an HLA blood test into evidence if the proponent husband shows good cause for their admission.¹²⁶ Because the husband is attempting to prove he is not the father of the child, he would use the test results as exclusionary evidence. If the test, however, indicated a high likelihood of paternity the mother or child would want to use the results as affirmative evidence. The controversy surrounding the use of HLA blood tests in divorce situations, however, does not concern whether the results are exclusionary or affirmative as it does in paternity determinations in illegitimacy cases,¹²⁷ but instead focuses more on the applicable statute of limitations. Recent judicial decisions and modern policy considerations support a statute of limitations in paternity determinations of illegitimate children, the only limit of which should be a child's reaching the age of majority.¹²⁸ However, similar policy considerations—the best interests of the

121. See *supra* note 105 and accompanying text.

122. See CAL. EVID. CODE § 621 (West Supp. 1983) (statute provides that a conclusive presumption of paternity exists if conception occurs while husband and wife cohabit, but that no such presumption exists if one can show that the alleged father was impotent or sterile). *Vincent B. v. Joan R.*, 126 Cal. App. 3d 619, 179 Cal. Rptr. 9 (1981) (The burden of showing impotence or sterility is upon the person claiming paternity. The court will allow blood test results to rebut conclusive presumption of a cohabitant's paternity only after the claiming party carries the burden of proof and shows sterility or impotence.). See also Recent Development, *California's Tangled Web: Blood Tests and the Conclusive Presumption of Legitimacy*, 20 STAN. L. REV. 754, 757 (1968).

123. See, e.g., *County of Fresno v. Superior Court*, 92 Cal. App. 3d 133, 137-38, 154 Cal. Rptr. 660, 662-63 (1979) (court allows use of the HLA blood test to prove nonpaternity if parties lay adequate foundation and demonstrate good cause); *J.H. v. M.H.*, 177 N.J. Super. 436, 441, 426 A.2d 1073, 1076 (1980) (holding that HLA testing was not precluded in divorce case); *Wake County v. Green*, 53 N.C. App. 26, 279 S.E.2d 901 (1981) (holding that the HLA test may be used in a divorce case to show that a man other than the husband was the father of a child). See also *infra* note 124 and accompanying text.

124. See *supra* note 25 and accompanying text.

125. See cases cited *supra* notes 65-66 for purposes of analogy. The cases illustrate instances in which courts have allowed older means of proof in paternity disputes. These means of proof rely heavily upon self-serving statements and past recollections.

126. See, e.g., *Balfour v. Balfour*, 413 So. 2d 1167, 1169 (Ala. Civ. App. 1982); *County of Fresno v. Superior Court*, 92 Cal. App. 3d 133, 137-38, 154 Cal. Rptr. 660, 662-63 (1979) (existence of better blood test than ABO test found sufficient for good cause).

127. See *supra* notes 53-56 and accompanying text.

128. See *supra* notes 97-104 and accompanying text.

child¹²⁹ and the preservation of family integrity¹³⁰—bolster the argument for a very brief statute of limitations in paternity disputes at the time of divorce.

1. The Child's Best Interests

Two of a child's possible interests at the time of a divorce are maintenance of his status as legitimate¹³¹ and insulation from routine disavowals of his legitimacy in divorce proceedings.¹³² Because recent Supreme Court cases have sought to equalize the status of illegitimate and legitimate children,¹³³ arguments based on the legal consequences of being labelled "illegitimate" command less support.¹³⁴ The child's interest in preserving his legitimacy, therefore, is more defensible if parties present it to reflect a concern with the maintenance of a legal relationship with the father as well as the avoidance of any social stigmas of illegitimacy. Maintenance of a child's legal relationship with his father achieves an important and practical economic goal: as long as the child is the legal dependent of the husband seeking a divorce, the husband must support the child financially during minority.¹³⁵ A child's interest in protection from courtroom denials of his legitimacy in divorce proceed-

129. See *supra* note 10 and accompanying text.

130. See, e.g., *Ferguson v. Ferguson*, 126 Cal. App. 3d 744, 179 Cal. Rptr. 108, 110 & n.2 (1981).

131. See *infra* notes 133-35 and accompanying text.

132. See Recent Development, *supra* note 122, at 758, 759 n.23.

133. See *supra* note 85 and accompanying text. See also Comment, *The Expanding Rights of the Illegitimate*, 3 CREIGHTON L. REV. 135 (1970).

134. See, e.g., *Mathews v. Lucas*, 427 U.S. 495, 512-13 (1976) (stigma of illegitimacy is not determinative of illegitimate's rights to social security benefits); *Gomez v. Perez*, 409 U.S. 535, 538 (1973) (discrimination results if state denies substantial benefits for an illegitimate because of that child's status compared to legitimate children); Comment, *supra* note 133. *Contra Serafin v. Serafin*, 401 Mich. 629, 258 N.W.2d 461 (1977) (Coleman, J., concurring). In his concurring opinion of the case, Justice Coleman wrote in dicta:

Despite these enlightened advances, there still are, unfortunately, social distinctions made between the legitimate and illegitimate child which continue to stigmatize the illegitimate child and scar his or her psychological development. We need no learned treatise to know that many children branded as illegitimate suffer painful and sometimes crippling emotional damage at the hands of cruel or thoughtless peers and adults. The work "bastard" has not yet lost its sting to the children against whom it is too often applied. Moreover, feelings of parental rejection and abandonment are realities that often continue to plague the illegitimate child. Related neglect and even abuse are not uncommon. It is no accident that many of these children strike back by committing antisocial or criminal acts.

Id. at 637-38, 258 N.W.2d at 464.

135. See, e.g., *In re Paternity of M.D.H.*, 437 N.E.2d 119, 128 (Ind. Ct. App. 1982); *County of Lenoir ex rel. Cogdell v. Johnson*, 46 N.C. App. 182, 186, 264 S.E.2d 816, 819 (1980).

ings, in comparison, reflects a concern for his emotional well-being. If courts admit HLA blood test results in all divorce proceedings for the purpose of denying paternity, then husbands could engage routinely in "fishing expeditions" by requesting that the blood tests be taken and the results admitted in efforts to avoid child support.¹³⁶ The traumatic effect of divorce on children is obviously evident, and courts should recognize the repugnance of additionally subjecting a minor to a blood test and thereby forcing the child to acknowledge that his father is attempting to deny paternity. The failure of some states to have a statute of limitations that bars a husband's action to disprove his paternity¹³⁷ makes such scenarios inevitable. Because the economic and emotional interests of children are so significant, all states should adopt some statute of limitations to limit such action.

2. Preservation of Family Integrity

The public policy of maintaining family integrity provides further support for a short statute of limitations in the divorce context. Some courts have characterized family integrity as "actual" and "legal."¹³⁸ Involvement in a divorce can destroy "actual" family integrity,¹³⁹ but courts feel that "legal" family integrity can survive.¹⁴⁰ A reasonable statute of limitations that excludes some blood test evidence in divorce proceedings would help maintain the legal father-child relationship and, although the child would not benefit from his father's presence, the existence of a legal relationship hopefully would preserve at least some of the psychological and emotional bonds between the two.¹⁴¹ An important consideration that courts should include in their analysis of this policy-dominated area is that, although the impact and accuracy of other forms of evidence¹⁴² may diminish, the utility and accuracy of the HLA blood test never waivers.¹⁴³ Thus, the HLA blood test could affect dramatically the number of paternity disavowals and destroy some legal relationships that courts should maintain to protect the

136. See Recent Development, *supra* note 122, at 759 n.23; *R.D.S. v. S.L.S.*, 402 N.E.2d 30, 37 n.13 (Ind. Ct. App. 1980).

137. See, e.g., *In re Paternity of M.D.H.*, 437 N.E.2d 119 (Ind. Ct. App. 1982).

138. See, e.g., *Stephen B. v. Sharyne B.*, 124 Cal. App. 3d 524, 530, 177 Cal. Rptr. 429, 432 (1981).

139. *Id.*, at 530, 177 Cal. Rptr. at 432.

140. *Id.*

141. *Id.* at 531, 177 Cal. Rptr. at 433.

142. See *supra* notes 120-22 and accompanying text.

143. See *supra* note 25.

economic and emotional interests of children and the legal integrity of the family. The best method for states to minimize the effect of HLA test results on children of divorce is to adopt short statutes of limitation that begin to run at the birth of the child. Present state statutes of limitation range from two years¹⁴⁴ to non-existent.¹⁴⁵ Interestingly, the Uniform Act on Parentage contains a statute of limitations of five years starting to run at the birth of the child¹⁴⁶ that apparently contradicts the express admonition in the Act's Comments that "paternity actions should be brought promptly"¹⁴⁷ to determine the paternity of illegitimate children. Although a child must be at least six months old before he can undergo an HLA blood test,¹⁴⁸ a two to three year statute of limitations beginning at the date of the child's birth would be reasonably adequate. A time limit of two or three years would afford a husband enough time to disprove the paternity of a child whom he discovered was not his biological offspring, yet would be short enough to protect the economic and emotional interests of a child who had achieved the longevity within the family unit that courts equitably should recognize as worthy of the protection that they give legal family integrity. In those divorce actions in which a man seeks to disprove paternity within a two or three year statute of limitations courts should admit HLA blood test results because they most objectively can exclude the man as the biological father.¹⁴⁹ In those instances, however, when a husband seeks to disprove paternity of a child who is over two or three years old, states should have statutes of limitation that toll not only the admission

144. See, e.g., *Stephen B. v. Sharyne B.*, 124 Cal. App. 3d at 527, 177 Cal. Rptr. at 431 (1981).

145. See *supra* note 137.

146. Section (6)(a) of the Uniform Parentage Act provides:

(a) A child, his natural mother, or a man presumed to be his father under Paragraph (1), (2), or (3) of Section 4(a), may bring an action

. . . .

(2) for the purpose of declaring the nonexistence of the father and child relationship presumed under Paragraph (1), (2), or (3) of Section 4(a) only if the action is brought within a reasonable time after obtaining knowledge of relevant facts, but in no event later than five years after the child's birth. After the presumption has been rebutted, paternity of the child by another man may be determined in the same action, if he has been made a party.

UNIF. PARENTAGE ACT § 6(a), 9A U.L.A. 593 (1979).

147. UNIF. PARENTAGE ACT § 7 Commissioners' Comment (1979).

148. See, e.g., *Elzey v. Smith*, 412 So. 2d 918, 919 (Fla. Dist. Ct. App. 1982) (baby had to be six months old to undergo HLA blood test because the baby needed to weigh enough to give blood).

149. See *supra* note 25 and accompanying text.

of HLA blood test results, but also the cause of action itself.

C. Unwed Fathers Attempting to Establish Paternity Rights

Unwed fathers recently have attempted to establish various rights respecting their children in the same actions in which mothers seek paternity determinations of illegitimate children. The rights that fathers usually seek are custody of the child,¹⁵⁰ visitation with the child,¹⁵¹ or prevention of the child's adoption.¹⁵² Many unwed fathers have achieved these goals since the Supreme Court held in *Stanley v. Illinois*¹⁵³ that an unmarried father was entitled to notice and a hearing in child custody proceedings for his illegitimate children.¹⁵⁴ Because the unwed father claiming paternity and seeking adoption of his illegitimate child bears the bur-

150. See, e.g., *W.E.J. v. Superior Court*, 100 Cal. App. 3d 303, 160 Cal. Rptr. 862 (1979); *In re Kelvin M.*, 77 Cal. App. 3d 396, 143 Cal. Rptr. 561 (1978); *In re Tricia M.*, 74 Cal. App. 3d 125, 141 Cal. Rptr. 554 (1977), cert. denied, 435 U.S. 996 (1978); *Pi v. Delta*, 175 Conn. 527, 400 A.2d 709 (1978); *Stevens v. Leone*, 35 Conn. Supp. 237, 406 A.2d 402 (1979); *In re Ozment*, 61 Ill. App. 3d 1044, 378 N.E.2d 409 (1978); *Vanderlaan v. Vanderlaan*, 9 Ill. App. 3d 260, 292 N.E.2d 145 (1972); *Hrouda v. Winne*, 77 A.D.2d 62, 432 N.Y.S.2d 643 (1980). See generally Note, *Putative Fathers: Unwed, But No Longer Unprotected*, 8 HOFSTRA L. REV. 425 (1980) (general discussion of the increased rights presently afforded putative fathers).

151. See, e.g., *Roque v. Frederick*, 272 Ark. 392, 614 S.W.2d 667 (1981); *Donald J. v. Evna M.*, 81 Cal. App. 3d 929, 147 Cal. Rptr. 15 (1978); *Griffith v. Gibson*, 73 Cal. App. 3d 465, 142 Cal. Rptr. 176 (1977); *Sullivan v. Bonafonte*, 172 Conn. 612, 376 A.2d 69 (1977); *Forestiere v. Doyle*, 30 Conn. Supp. 284, 310 A.2d 607 (1973); *People ex rel. Ritchie v. Ritchie*, 58 Ill. App. 3d 1045, 374 N.E.2d 1292 (1978); *People ex rel. Vallera v. Rivera*, 39 Ill. App. 3d 775, 351 N.E.2d 391 (1976); *Dustin v. Belanger*, 429 A.2d 212 (Me. 1981); *Pennsylvania v. Rozanski*, 206 Pa. Super. 397, 213 A.2d 155 (1965). See generally Note, *The Rights of Fathers of Non-Marital Children to Custody, Visitation and to Consent to Adoption*, 12 U.C.D. L. REV. 412 (1979) (discussion of visitation rights under the Uniform Parentage Act).

152. See, e.g., *Caban v. Mohammed*, 441 U.S. 380 (1979); *Quilloin v. Walcott*, 434 U.S. 246 (1978); *Rothstein v. Lutheran Social Servs.*, 405 U.S. 1051 (1972); *Adoption of Rebecca B.*, 68 Cal. App. 3d 193, 137 Cal. Rptr. 100 (1977); *Cheryl H. v. Superior Court*, 41 Cal. App. 3d 273, 115 Cal. Rptr. 849 (1974); *People ex rel. Slawek v. Covenant Children's Home*, 52 Ill. 2d 20, 284 N.E.2d 291 (1972). See generally Note, *Unwed Fathers: Conflict of Rights in Adoption Proceedings*, 7 FLA. ST. U.L. REV. 559 (1979); Recent Development, *Constitutional Law—Fourteenth Amendment—Equal Protection Clause—Adoption—Rights of Putative Fathers*, 18 DUQ. L. REV. 375 (1980) (discussion of statute requiring the mother's, but not the putative father's consent before adoption as violative of equal protection).

153. 405 U.S. 645 (1972).

154. *Id.* *Stanley* found that a father's interest in his illegitimate child was "cognizable and substantial." Prior to this 1972 decision, unwed fathers were predominately unsuccessful in attempting to claim custody of their biological children. See, e.g., *Jolly v. Queen*, 142 S.E.2d 592 (N.C. 1965). See generally Marcus, *Equal Protection: The Custody of the Illegitimate Child*, 11 J. FAM. L. 1 (1971); Recent Development, *Family Law—Voluntary Legitimation—Father Has No Absolute Right to Legitimate Child Solely on Proof of Biological Fatherhood*, 8 ST. MARY'S L.J. 392 (1976); *Stanley v. Illinois*, 405 U.S. 645, 652 (1972).

den of proof,¹⁵⁵ he needs a precise method of proving paternity like the HLA blood test.¹⁵⁶

Unwed fathers also have attempted to utilize the results of the HLA test to prove their paternity when litigating other rights.¹⁵⁷ On the basis of the *Stanley* decision and state statutes that adopt provisions of the Uniform Parentage Act,¹⁵⁸ for example, unwed fathers successfully have sought custody of their children by using the clear and convincing proof of the HLA blood test.¹⁵⁹ Thus, the HLA blood test has extended the rights of unwed fathers and, because admittance of the test has not proved controversial when unwed fathers seek to establish paternity of a clearly illegitimate child, courts generally have denied admittance only intermittently on the basis of outdated exclusionary statutes.¹⁶⁰

The attempted extension of the use of the HLA test by an alleged biological father to challenge the paternity of a presumably legitimate child, however, has caused courts to consider important policy considerations that might limit this controversial use. In *Happel v. Mecklenburger*,¹⁶¹ for instance, the alleged biological father of an eight year old child brought suit to establish paternity and visitation rights. The lower court denied his request to require the parties to submit to the HLA blood test.¹⁶² The court of appeals upheld the trial court's decision because the HLA blood test was only one factor in the determination of paternity, and the denial of the motion to require the test, therefore, was not reversible error.¹⁶³ Extensive dicta which detailed strong policy considerations indicated that the appellate court preferred to deny use of tests like the HLA which could identify the biological father of a

155. See H. KRAUSE, *supra* note 6, at 140-47.

156. *Stanley v. Illinois*, 405 U.S. at 657 n.9.

157. See, e.g., *Marticoarena v. Miller*, 597 P.2d 1349 (Utah 1979).

158. See UNIF. PARENTAGE ACT, §§ 2 & 6(a)(1), 9A U.L.A. 588, 593 (1979). An unwed father can bring an action under §§ 2 and 6(a)(1), which provide that the "parent and child relationship extends equally to every child and to every parent, regardless of the marital status of the parents" and that "a child, his natural mother, or a man presumed to be his father under Paragraph (1), (2), or (3) of Section 4(a), may bring an action . . . at any time for the purpose of declaring the existence of the father and child relationship presumed under Paragraph (1), (2), or (3) of Section 4(a)."

159. See, e.g., *LaCroix v. Deyo*, 113 Misc. 2d 89, 447 N.Y.S.2d 864 (N.Y. Fam. Ct. 1981), *aff'd*, 88 A.D.2d 1077, 452 N.Y.S.2d 726 (1982).

160. See, e.g., *J.B. v. A.F.*, 92 Wis. 2d 696, 698-705, 285 N.W.2d 880, 881-84 (Wis. Ct. App. 1979).

161. 101 Ill. App. 3d 107, 427 N.E.2d 974 (1981).

162. *Id.* at 109, 427 N.E.2d at 977.

163. *Id.* at 112, 427 N.E.2d at 981.

child whose paternity was in doubt even by his mother.¹⁶⁴ The strongest public policy argument against a paternity claim by the alleged biological father that the court recognized was the preservation of the family unit.¹⁶⁵ Although the parents of the child in *Happel* had divorced, the legal father and child regularly visited each other and the legal father supplied child support.¹⁶⁶ Thus, in denying the unwed father's cause of action the court gave more weight to the established emotional family ties between the legal father and the child than to the possibly destructive recognition of the unwed father's paternity rights.¹⁶⁷

Not all courts recognize the *Happel* rationale. In *R. McG. v. J.W.*¹⁶⁸ the dissent voiced the same policy considerations that the *Happel* court found determinative,¹⁶⁹ but the majority allowed an unwed father to use the HLA blood test successfully to claim paternity of a child born to a married couple.¹⁷⁰ The dissent strongly criticized the majority's holding and noted the state's strong interest in promoting durable family ties.¹⁷¹ The majority, however, based its decision on the equal protection grounds that an unwed father has the right to seek his paternity rights concerning a child born to a marriage, just as an unwed mother can institute a paternity suit against a married man.¹⁷² Although the court did not discuss specifically the natural father's good faith, it did note that the guardian *ad litem* consistently advocated allowance of the action,¹⁷³ and that the plaintiff instituted his claim one and one-half years after the birth of the child.¹⁷⁴ If factors such as these indicate an earnest and persistent interest by the biological father to establish a parent-child relationship, then the effects of such a relationship would be positive because the child would benefit from the

164. *Id.* at 115, 427 N.E.2d at 983. The mother had undergone artificial insemination with sperm of both her husband and an anonymous donor near the time of conception. She also had sexual intercourse with her husband and the plaintiff near the same dates. 427 N.E.2d 979.

165. *Id.* at 115, 427 N.E.2d at 983.

166. *Id.*

167. The unwed father sought both a determination of paternity and visitation rights. 427 N.E.2d at 977.

168. 615 P.2d 666 (Colo. 1980).

169. *Id.* at 676 (Lohr, J., dissenting).

170. *Id.* at 669.

171. *Id.* at 676-77.

172. *Id.* at 671.

173. *Id.* at 672.

174. *Id.* at 667-68.

love and support of two "fathers" instead of one.¹⁷⁵

Like the cases in which a husband wants to disprove his paternity in a divorce, unwed father cases in which the plaintiff attempts to prove he is a child's biological parent require the plaintiff to overcome a rebuttable presumption of legitimacy.¹⁷⁶ The plaintiffs in both *Happel*¹⁷⁷ and *R. McG.*¹⁷⁸ tried to use the HLA blood test because it provides the most reliable evidence in establishing paternity. The differing conclusions on admissibility in these cases, however, indicate the importance that courts place on the values of family preservation. This Note proposes that an appropriate statute of limitations would govern most effectively the use of the HLA blood test in unwed father cases. No reason exists to bar an unwed father, for instance, from claiming visitation rights with his biological child if he institutes the action within a reasonable time.¹⁷⁹ A suggested reasonable time of two years from the date of the child's birth would give the unwed father enough time to seek judicial recourse that used the precise, objective evidence which the HLA test provides, and would also protect the established family unit as deeper emotional ties between the child and his legal father developed.

IV. CONCLUSION

The use of the HLA blood test as evidence in cases of disputed parentage involves very emotional issues. Different types of disputed paternity cases present similar policy considerations with one common denominator: the best interests of the child. Determination of those interests, however, is difficult and scientifically inexact. Medical and bio-medical engineering developments in the determination of paternity often outrace legal reasoning. Because of this unavoidable dichotomy, American courts have shown mixed reactions to the use of the HLA blood test as evidence in disputed paternity cases. While many courts have recognized its accuracy and extolled its use as objective evidence, other courts have expressed a more tentative attitude. In general, however, courts have

175. See Case Comment, *Bastardizing the Legitimate Child: The Colorado Supreme Court Invalidates the Uniform Parentage Act Presumption of Legitimacy in R. McG. v. J.W.*, 59 DEN. L.J. 157, 171-72 (1981).

176. See *supra* notes 115-19 and accompanying text.

177. See *supra* note 161 and accompanying text.

178. See *supra* note 169 and accompanying text.

179. See *supra* notes 129-48 and accompanying text for an analogous argument in the area of paternity disputes in divorce actions.

admitted the HLA blood test as both exclusionary and affirmative evidence in various paternity claims.

Yet, the common denominator—the best interests of the child—has resulted in courts questioning the appropriateness of the evidence in certain cases. Concerns for family values and the child's emotional and financial health constitute the major public policy considerations in this debate. Because the accurate, reliable evidence that the HLA blood tests provides may clash with a child's best welfare, admittance of the HLA blood test as evidence in disputed paternity cases may not always be appropriate.

First, this Note proposes that the HLA blood test should be unconditionally admissible in all paternity claims involving illegitimate children who do not have a legal father. Requiring admission of the HLA blood test in AFDC and non-AFDC paternity suits would serve both public and private interests because it is the most reliable and accurate evidence available. Positive identification of the biological father would benefit the illegitimate child emotionally and financially; and if the father continued to support the child willingly, the public sector would realize financial benefits. Litigation also would decrease if HLA blood test results convinced a father to accept his parental duties without a judicial mandate.

Second, this Note suggests that courts admit the HLA blood test as evidence in paternity disputes which arise out of divorce proceedings only if those actions fall within a two or three year statute of limitations. Such actions involve different policy considerations than do paternity actions involving illegitimate children who do not have a legal father. A supposedly legitimate child, for example, is subject to illegitimacy and loss of financial support if the father who has supported him can prove he is not the biological father. The law, therefore, should afford some protection to the child from evidence—albeit accurate and reliable—which could seriously harm his best interests. The law, however, should not bar such claims by the father completely, and arguably should protect him from unjustified responsibility. Policy considerations, therefore, favor requiring the father to forego his paternity action within two to three years from the child's birth.

Finally, this Note advocates a two year statute of limitations for actions by unmarried fathers who attempt to introduce HLA blood test results to assert paternity rights concerning a presumably legitimate child who has a legal father. Two very strong policy considerations—the rights of unwed fathers and the best interests

of the child—clash in this situation. Although an unwed father undoubtedly has a right to seek paternity of his illegitimate child, such an action poses no difficulty to the use of the HLA blood test as evidence. Indeed, courts should encourage the use of the HLA test in this instance because of the accuracy and objectivity of the evidence. In situations in which an unmarried father attempts to introduce HLA blood test results in order to assert paternity rights concerning a presumably legitimate child who has a legal father, however, the public policies clash. The growing support for equal protection of the rights of unwed fathers militates in favor of admission, but a child's emotional and economic interests and society's interest in the preservation of the family constitute powerful elements of the argument for inadmissibility. The best solution, again, seems to be reliance on the statute of limitations as a line of demarcation for the use of the HLA blood test as evidence. Thus, if an unwed father wants to seek paternity rights of a child already born to a marriage, courts should require him to bring his claim within a two year period commencing at the birth of the child. Thereafter, the courts should not permit the unmarried father to disrupt the established family unit.

Thus, although the technology is available to provide accurate and reliable information about paternity, courts should not allow the unchecked use of this evidence. Legislatures and courts should weigh carefully the availability and capability of the HLA blood test against the desirability and necessity of the information it provides. Although the test can clarify biological relationships, courts should not neglect the psychological, emotional, and economic aspects of the human condition in the blind rush to use all the evidence that science can provide.

PATRICIA BUNDSCHUH BLUMBERG

Twisting the Purposes of Discovery: Expert Witnesses and the Deposition Dilemma

I.	INTRODUCTION	1616
II.	LEGAL HISTORY CONCERNING DISCOVERY OF EXPERT WITNESSES	1618
	A. <i>Common Law Disagreement Over the Appropriateness of Discovering Expert Witnesses</i> ..	1618
	B. <i>Enactment of Federal Rule 26(b)(4): Discovery of Expert Witnesses Permitted in Federal Courts</i>	1621
III.	DEPOSING EXPERT WITNESSES: HIDDEN DANGER FOR DEPOSING PARTY IN FEDERAL RULE 32(A)(3)	1621
	A. <i>The Process of Discovering Expert Witnesses and the Deposition Problem</i>	1622
	B. <i>The Unavailability Requirement of Rule 32(a)(3)</i>	1623
	1. Rule 32(a)(3)(A): Death of the Deposed Witness	1623
	2. Rule 32(a)(3)(B): Absence of the Deposed Witness	1624
	3. Rule 32(a)(3)(C): Age, Illness, Infirmary, and Imprisonment of the Deposed Witness	1626
	4. Rule 32(a)(3)(D): Inability to Procure Attendance of the Deposed Witness by Subpoena	1626
	5. Rule 32(a)(3)(E): Admissibility under Exceptional Circumstances	1628
IV.	JUDICIAL PERCEPTIONS OF THE WITNESS DEPOSITION PROBLEM	1630
V.	ANALYSIS	1633
	A. <i>The Expert Witness Deposition Problem: The Degree of Unfairness</i>	1633
	1. Absence of Proper Opportunity for Cross-Examination	1633
	(a) <i>Dependence of Admissibility on Prior Cross-Examination</i>	1634
	(b) <i>Adequacy of Opportunity for Cross-Examination</i>	1635

2.	Practical Implications of the Expert Witness Deposition Problem	1637
	(a) <i>Potential for Abuse</i>	1637
	(b) <i>Inhibition of Discovery</i>	1637
B.	<i>Potential Solutions to the Expert Witness Deposition Problem</i>	1638
1.	Court Orders Limiting Use of Depositions	1638
2.	Stipulation by the Parties	1640
3.	Cautionary Instructions	1640
4.	A Stricter Test for Unavailability of Expert Witnesses	1641
5.	Local Court Rules	1642
6.	Amendment to the Federal Rules of Civil Procedure	1643
VI.	CONCLUSION	1644

I. INTRODUCTION

The Federal Rules of Civil Procedure (hereinafter the "Federal Rules") permit parties to civil actions in federal courts to discover the opinions and conclusions of any expert witness that a party-opponent expects to call as a witness at trial.¹ Federal rule 26 governs discovery of these experts and provides several different but integrated techniques for compelling disclosure of information,² including written interrogatories³ and court ordered depositions.⁴

The provisions of rule 26(b)(4) resulted from a protracted debate by courts and commentators over the merits of permitting discovery of expert witnesses.⁵ Federal rule 26(b)(4) strongly affirms the idea that adequate discovery of expert witnesses is essential to their effective cross-examination in today's complex and highly technical civil cases.⁶

1. FED. R. CIV. P. 26(b)(4).

2. C. WRIGHT, LAW OF FEDERAL COURTS § 81, at 543 (4th ed. 1983).

3. See FED. R. CIV. P. 26(a) & 33.

4. See *id.* at 26(a) & 27-32. A 1941 study of pretrial discovery revealed that parties used oral depositions more than any other discovery device. See A. Stockman, Some Statistical Observations on the Operation of Discovery and Related Provisions of the Federal Rules of Civil Procedure 21 (1942) (unpublished study, Administrative Office of the U.S. Courts), cited in Note, *Dead Men Tell No Tales: Admissibility of Civil Depositions upon Failure of Cross-examination*, 65 VA. L. REV. 153, 154 n.11 (1979).

5. See *infra* notes 14-27 and accompanying text.

6. "Effective cross-examination of an expert witness requires advance preparation.

Parties to civil actions enjoy an absolute right to cross-examine witnesses at trial.⁷ A party seeking discovery of a party-opponent's expert witness in preparation for cross-examination and rebuttal of the expert at trial, however, lacks sufficient preparation to cross-examine or rebut the expert's responses to discovery questions. An expert's testimony during a discovery deposition,⁸ therefore, is not subject to true cross-examination but, nevertheless, still is potentially admissible at trial against the discovering party. This anomaly arises because rule 32(a)(3) permits any party to use the deposition of a witness "for any purpose" if the witness becomes unavailable⁹ to testify at trial.¹⁰ Thus, if a party to a civil action deposed an expert witness pursuant to rule 26(b)(4), and the expert subsequently became unavailable to testify at trial within the meaning of rule 32(a)(3), a federal district court judge, under the Federal Rules as they presently exist, probably would permit introduction of the deposition into evidence against the deposing party.¹¹ A party who desires to depose a party-opponent's expert witness in such cases faces the quandary of deposing the opponent's expert witness to prepare effectively for cross-examination and thereby risking that opposing counsel will use the deposition against him at trial.¹²

The system of discovery that the Federal Rules establish theoretically entitles all parties in civil actions, prior to commencement of trial, to disclosure of all relevant nonprivileged information in

The lawyer even with the help of his own experts frequently cannot anticipate the particular approach his adversary's expert will take or the data on which he will base his judgment on the stand." FED. R. CIV. P. 26(b)(4) advisory committee note.

7. See, e.g., *Alford v. United States*, 282 U.S. 687, 691-92 (1931); *United States v. Stoehr*, 196 F.2d 276, 280 (3d Cir.), cert. denied, 344 U.S. 826 (1952); 4 B. JONES, JONES ON EVIDENCE § 25:1, at 106 (6th ed. 1972); 5 J. WIGMORE, EVIDENCE IN TRIALS AT COMMON LAW § 1367, at 32 (J. Chadbourn ed. 1974).

8. Although the Federal Rules do not distinguish between discovery and evidentiary depositions, recognizable practical differences distinguish them. See *infra* notes 125-28 and accompanying text.

9. See *infra* notes 44-86 and accompanying text for a discussion of the unavailability requirement in rule 32(a)(3).

10. FED. R. CIV. P. 32(a).

11. The problem of using discovery deposition testimony against a discovering party exists with depositions of all witnesses, but the arguments against permitting this technique are most compelling in the context of expert witness depositions.

12. A 1976 survey revealed that counsel deposes the opponent's expert in approximately 60% of the cases in which opposing counsel expects the expert to testify at trial. Graham, *Discovery of Experts Under Rule 26(b)(4) of the Federal Rules of Civil Procedure: Part Two, an Empirical Study and a Proposal*, 1977 U. ILL. L.F. 169, 175 (1977) [hereinafter cited as *Graham, Part Two*].

the possession of any person.¹³ Thus, federal discovery rules should not force litigants to choose between failing to depose a party-opponent's expert witness and thereby preparing inadequately for trial, and deposing the expert witness and consequently risking that opposing counsel will use the deposition against him at trial without the benefit of cross-examination. Part II of this Note reviews common law disagreement over the appropriateness of expert witness discovery and the acceptance of the principle under the Federal Rules. Part III discusses the procedural mechanics of discovering an expert witness and demonstrates the potential for use of such a deposition against the discovering party. Part IV reviews the judicial history of the expert witness deposition problem and demonstrates that courts have ignored the policy reasons that favor remedying this procedural problem. Part V discusses the inequities that the expert witness deposition problem causes. It concludes that courts, by admitting discovery depositions of expert witnesses against the deposing party, not only have controverted the basic purposes of pretrial discovery, but also erroneously have eliminated the process of deposing the witness of a party-opponent as a pretrial discovery technique by effectively recharacterizing it as an extension of formal trial proceedings that presume the deposing party engaged in full pretrial preparation. Part V then explores several possible methods courts and rulemakers could employ to solve this problem, including an amendment to rule 26.

II. LEGAL HISTORY CONCERNING DISCOVERY OF EXPERT WITNESSES

A. *Common Law Disagreement Over the Appropriateness of Discovering Expert Witnesses*

Before adoption of the 1970 amendments to the Federal Rules, most courts did not permit parties to discover information from

13. C. WRIGHT, *supra* note 2, at 540. Professor Charles Wright described the scope of discovery that Federal Rule 26 authorizes as follows:

The scope of discovery contemplated by Rule 26 is extremely broad. "No longer can the time-honored cry of 'fishing expedition' serve to preclude a party from inquiring into the facts underlying his opponent's case." Discovery may extend to matters relating to the claim or defense of any party, unlike the former equity practice that limited discovery to matters in support of the proponent's case. It is no objection that the examining party already knows the facts as to which he seeks discovery, since one of the purposes of discovery is to ascertain the position of the adverse party on the controverted issues.

Id. at 543-44 & nn.27-30 (quoting *Hickman v. Taylor*, 329 U.S. 495, 507 (1947)) (footnotes omitted).

experts whom party-opponents hired as witnesses or consultants.¹⁴ These courts characterized expert information as privileged and articulated several reasons for this characterization. First, some courts stated that an expert's status as an expert rendered the information in his possession privileged.¹⁵ Second, some courts reasoned that expert information fell within the scope of the attorney-client privilege.¹⁶ Third, some courts felt that expert information constituted attorney work product and protected it with a qualified privilege.¹⁷ Last, some courts prohibited discovery simply because they felt that allowing a party to obtain free access to information for which a party-opponent had paid by hiring an expert was unfair.¹⁸ Courts that did allow discovery of expert witnesses frequently restricted it to certain circumstances or limited the amount or type of disclosure a discovering party could obtain. Some courts, for example, permitted parties to discover the information that formed the basis of an expert's conclusions, but not the conclusions themselves.¹⁹ Other courts allowed parties to discover only expert employees of a party-opponent.²⁰ Although

14. See Friedenthal, *Discovery and Use of an Adverse Party's Expert Information*, 14 STAN. L. REV. 455 (1962); Long, *Discovery and Experts Under the Federal Rules of Civil Procedure*, 38 F.R.D. 111, 116-119 (1966).

15. See *American Oil Co. v. Pennsylvania Petroleum Prod. Co.*, 23 F.R.D. 680, 685 (D.R.I. 1959); *Walsh v. Reynolds Metals Co.*, 15 F.R.D. 376, 378-79 (D.N.J. 1954); *Lewis v. United Air Lines Corp.*, 32 F. Supp. 21, 23 (W.D. Pa. 1940).

16. Some courts have applied the attorney-client privilege not only to communications between expert witnesses and attorneys, but also to the expert's knowledge. See Friedenthal, *supra* note 14, at 462-69. One line of decisions based this application of the attorney-client privilege on the theory that the expert was acting as a mere "conduit," relaying and interpreting the client's information to the attorney. *Id.*; see *City & County of San Francisco v. Superior Court*, 37 Cal. 2d 227, 231 P.2d 26 (1951); *Webb v. Francis J. Lewald Coal Co.*, 214 Cal. 182, 4 P.2d 532 (1931); *Rust v. Roherts*, 171 Cal. App. 2d 772, 341 P.2d 46 (3d Dist. Ct. App. 1959). Another court characterized an expert as an assistant to an attorney and, therefore, held that the attorney-client privilege protected the expert from being subject to questioning at a deposition. *Cold Metal Process Co. v. Aluminum Co. of America*, 7 F.R.D. 684, 686 (D. Mass. 1947).

17. The attorney work product privilege prohibits litigants from discovering the "written statements and mental impressions of opposing counsel" absent a showing of necessity. See *Hickman v. Taylor*, 329 U.S. 495, 509 (1947). Some courts have extended this privilege to information that an attorney obtains from an expert, reasoning that this information becomes protected once it reaches the hands of an attorney. See, e.g., *United Air Lines, Inc. v. United States*, 26 F.R.D. 213, 218 (D. Del. 1960); *United States v. Certain Parcels of Land*, 25 F.R.D. 192, 193 (N.D. Cal. 1959); *White Pine Copper Co. v. Continental Ins. Co.*, 166 F. Supp. 148, 162-63 (W.D. Mich. 1958).

18. See Friedenthal, *supra* note 14, at 479-88. See also *United States v. 23.76 Acres of Land*, 32 F.R.D. 593, 596-97 (D. Md. 1963) (suggesting that the court can alleviate any unfairness in discovery costs by ordering the discovering party to pay).

19. Long, *supra* note 14, at 118.

20. *Id.* at 119.

courts treated discovery of expert witnesses inconsistently,²¹ courts that followed the predominant judicial trend either completely restricted discovery of experts or placed severe limitations upon it.²²

The primary factor that motivated courts to prohibit or restrict discovery of expert witnesses was the fear that it would promote lazy advocacy by allowing attorneys to rely on the work of opposing counsel to help them prepare their cases.²³ By imposing stringent limitations on discovery of expert witnesses, courts left attorneys inadequately prepared to face modern cases that concern complex, highly technical matters, and to counter the increased use of expert testimony that these cases often demand.²⁴ Attorneys, therefore, increasingly faced the prospect of cross-examining expert witnesses at trial without previously obtaining the information necessary to develop an effective strategy of rebuttal.²⁵ This situation totally subverted one of the main purposes of discovery: to "make a trial less a game of blind man's buff [sic] and more a fair contest with the basic issues and facts disclosed to the fullest practicable extent."²⁶

21. See *id.* at 117-19.

22. *Smith v. Ford Motor Co.*, 626 F.2d 784, 792 (10th Cir. 1980), *cert. denied*, 450 U.S. 918 (1981).

23. Friedenthal, *supra* note 14, at 488. The Advisory Committee for the Federal Rules, recognizing that this concern was the basis for the restrictions on discovery of expert witnesses, stated that "[p]ast judicial restrictions on discovery of an adversary's expert, particularly as to his opinions, reflect the fear that one side will benefit unduly from the other's better preparation." FED. R. CIV. P. 26(b)(4) advisory committee note.

24. Cases that concern complex subject matter often create a "battle of the experts" scenario. See Friedenthal, *supra* note 14; Getzoff, *Direct and Cross-examination of an Expert—Some Suggestions Concerning "How to Do It"*, 22 TRIAL LAW GUIDE 267, 267 (1978). In these cases, expert testimony often determines which party ultimately will prevail. FED. R. CIV. P. 26(b)(4) advisory committee note; see *United States v. Nysco Laboratories, Inc.*, 26 F.R.D. 159, 162 (E.D.N.Y. 1960).

25. In *United States v. 23.76 Acres of Land*, 32 F.R.D. 593 (D. Md. 1963), the court stated "that an expert is the most difficult witness to cross-examine, particularly if one is unaware until trial of the substance of his testimony." *Id.* at 596.

26. *United States v. Proctor & Gamble Co.*, 356 U.S. 677, 682 (1958). In light of this disturbing trend, Professor Jack Friedenthal, see Friedenthal, *supra* note 14, and others, see Long, *supra* note 14, posed vehement arguments for reform of laws governing discovery of expert witnesses. In addition, some courts became more receptive to the idea of discovering expert witnesses, and several opinions revealed greater judicial awareness of the relationship between pretrial discovery and effective cross-examination and rebuttal at trial. See *United States v. Meyer*, 398 F.2d 66 (9th Cir. 1968); *Franks v. National Dairy Prod. Corp.*, 41 F.R.D. 234 (W.D. Texas 1966); *United States v. 23.76 Acres of Land*, 32 F.R.D. 593 (D. Md. 1963). The increasing judicial awareness of the positive need for expert discovery, however, failed to rectify inconsistent treatment of the expert witness issue among jurisdictions. See *infra* notes 87-114 and accompanying text.

B. Enactment of Federal Rule 26(b)(4): Discovery of Expert Witnesses Permitted in Federal Courts

Congress responded to increasing judicial awareness of the need for discovery of expert witnesses by adopting rule 26(b)(4) as part of the 1970 amendment to the Federal Rules.²⁷ In addition to liberalizing discovery of expert witnesses, rule 26(b)(4) also contains provisions that prevent discovering parties from enjoying unfair benefits in the discovery process. For example, rule 26(b)(4) grants courts discretion to impose limited restrictions on the scope of discovery²⁸ and to order the discovering party to pay the party-opponent "a fair portion of the fees and expenses" that the party-opponent reasonably incurs "in obtaining facts and opinions from the expert."²⁹ The rule also limits discovery, except in special circumstances,³⁰ to expert witnesses the party-opponent expects to call at trial.³¹ Moreover, discovering parties may depose expert witnesses only by securing a court order.³² Thus, rule 26(b)(4)'s time limitations, which allow depositions of expert witnesses only after the party-opponent designates such witnesses for trial, coupled with the power it gives judges to limit the scope of disclosure and to require the discovering party to share in paying the expert witness' fees, effectively eliminates the danger of discovering parties benefiting unfairly from a party-opponent's trial preparation.

III. DEPOSING EXPERT WITNESSES: HIDDEN DANGER FOR DEPOSING PARTIES IN FEDERAL RULE 32(a)(3)

Although rulemakers drafted rule 26(b)(4) to keep parties who sought discovery of expert witnesses from benefiting unfairly, they ironically overlooked the possibility that courts, pursuant to rule 32(a)(3), would bestow inequitable rewards upon discovered parties. Rule 32(a)(3) provides that the "deposition of a witness . . . may be used by any party for any purpose . . ." if certain conditions arise—primarily unavailability of a witness to testify at

27. FED. R. CIV. P. 26(b)(4) advisory committee note.

28. FED. R. CIV. P. 26(b)(4)(A).

29. *Id.* at 26(b)(4)(C).

30. Rule 26(b)(4)(B) permits discovery of experts that a party-opponent does not expect to call at trial as provided in rule 35(b) "or upon a showing of exceptional circumstances under which it is impracticable for the party seeking discovery to obtain facts or opinions on the same subject by other means." *Id.* at 26(b)(4)(B).

31. *Id.* at 26(b)(4).

32. *Id.* at 26(b)(4)(A).

trial.³³ Rule 32(a)(3) does not "evince a distinction as to admissibility at trial between a deposition taken solely for purposes of discovery and one taken for use at trial . . .";³⁴ nor does a review of its legislative history reveal such a distinction.³⁵ If a party deposes a party-opponent's expert witness solely for discovery purposes in preparation for trial, therefore, the party-opponent can submit the deposition into evidence against the deposing party under rule 32(a) if the expert witness becomes unavailable to testify.³⁶ Consequently, the deposed party could admit into evidence the deposition of his unavailable expert's opinions and conclusions without subjecting them to in-court cross-examination by the deposing party.³⁷

A. *The Process of Discovering Expert Witnesses and the Deposition Problem*

A party who anticipates that a party-opponent plans to use an expert witness at trial typically submits an interrogatory to that party requiring him to identify any experts which he expects to call at trial and to state the subject matter and summarize the substance of the expert's anticipated testimony.³⁸ The responses to the interrogatories typically are not adequate to help the requesting party prepare an effective cross-examination.³⁹ Thus, the requesting party, by order⁴⁰ or by agreement with the party-opponent, usually deposes the expert identified in the party-opponent's interrogatory responses. Although the parties may specify whether the deposition is for discovery purposes only, for use at trial, or for unlimited purposes,⁴¹ these designations do not seem to affect the operation of rule 32.⁴² The Federal Rules do not require absolutely that a party who identified an expert as a trial witness produce the witness at trial because uncontrollable circumstances often prevent

33. *Id.* at 32(a)(3).

34. *Rosenthal v. Peoples Cab Co.*, 26 F.R.D. 116, 117 (W.D. Pa. 1960).

35. *See United States v. International Business Machs. Corp.*, 90 F.R.D. 377, 381 n.7 (S.D.N.Y. 1981); 4A J. MOORE, MOORE'S FEDERAL PRACTICE ¶ 30.50, at 30-38 (2d ed. 1981).

36. *See infra* notes 39-86 and accompanying text.

37. *See supra* notes 7-8 and accompanying text.

38. *See* FED. R. CIV. P. 26(b)(4)(A)(i).

39. Professor Graham's survey suggests that most practitioners find the interrogatory an unsatisfactory method of providing trial preparation. Graham, *Part Two, supra* note 12, at 172.

40. *See* FED. R. CIV. P. 26(b)(4)(A)(ii).

41. *See, e.g., Inland Bonding Co. v. Mainland Nat'l Bank*, 3 F.R.D. 438 (D.N.J. 1944).

42. *See infra* notes 87-114 and accompanying text.

the expert's attendance.⁴³ When such circumstances arise, therefore, the party who hired the expert can invoke rule 32(a)(3) and introduce into evidence the deposition that opposing counsel took for discovery purposes.

B. *The Unavailability Requirement of Rule 32(a)(3)*

The essential requirement under rule 32(a)(3) for admitting the deposition of a witness into evidence is unavailability of the deposed witness for testimony at trial. The unavailability requirements of rule 32(a)(3) reflect the traditional judicial preference for live testimony over recorded testimony⁴⁴ by permitting the admission of deposition evidence only when appearance of the witness is impossible or impractical, or when "exceptional circumstances" render admission of the deposition desirable.⁴⁵ The unavailability provisions of rule 32(a)(3) provide insight into the degree that the rule emphasizes the preference for live testimony, and illustrate the ways an expert deposition can become admissible at trial.

1. Rule 32(a)(3)(A): Death of the Deposed Witness

Rule 32(a)(3)(A) permits parties to use depositions for any purpose if the court finds that the deposed witness is dead. Admission of depositions under this provision has provoked very little controversy⁴⁶ because unavailability due to death is easy to demon-

43. See FED. R. CIV. P. 32(a)(3).

44. See, e.g., *Lamb v. Globe Seaways, Inc.*, 516 F.2d 1352 (2d Cir. 1975); *Salsman v. Witt*, 466 F.2d 76 (10th Cir. 1972). The traditional judicial preference for live testimony derives from the desire of courts to afford factfinders the opportunity to observe the demeanor of witnesses during testimony. See *Arnstein v. Porter*, 154 F.2d 464, 470 (2d Cir. 1946) (stating that "the demeanor of witnesses is recognized as a highly useful, even if not an infallible, method of ascertaining the truth and accuracy of their narratives."); *Napier v. Bossard*, 102 F.2d 467, 468-69 (2d Cir. 1939) (in which Judge Learned Hand stated that "[t]he deposition has always been, and still is, treated as a substitute, . . . not to be used when the original is at hand" because the party against whom the deposition is used is denied the advantage of having the witness appear in front of the jury). Federal Rules 32(a)(3)(E) and 43(a) expressly recognize the importance of allowing factfinders to observe the demeanor of witnesses. Rule 32(a)(3)(E), which governs the use of depositions in trial proceedings, emphasizes "the importance of presenting the testimony of witnesses orally in open court." Rule 43(a), which concerns the form and admissibility of evidence at trial, provides that "[i]n all trials the testimony of witnesses shall be taken orally in open court, unless otherwise provided . . . by these rules."

45. FED. R. CIV. P. 32(a)(3)(E).

46. Cases in which the deposed witnesses died before the deposition was complete, however, present a problem analogous to the principle issue in this Note. See cases cited *infra* note 87. For a full discussion of these cases, see Note, *supra* note 4.

strate and seldom is subject to abuse.⁴⁷ As with all other forms of unavailability under rule 32(a)(3), the party seeking to admit the deposition into evidence bears the burden of proving that the witness is dead.⁴⁸

2. Rule 32(a)(3)(B): Absence of the Deposed Witness

Rule 32(a)(3)(B), which permits parties to introduce a deposition into evidence when the witness is more than one hundred miles from the place of trial or hearing, has spawned more litigation than any other unavailability provision and probably is the most frequently abused criteria in the area of expert testimony. Many courts have interpreted the 100 mile rule, and most have construed it liberally.⁴⁹ The party offering the deposition into evidence bears the burden of establishing that the deponent meets the distance requirements for unavailability.⁵⁰ This burden of proof, however, is not difficult to overcome. In *Ikerd v Lapworth*,⁵¹ for example, the United States Court of Appeals for the Seventh Circuit held that the trial court could judicially notice that the witness' place of residence was more than 100 miles from the place of trial for purposes of admitting his deposition.⁵² More recently in *SCM Corp. v. Xerox Corp.*,⁵³ however, the Federal District Court of Connecticut adopted a narrower view of the 100 mile rule. The court held that it would not admit a deposition into evidence if the deposed witness either resided or regularly worked within 100 miles of the place of trial "when the deposition [was] offered [in evidence], or at a time during the proponent's case when a trial subpoena could have been served" upon the deposed witness.⁵⁴

47. The death of deposed witness provision in rule 32(a)(3) is not subject to abuse presumably because few parties would procure the death of a witness so that they could use a deposition instead of live testimony.

48. See, e.g., *National Screw & Mfg. Co. v. Voi-Shan Indus., Inc.*, 347 F.2d 1 (9th Cir. 1965) (holding that the lower court properly excluded a deposition from evidence when a party offered it pursuant to former rule 26 without showing the existence of circumstances making the rule applicable).

49. See *infra* notes 51-54 and accompanying text.

50. See, e.g., *Transcontinental Energy Corp. v. Pacific Energy Resources*, 683 F.2d 326, 330 (9th Cir. 1982).

51. 435 F.2d 197 (7th Cir. 1970).

52. *Id.* at 205. See also *Frederick v. Yellow Cab Co.*, 200 F.2d 483 (3d Cir. 1952) (court admitted deposition upon plaintiff's representation that the witness was out of town).

53. 77 F.R.D. 16 (D. Conn. 1977).

54. *Id.* at 18. See *Hartman v. United States*, 538 F.2d 1336, 1345 (8th Cir. 1976); *Ikerd v. Lapworth*, 435 F.2d at 205. *But see* *United States v. International Business Machs. Corp.*, 90 F.R.D. 377, 383 (S.D.N.Y. 1981).

The 100 mile requirement in rule 32(a)(3)(B) contains, as the sole means of controlling abuse, the qualification that a deposition is inadmissible when the party offering the deposition into evidence procures the witness' absence.⁵⁵ Courts, however, have been very reluctant to find that a party procured the absence of a deposed witness. Courts carefully have distinguished a party's act of procuring a witness' absence from the failure to facilitate his presence.⁵⁶ To exclude a deposition under this provision, a party apparently must meet the difficult burden of showing that the party who sought admission of the deposition actively tried to keep the deposed witness out of the courtroom.⁵⁷

Courts arguably should scrutinize attempted admissions of expert witness depositions more closely under the 100 mile requirement of rule 32(a)(3)(B) than attempted admissions of depositions of nonexpert witnesses. Today, parties frequently hire expert witnesses who live and work more than 100 miles from the place of trial.⁵⁸ These experts typically receive substantial compensation for consulting, testifying, and travel expenses.⁵⁹ Ordinary witnesses, however, usually receive only nominal statutory compensation because they accidentally observed a particular event, not because they possess any valuable expertise.⁶⁰ Thus, because experts are very costly and often live outside the 100 mile zone, parties who hire them are more likely to abuse the 100 mile requirement by not paying the expert to return for trial than parties who seek attendance of nonexpert witnesses.⁶¹ Nevertheless, most authorities treat expert and nonexpert witnesses identically under the 100 mile requirement for purposes of determining their unavailability and whether their depositions are admissible.⁶²

55. FED. R. CIV. P. 32(a)(3)(B).

56. See *Houser v. Snap-On Tools Corp.*, 202 F. Supp. 181, 189 (D. Md. 1962); see also *United States v. International Business Machs. Corp.*, 90 F.R.D. at 383 n.10 (citing *Houser* 202 F. Supp. at 189); *Richmond v. Brooks*, 227 F.2d 490, 493 (2d Cir. 1955) (court held that party did not procure his own absence); *M.S.D. Inc. v. United States*, 434 F. Supp. 85, 91 n.16 (N.D. Ohio 1977).

57. *Houser v. Snap-On Tools Corp.*, 202 F. Supp. at 189.

58. For a nationwide list of expert witness sources, see Kirk, *Locating Scientific and Technical Experts*, in 2 AM. JUR. TRIALS 293 (1964).

59. See Strodel, *The Expert Witness: The Cornerstone of the Medical Negligence Case*, TRIAL No. 6, June 1982, at 37, 38.

60. *Carter-Wallace, Inc. v. Otte*, 474 F.2d 529, 536 (2d Cir. 1972), cert. denied, 412 U.S. 929 (1973).

61. See *supra* notes 58 & 59 and accompanying text.

62. See *infra* notes 97-113 and accompanying text.

3. Rule 32(a)(3)(C): Age, Illness, Infirmary, and Imprisonment of the Deposed Witness

Rule 32(a)(3)(C) permits courts to admit a witness' deposition if "the witness is unable to attend or testify because of age, illness, infirmity, or imprisonment."⁶³ Again, the party seeking admission of the deposition bears a relatively light burden of proving that one of these factors prevents the witness' attendance at trial.⁶⁴ If a party can use this unavailability provision to admit a witness' deposition into evidence, other unavailability provisions such as the 100 mile and exceptional circumstances exceptions also may be available to the party.⁶⁵ If a witness is in prison, a party may seek to exclude the deposition by moving the court to use its discretion⁶⁶ to issue a writ of habeas corpus ad testificandum, which compels the prisoner to appear and testify at trial.⁶⁷

4. Rule 32(a)(3)(D): Inability to Procure Attendance of the Deposed Witness by Subpoena

If a party is unable to compel a deposed witness' attendance at trial by subpoena, the witness' deposition is admissible into evidence under rule 32(a)(3)(D).⁶⁸ This provision, together with the 100 mile provision in rule 32(a)(3)(B), presents uncertainty about whether a party who shows that a deposed witness is outside the 100 mile zone also must prove that he unsuccessfully tried to subpoena the witness. Rule 45(e)⁶⁹ suggests that witnesses who are more than 100 miles from the place of trial are not subject to subpoena. This simple answer, however, overlooks at least two situations that possibly could arise under the Federal Rules. First, rule 45(e) subjects witnesses to service of a subpoena at any place within the judicial district. Some judicial districts, because of their

63. FED. R. CIV. P. 32(a)(3)(C).

64. See *supra* notes 48 & 50 and accompanying text.

65. See *Scarfarotti v. Bache & Co.*, 438 F. Supp. 199, 202 n.3 (S.D.N.Y. 1977) (witness was ill at time of trial, and court admitted deposition, citing 32(a)(3)(B) and (E)).

66. See *Murray v. United States*, 138 F.2d 94, 97 (8th Cir. 1943).

67. On the other hand, a court need not let an imprisoned plaintiff come to trial when his deposition is available. See *Ball v. Woods*, 402 F. Supp. 803 (N.D. Ala. 1975), *aff'd without op.*, *Ball v. Shamblin*, 529 F.2d 520 (5th Cir.), *cert. denied*, 426 U.S. 940 (1976), *aff'd without op.*, *Ball v. Dwyer*, 538 F.2d 897 (5th Cir. 1976), *modified*, *Ball v. Woods*, 541 F.2d 279 (5th Cir. 1976).

68. FED. R. CIV. P. 32(a)(3)(D).

69. Rule 45(e) provides in pertinent part: "A subpoena . . . may be served at any place within the district, or at any place without the district that is within 100 miles of the place of the hearing . . ." *Id.* at 45(e).

geographical size, require witnesses to travel more than 100 miles from their place of residence to testify.⁷⁰ Second, in certain statutory causes of action, federal courts have nationwide subpoena power under rule 45(e)(1).⁷¹ In cases in which either of these two situations has arisen parties unsuccessfully have argued that the 100 mile provision which permits admissibility of depositions is inapplicable when the deposed witness is subject to a court's subpoena power.⁷² Federal courts, however, currently hold that parties may admit depositions of witnesses into evidence under the 100 mile requirement without demonstrating that they attempted to subpoena the witness.⁷³ Thus, the significance of the subpoena requirement in determining whether a deposed witness is unavailable is questionable, and very few cases construing the subpoena requirement have arisen.

The question whether a party can subpoena an expert witness to testify when his only connection with the case is that he consulted with one of the parties has gained considerable attention. If a party cannot compel such testimony, then any time he wishes to introduce the deposition of an expert witness the party simply must motion the court to issue a subpoena and, if the witness fails to respond, offer the deposition into evidence under rule 32(a)(3)(D). The United States Supreme Court stated in *Branzburg v. Hayes*⁷⁴ that "the public . . . has a right to every man's evidence," except for those persons protected by a constitu-

70. *SCM Corp. v. Xerox Corp.*, 76 F.R.D. 214, 215 (D. Conn. 1977).

71. See, e.g., 15 U.S.C. § 23 (1976) (nationwide subpoena power in antitrust suits when the United States is plaintiff). The note on amendments to rule 45(e)(1), which discusses justifications for giving federal courts nationwide subpoena power, states the following:

The last clause of the second sentence in Rule 45(e)(1) was added in 1980. Previously a person could not be subject to a subpoena served on him more than 100 miles from the place of the hearing even though the place of service and the site of the hearing were within the same state. It was felt, however, that state rules of service would adequately reflect the degree of hardship in travelling more than 100 miles within the state, and, therefore, there was no justification for greater federal restrictions.

72. *United States v. International Business Machs. Corp.*, 90 F.R.D. 377, 379-80 (S.D.N.Y. 1981); *SCM Corp. v. Xerox Corp.*, 77 F.R.D. 16, 17 n.2 (D. Conn. 1977); *Houser v. Snap-On Tools, Inc.*, 202 F. Supp. 181, 189 (D. Md. 1962). But see *United States v. Empire Gas Corp.*, 393 F. Supp. 903 (W.D. Mo. 1975), *aff'd*, 537 F.2d 296 (8th Cir. 1976), *cert. denied*, 429 U.S. 1122 (1977). In *Empire Gas* the court, in a federal antitrust action, refused to admit into evidence depositions that the government offered because the government did not attempt to show unavailability of deposed witnesses. *Id.* at 912. The court remarked that the government acted "in the face of its nationwide subpoena power," and, thus, suggested that the government should have attempted to subpoena the witnesses. *Id.*

73. See *supra* authorities cited in note 72.

74. 408 U.S. 665 (1972).

tional, common-law, or statutory privilege."⁷⁵ Experts under subpoena, however, do not enjoy a general privilege to refuse to testify concerning their expert knowledge.⁷⁶ Rather, courts generally will subpoena and compel an expert witness to testify concerning any previously formed expert opinions.⁷⁷ Most courts, however, probably would not force experts to testify against their will and, considering the value of an expert's time, would not subpoena an expert unless circumstances warranted otherwise.⁷⁸

5. Rule 32(a)(3)(E): Admissibility under Exceptional Circumstances

Rule 32(a)(3)(E) is a catch-all provision that permits courts to admit the deposition of a witness because of "exceptional circumstances."⁷⁹ Unlike other unavailability provisions of rule 32(a)(3) which require the deposed witness to be physically unavailable to testify at trial, the exceptional circumstances provision only requires the party offering the deposition into evidence to show the court that forcing a witness to testify at trial would be unfair. This burden seems more difficult because proving unfairness necessitates that parties satisfy more than simply factual tests which the other provisions in rule 32(a)(3)(A)-(D) establish.⁸⁰

In *United States v. Rollins*,⁸¹ for example, the federal government charged the appellee with selling heroin to an undercover agent. The government tried to avoid producing the undercover agent as a witness at trial. The government alleged "exceptional circumstances" for introducing the agent's deposition into evidence because the agent himself was under investigation, in an unrelated case, for permitting an informant to take some heroin that the agent had purchased.⁸² The government was unwilling to risk the serious danger that cross-examination of the agent unduly would have focused the investigation on the agent and caused harm to

75. *Id.* at 688 (quoting *United States v. Bryan*, 339 U.S. 323, 331 (1950)).

76. *Kaufman v. Edelstein*, 539 F.2d 811, 820 (2d Cir. 1976).

77. *See Carter-Wallace, Inc. v. Otte*, 474 F.2d 529, 536 (2d Cir. 1972), *cert. denied*, 412 U.S. 929 (1973).

78. *See Kaufman v. Edelstein*, 539 F.2d at 821. Parties that face the problem of having their discovery deposition of an expert witness used against them at trial should try to compel the expert to testify by moving the court to subpoena the expert.

79. FED. R. CIV. P. 32(a)(3)(E).

80. *See infra* notes 81-85 and accompanying text.

81. 487 F.2d 409 (2d Cir. 1973).

82. *Id.* at 411.

the agent's reputation.⁸³ The Second Circuit held that the agent was not unavailable within the meaning of rule 32(a)(3)(E)'s exceptional circumstances provision and thus, refused to admit the deposition.⁸⁴ The Second Circuit stated that a party could not argue "that a deposition would be usable if the witness were alive and well within the jurisdiction and subject to subpoena, simply because the party calling him wanted to preserve his identity or wished not to expose him to unfair cross-examination."⁸⁵ The *Rollins* decision suggests that the exceptional circumstances provision in rule 32(a)(3)(E) seldom will provide an avenue for parties to admit depositions of expert witnesses into evidence against the deposing party. Moreover, the language of rule 32(a)(3)(E), which contains the only statement in rule 32 that articulates a clear preference for live testimony over deposition evidence,⁸⁶ supports this suggested proposition. Thus, parties who wish to introduce depositions of witnesses into evidence probably will have to rely primarily on the provisions in rule 32(a)(3)(A)-(D) to achieve this objective.

83. *Id.*

84. *Id.* at 412.

85. *Id.* See also *Congoleum Industries, Inc. v. Armstrong Cork Co.*, 319 F. Supp. 714, 716 (E.D. Pa. 1970), a patent infringement suit, in which the court denied defendant's request to use at trial depositions of experts and other witnesses who would have observed the defendant conducting *inter partes* tests on certain commercial material it produced if the court had granted the defendant permission to conduct the tests. The defendant argued, under the exceptional circumstances provision of rule 32(a)(3)(E), that because its commercial equipment was too large to bring into and test in court, it could only give the court a present sense impression of the test procedures by taking depositions of the test witnesses while conducting the tests and introducing the depositions into evidence at trial. *Id.* In denying the defendant's request to use the depositions at trial, the court reasoned as follows:

Witnesses frequently, in fact almost always, testify to impressions and observations of events which take place outside the courtroom. The court must regularly make its decision from testimony concerning technical matters based on observations of events occurring outside of the courtroom. This is hardly an "exceptional circumstance". There is therefore no need to reduce defendant's experts' testimony to depositions before trial.

Id. But see *SCM Corp. v. Xerox Corp.*, 76 F.R.D. 214, 216 n.2 (D. Conn. 1977) (in which the court stated that the length of an antitrust suit was an exceptional circumstance that justified using the deposition of a witness at trial rather than compelling the witness to testify in court about "relatively non-controversial matters as to which the opportunity to assess credibility by observing the witness is not important.")

86. Rule 32(a)(3)(E) requires courts to give "due regard to the importance of presenting the testimony of witnesses orally in open court . . ." FED. R. CIV. P. 32(a)(3)(E).

IV. JUDICIAL PERCEPTIONS OF THE WITNESS DEPOSITION PROBLEM

Courts have expressed varying degrees of sensitivity for the arguments of parties who challenge the propriety of admitting depositions of witnesses into evidence against a deposing party. In several cases parties argued that allowing party-opponents to use depositions at trial denied them their right to cross-examine the unavailable witness.⁸⁷ Despite such arguments, courts universally have admitted these depositions into evidence while offering aggrieved deposing parties little or no protection from the inequities that resulted from these decisions.⁸⁸

In *Rosenthal v. Peoples Cab Co.*,⁸⁹ the Federal District Court for the Western District of Pennsylvania recognized the potential unfairness of admitting depositions of witnesses into evidence. The plaintiff in *Rosenthal* deposed the driver of defendant's taxicab solely for purposes of discovery.⁹⁰ The witness later died and defendant offered the deposition into evidence at the trial under rule 26(d), the predecessor to rule 32.⁹¹ Plaintiff moved to exclude the deposition, arguing that he deposed the taxicab driver only for discovery and that his counsel had no opportunity to cross-examine the driver with leading questions because the witness was not "hostile or recalcitrant."⁹² The court admitted the deposition despite plaintiff's arguments by reasoning that it did not have the power to read a nonexistent restriction into rule 26(d).⁹³ In recognition of defendant's inability to cross-examine the taxicab driver, however, the court admitted the deposition subject to "cautionary

87. See *infra* notes 89-113. A similar problem occurs concerning the admissibility of a deposition of a witness at trial when the witness dies before the deposition is complete. In these cases, parties have argued that the lack of an opportunity to cross-examine the witness before he died should foreclose admission of the deposition. See, e.g., *Derewecki v. Pennsylvania R.R. Co.*, 353 F.2d 436, 440 (3d Cir. 1965); *Continental Can Co. v. Crown Cork & Seal, Inc.*, 39 F.R.D. 354, 356 (E.D. Pa. 1965); *Inland Bonding Co. v. Mainland National Bank*, 3 F.R.D. 438, 439 (D.N.J. 1944). In each case, the court admitted the deposition by finding either that the objecting party consented to the omission of cross-examination or by holding that the value of the evidence outweighed the value of cross-examination. See Note, *supra* note 4, at 159-71.

88. See *infra* notes 89-114 and accompanying text.

89. 26 F.R.D. 116 (W.D. Pa. 1960).

90. *Id.* at 117.

91. As part of a general restructuring of discovery rules in 1970, the Federal Rules Advisory Committee moved the portions of 26(d)(3) concerning use of depositions at trial to rule 32(a)(3)(A). Note, *supra* note 4, at 155 n.17.

92. *Rosenthal*, 26 F.R.D. at 117. Ordinarily, counsel may not employ leading questions on direct examination except when witness appears hostile. See, e.g., W. RICHARDSON, RICHARDSON ON EVIDENCE § 483 (J. Prince ed. 1973).

93. *Rosenthal*, 26 F.R.D. at 117.

instructions . . . and such observations as the circumstance may require”⁹⁴ Thus, the *Rosenthal* court recognized the unfairness inherent in admitting discovery depositions of witnesses who have not undergone true cross-examination.⁹⁵ The court, however, felt that the silence of the Federal Rules on this issue constrained it and implicitly held that cautionary instructions were an adequate protection against any unfairness to the discovering party.⁹⁶

In *Wright Root Beer Co. v. Dr. Pepper Co.*,⁹⁷ however, the United States Court of Appeals for the Fifth Circuit expressed little concern for the plaintiff’s argument that the court should not admit the deposition of an inadequately cross-examined witness. Defendant Dr. Pepper, upon notice to plaintiff Wright, had taken the deposition of a witness whose testimony was crucial to Dr. Pepper’s case.⁹⁸ The witness died prior to trial and Dr. Pepper offered the deposition in evidence.⁹⁹ Plaintiff argued that although it had cross-examined the witness during the deposition, it had not done so vigorously because it felt that the deposition was to serve only discovery purposes.¹⁰⁰ The trial court admitted the deposition, but on several occasions told the jury that it could give less weight to the deposition than to live testimony which a party would have subjected to in-depth cross-examination.¹⁰¹ On appeal, the Fifth Circuit held that giving such instructions was reversible error.¹⁰² The court stated that under rule 26(d)(3),¹⁰³ “as a matter of right, a party may introduce the deposition of a deceased witness with no

94. *Id.* at 118.

95. *See infra* notes 115-23 and accompanying text.

96. *Rosenthal*, 26 F.R.D. at 118.

97. 414 F.2d 887 (5th Cir. 1969).

98. *Id.* at 889.

99. *Id.*

100. *Id.*

101. *Id.* at 889-90. When Dr. Pepper objected to the court’s cautionary instructions, the trial judge said:

I think it was proper to point out to the jury all of the circumstances that might be attended at the taking of a deposition for discovery purposes, versus a deposition for perpetuation of testimony. *Even with the presence of opposing counsel*, so forth, certainly at a deposition for discovery, it’s not unusual that you would not go into an in depth cross-examination that you would, where you are going to perpetuate the testimony. I think the jury was sufficiently charged on the point. I may not have mentioned specifically, I cannot remember, now, that definitely counsel for both parties were present.

Id. at 889-90 (emphasis added).

102. *Id.* at 890-91.

103. Old rule 26(d)(3) was the predecessor to rule 32(a)(3). *See supra* note 91.

strings attached."¹⁰⁴ The court reasoned that Wright's counsel purposefully chose to limit his cross-examination of the witness at the deposition and should have anticipated unexpected occurrences, "including the necessity for using depositions when the deponent has met an untimely death before trial."¹⁰⁵ The court noted that Wright failed to show that its failure to cross-examine the witness during the deposition, as it had a right to do, prejudiced it at trial.¹⁰⁶ In addition, the court stated that Wright did not prove that it could have obtained any evidence during cross-examination that would have impeached the witness' testimony.¹⁰⁷ Thus, unlike the district court in *Rosenthal*, the Fifth Circuit held that a party who fails to cross-examine a witness completely during a discovery deposition has no right to cautionary instructions when a party-opponent seeks to admit the deposition at trial.¹⁰⁸

The District Court for the Southern District of New York rendered the most recent decision on the expert witness deposition problem in *United States v. International Business Machines Corp. (IBM)*.¹⁰⁹ In *IBM* the government argued that it deposed several prospective IBM witnesses purely for discovery purposes to prepare for cross-examination.¹¹⁰ Moreover, the government argued that the court should not permit IBM to use the depositions against it at trial because rule 32(a)(3)(B) did not contemplate admission of discovery depositions.¹¹¹ The court, however, stated that rule 32 does not distinguish between depositions for discovery purposes and those for use at trial.¹¹² The court admitted the government's deposition into evidence, reasoning that "admission of un-

104. *Wright Root Beer Co. v. Dr. Pepper Co.*, 414 F.2d at 890.

105. *Id.*

106. *Id.*

107. *Id.*

108. *Id.* at 890-91. Eleven years later, the Fifth Circuit in *Savoie v. LaFourche Boat Rentals, Inc.*, 627 F.2d 722 (5th Cir. 1980), again demonstrated its unwillingness to go beyond the express language of rule 32(a)(3) by admitting the deposition of a witness against the deposing party. In *Savoie* the defendant had deposed a witness who was outside the United States during the trial and argued that admission of the deposition against it would act as a disincentive for parties in future litigation to prepare properly for trial by taking discovery depositions. *Id.* at 724. The Fifth Circuit, in a per curiam opinion, held that the language of rule 32(a)(3) foreclosed this type of argument, *id.*, and admitted the deposition under the relevant provision of rule 32(a)(3)(B). *Id.*

109. 90 F.R.D. 377 (S.D.N.Y. 1981).

110. *Id.* at 381.

111. *Id.*

112. *Id.* at 381. The court noted that the 1970 revisions to the Federal Rules removed language from the Rules that distinguished between depositions purely for discovery purposes and those for use at trial. The court perceived that this revision extinguished any

favorable deposition records was a risk the government assumed when it chose to limit its questioning."¹¹³

The courts that have addressed the witness deposition problem increasingly have been unwilling to consider seriously the potential prejudice to the party against whom the court admits the deposition. These courts apparently have embraced the idea that parties who fail to conduct full cross-examinations during depositions assume the risk of having depositions of subsequently unavailable witnesses used against them at trial. This theory fails to recognize the distinction between discovery and preparation of evidence that exists in practice although rule 32(a)(3) fails to reflect it.¹¹⁴ The practical difference between these two procedures makes use of expert witness depositions at trial unfair when the parties intended to use them only for discovery.

V. ANALYSIS

A. *The Expert Witness Deposition Problem: The Degree of Unfairness*

1. Absence of Proper Opportunity for Cross-Examination

Parties in the American judicial system enjoy an absolute right to cross-examine opposing witnesses.¹¹⁵ The purpose of cross-examination is to give factfinders an opportunity to evaluate the reliability of a witness' testimony by allowing the cross-examining party to impeach the witness and elicit facts favorable to the cross-examiner's case.¹¹⁶ The prior opportunity for cross-examination of a witness, therefore, is the primary factor underlying the admission of hearsay under the prior testimony exception.¹¹⁷ Similarly, this opportunity also justifies the provisions of rule 32(a)(3) that allow courts to admit depositions as substantive evidence when the deponent subsequently becomes unavailable to testify at trial.¹¹⁸ Ad-

argument for a distinction between discovery and evidentiary depositions. *Id.* at 381 n.7. See *Rosenthal v. Peoples Cab Co.*, 26 F.R.D. at 117.

113. *United States v. International Business Machs. Corp.*, 90 F.R.D. at 381.

114. See *infra* notes 125-29 and accompanying text.

115. See *supra* note 7.

116. 4 B. JONES, *supra* note 7, at § 25:1.

117. See *infra* note 122 and accompanying text. "According to Dean McCormick, 'the premise that the opportunity of cross-examination is an essential safeguard has been the principle justification for the exclusion generally of hearsay statements, and for the admission as an exception to the hearsay rule of reported testimony taken at a former hearing.'" Note, *supra* note 4, at 156 n.19 (citing C. MCCORMICK, MCCORMICK'S HANDBOOK ON THE LAW OF EVIDENCE § 19, at 43 (2d ed. E. Cleary 1972)).

118. See Note, *supra* note 4, at 156 n.19.

mission of discovery depositions into evidence against deposing parties who only directly examine the deponent, therefore, infringes upon the deposing parties' right of cross-examination.

(a) *Dependence of Admissibility on Prior Cross-Examination*

Rule 32(a)(3) does not contain an express requirement that cross-examination occur as a prerequisite to admissibility of a deposition. The rule, however, does permit use of depositions at trial against only parties who were "present or represented . . . or who had reasonable notice" of the taking of the deposition.¹¹⁹ This language implies that the party against whom a party-opponent seeks to offer the deposition into evidence must have had an opportunity to temper the deposition's impact through cross-examination.

Depositions, however, basically remain a form of hearsay,¹²⁰ and courts and commentators have agreed that the total absence of cross-examination precludes admission of a deposition.¹²¹ The Federal Rules of Evidence recognize the fundamental hearsay nature of depositions and allow their admission only against a party who previously "had an opportunity and similar motive to develop the testimony by direct, cross, or redirect examination."¹²² Because courts and commentators have read an opportunity to cross-examine requirement into rule 32(a),¹²³ courts should admit depositions into evidence only against parties who had the opportunity to "develop the testimony" through cross-examination. This requirement will ensure a deposition's reliability as evidence, and thus, its admissibility as an exception to the hearsay rule.¹²⁴

119. FED. R. CIV. P. 32(a).

120. 3 B. JONES, *supra* note 7, § 18:26 at 498.

121. See 4A, J. MOORE, MOORE'S FEDERAL PRACTICE ¶ 30.58, at 30-133 (2d ed. 1983); *Bobb v. Modern Products, Inc.*, 648 F.2d 1051, 1055 (5th Cir. 1981). For example, in *Du-Beau v. Smither & Mayton, Inc.*, 203 F.2d 395, 396-97 (D.C. Cir. 1953), the court excluded the deposition of a witness from evidence because the witness refused to disclose his residence or occupation when cross-examined at the deposition proceedings. The court reasoned that the deponent's refusal to reveal his place of abode or occupation substantially deprived plaintiff of her right of cross-examination. *Id.* at 397.

122. FED. R. EVID. 804(b)(1). "Former testimony does not rely upon some set of circumstances to substitute for oath and cross-examination, since both oath and opportunity to cross-examine were present in fact." *Id.* advisory committee note.

123. See *supra* note 121 and accompanying text.

124. Dean McCormick contends that lack of any opportunity to cross-examine an absent declarant is one of the main justifications for the hearsay exclusion of evidence:

It would be generally agreed today that the third factor is the main justification for the exclusion of hearsay. This is the lack of any opportunity for the adversary to cross-examine the absent declarant whose out-of-court statement is reported by the witness. Thus as early as 1668 we find a court rejecting hearsay because "the other party could

(b) *Adequacy of Opportunity for Cross-Examination*

The evaluation of whether a discovering party had an adequate opportunity to cross-examine a deposed witness requires an understanding of the practical difference between taking a deposition for discovery purposes and taking one for evidentiary purposes. In choosing to depose a party-opponent's expert witness, a party obviously does not desire to preserve the contents of the deposition as evidence for use at trial because the expert's opinion probably is harmful to the discovering party's position.¹²⁵ Rather, a party's primary purpose for deposing an expert witness is to ascertain the crux of the expert's future testimony, which often exceeds the scope of the deposing party's expertise, and to prepare an effective cross-examination strategy that rebuts the expert's testimony.¹²⁶ When taking evidentiary depositions, however, the deposing party's counsel invariably seeks to preserve the contents of the deposition for use at trial in the light most favorable to his client—a task that is very difficult when examining an adverse expert

not cross-examine the party sworn." Judicial expressions stress this as a principal reason for the hearsay rule. Cross-examination, as Bentham pointed out, was a distinctive feature of the English trial system, and the one which most contributed to the prestige of the institution of jury trial. He called it "a security for the correctness and completeness of testimony." The nature of this safeguard which hearsay lacks is indicated by Chancellor Kent: "Hearsay testimony is from the very nature of it attended with . . . doubts and difficulties and it cannot clear them up. 'A person who relates a hearsay is not obliged to enter into any particulars, to answer any questions, to solve any difficulties, to reconcile any contradictions, to explain any obscurities, to remove any ambiguities; he entrenches himself in the simple assertion that he was told so, and leaves the burden entirely on his dead or absent author.' . . . The plaintiff by means of this species of evidence would be taken by surprise and be precluded from the benefit of a cross-examination of S. as to all those material points which have been suggested as necessary to throw full light on his information." Similarly, a Georgia judge has said that cross-examination "is the most efficacious test which the law has devised for the discovery of truth." Morgan analyzed the protective function of cross-examination and concluded (1) that while the fear of exposure of falsehoods on cross-examination is a stimulus to truth-telling by the witness, actual exposure of wilful falsehood is rarely accomplished in actual practice and (2) that the most important service of cross-examination in present day conditions is in affording the opportunity to expose faults in the perception and memory of the witness.

C. McCORMICK, *supra* note 116, at 583 (footnotes omitted).

125. The opponent would not hire an expert to testify at trial unless the expert's testimony is favorable to his case.

126. See, e.g., *Kerr v. United States Dist. Court*, 511 F.2d 192, 196-97 (9th Cir. 1975), *aff'd*, 426 U.S. 929 (1976); *United States v. Meyer*, 398 F.2d 66, 72 (9th Cir. 1968); see also *Hoover v. United States Dept. of Interior*, 611 F.2d 1132, 1141-42 (5th Cir. 1980) (discussing the "entitlement of right" to discover substance of the expert's anticipated testimony under rule 26(b)(4)(A)).

whose area of expertise is foreign to the deposing attorney.¹²⁷

Strategic concerns also may play an important role in defining the scope of questioning an expert receives during a deposition.¹²⁸ Because attorneys sometimes are unknowledgeable in an expert's specialty, they may choose not to question or challenge the expert on certain matters until they can familiarize themselves with the expert's ideas through further research or consultation with their own experts. Similarly, counsel also may limit cross-examination of an expert to avoid revealing information that he has gathered while preparing his case,¹²⁹ thereby forcing the opponent to elicit this information through its own discovery measures.

By asserting that parties have an adequate opportunity to cross-examine and rebut an expert's testimony while deposing him, courts completely ignore the reasons that compel parties to depose expert witnesses and the realities of the deposition process. The judicial theory that parties waive their right to cross-examine a deposed expert witness who later becomes unavailable by not cross-examining the expert with full vigor during the deposition is even more untenable.¹³⁰ Parties often lack sufficient preparation while taking depositions before trial to cross-examine party-opponents' expert witnesses effectively. In fact, the desire to prepare for cross-examination often is the key factor that motivates a party to depose a witness. By admitting discovery depositions of expert witnesses against discovering parties under rule 32(a)(3), therefore, courts not only have controverted the essential information-gathering and surprise-squelching purposes of pretrial discovery, but also effectively and erroneously have eliminated the process of deposing the witnesses of a party-opponent as a pretrial discovery technique. Instead, the courts effectively have recharacterized deposing of witnesses as an extension of formal trial proceedings which presume that the deposing party has engaged in full pretrial preparation. This ill-founded scenario is particularly egregious when the deposed witness is an expert whose testimony encompasses knowledge of which the deposing party is ignorant.

127. An expert, by definition, is well-versed in the particular matters of his testimony. Opposing counsel, even after availing himself of self-education, usually lacks sufficient preparation to discredit the expert's testimony.

128. See Kennelly, *Pretrial Discovery—The Courts and Trial Lawyers Are Finally Discovering That Too Much of It Can Be Counterproductive*, 21 TRIAL LAW. GUIDE 458, 474-89 (1977).

129. "Pretrial interrogation may serve only to educate the witness so that cross-examination becomes ineffectual at trial." *Id.* at 480.

130. See *supra* notes 97-113 and accompanying text; Note, *supra* note 4, at 159-66.

2. Practical Implications of the Expert Witness Deposition Problem

(a) *Potential for Abuse*

In most cases concerning the expert witness deposition problem under rule 32(a)(3), the parties either did not expect the deposed witness to become unavailable or they anticipated the unavailability and took the expert's deposition to preserve the evidence.¹³¹ Rule 32(a)(3), however, allows a party to permit opponents to depose his witness for discovery purposes while he never intends to compel the witness' presence at trial. Similarly, a party could choose this strategy after a party-opponent completes the deposition. A decision to submit the deposition into evidence and not to produce the witness to testify, especially in the context of expert witnesses, abuses the equitable purposes of discovery under the Federal Rules, particularly when the party who seeks admission of the deposition has identified the expert as a trial witness.¹³² In this situation the deposed party effectively has placed the discovering party in a trial-type situation that erroneously presumes the discovering party has prepared fully for cross-examination of the expert, and depositions occurring under these circumstances generally favor the deposed party.¹³³

(b) *Inhibition of Discovery*

As the appellant argued in *Savoie v. LaFourche Boat Rentals, Inc.*,¹³⁴ courts discourage parties from deposing the expert witnesses of party-opponents by admitting these depositions into evidence if the witnesses later become unavailable to testify at trial.¹³⁵ Courts in such cases clearly have stated that parties who depose an opponent's expert witness must anticipate the possibility of the witness subsequently becoming unavailable, thereby necessitating admission of the deposition against them.¹³⁶ These courts, however, ironically are forcing discovering parties to be fully prepared to cross-examine the party-opponent's expert witness at a stage in

131. See *supra* notes 128-29 and accompanying text.

132. See *supra* text accompanying note 30.

133. Questioning that attempts to record the expert's opinion and the basis for it often does not elicit any significant discrediting testimony. See, e.g., *United States v. International Business Mach.*, 90 F.R.D. 377, 381 (S.D.N.Y. 1981).

134. 627 F.2d 722 (5th Cir. 1980).

135. See *supra* note 108.

136. See *Wright Root Beer Co. v. Dr. Pepper Co.*, 414 F.2d 887, 890 (5th Cir. 1969); *United States v. International Business Mach.*, 90 F.R.D. 377, 381 n.6 (S.D.N.Y. 1981).

pretrial discovery that should be producing the information necessary to prepare for such cross-examination. Because parties usually lack sufficient information to engage in meaningful cross-examination of expert witnesses before deposing them, courts in this situation effectively are discouraging parties from deposing an opponent's expert witnesses.

*B. Potential Solutions to the Expert Witness
Deposition Problem*

Future courts, parties, and rulemakers could employ several strategies to reduce the chances of parties using discovery depositions of expert witnesses against discovering parties at trial and to reduce the degree of unfairness that stems from such use. These strategies include the following: (1) court orders limiting the use of depositions at trial; (2) stipulation by the parties that limits the use of depositions to discovery; (3) cautionary instructions by the trial judge warning jurors to scrutinize deposition evidence carefully; (4) a stricter judicial test for determining unavailability of expert witnesses; (5) local court rules that mitigate unfair use of expert witness depositions; and (6) amendment to the Federal Rules addressing the expert witness deposition problem. Courts, legislators, and rulemakers who currently are considering the problems with and potential reforms of the discovery process should consider adopting one or more of these approaches to protect parties who wish to depose a party-opponent's expert witness against unfairness. Similarly, trial counsel operating under the Federal Rules or similar state rules should recognize the limited protections that already exist.

1. Court Orders Limiting Use of Depositions

District courts possess broad discretionary powers to make discovery and evidentiary rulings in the interest of conducting fair and orderly trials.¹³⁷ This discretion includes "the power to exclude or admit expert testimony . . . and to exclude testimony of witnesses whose use at trial is in bad faith or would unfairly prejudice an opposing party."¹³⁸ This Note advocates that the unfairness inherent in using the discovery deposition of an expert witness against the deposing party constitutes a colorable claim of

137. See *United Air Lines, Inc. v. United States*, 26 F.R.D. 213, 217-18 (D. Del. 1960); *Dipson Theatres, Inc. v. Buffalo Theatres, Inc.*, 8 F.R.D. 313, 314 (W.D.N.Y. 1948).

138. See cases cited *supra* note 137.

prejudice against which a judge could exercise his discretionary powers by excluding the deposition from evidence. Procedurally, parties may seek such a court order pursuant to the "protective order" provision of rule 26(c) before deposing the expert witness.¹³⁹ The moving party simply should show the court that it lacks sufficient preparation to conduct extensive cross-examination and accordingly, that the court should limit the deposition's use to discovery purposes.

Although the judicial power to issue protective orders is discretionary, at least two reasons make it unlikely that courts will issue such orders against the use of expert witness depositions at trial. First, past judicial treatment of expert witness discovery depositions indicates an unwillingness to recognize any distinction between discovery and evidentiary depositions.¹⁴⁰ Because courts traditionally have not distinguished these types of depositions and have not acknowledged the prejudice that results from using discovery depositions against deposing parties, courts probably would refuse to order any limitations on the use of depositions. Second, some courts might construe explicit unavailability provisions in rule 32(a) to prohibit this type of protective order.¹⁴¹ Nevertheless, discovering parties still should seek protective orders because, even if a court denies the party's motion, the mere attempt to restrict use of the deposition bolsters any subsequent argument of prejudice that the discovering party can make if his opponent later tries to offer the deposition into evidence. A discovering party could argue that he had not waived his right to cross-examine the unavailable witness because he had notified the court prior to taking the deposition that he was limiting his questioning to discovery. A court, however, could argue that its denial of the motion put counsel on notice of the deposition's potential use at trial and that the party therefore had waived his opportunity to cross-examine the expert witness. Thus, because courts issue protective orders in only limited circumstances, they do not protect discovering parties adequately.¹⁴²

139. Rule 26(c) provides in pertinent part: "Upon motion by a party or by the person from whom discovery is sought, and for good cause shown, the court . . . may make any order which justice requires to protect a party or person from annoyance, embarrassment, oppression, or undue burden or expense . . ." FED. R. CIV. P. 26(c).

140. See *supra* note 34 and accompanying text.

141. See *supra* note 44-86 and accompanying text.

142. Another argument in favor of protective orders is that courts increasingly have exercised control over the discovery phase of litigation. See Sherman & Kinnard, *Federal Court Discovery in the 80's—Making the Rules Work*, 95 F.R.D. 245, 247, 269 (1979). This

2. Stipulation by the Parties

Another strategic option currently available to discovering parties is stipulation with party-opponents limiting the use of expert witness depositions to discovery.¹⁴³ To make such a stipulation, discovering parties should request that party-opponents join them in submitting a written stipulation which states that neither party intends to introduce the deposition into evidence. The obvious problem with stipulations is that an opponent simply may refuse to enter into the agreement. As a practical matter, the party who hired the deposed expert witness should try to preserve the right to use the deposition at trial if the expert witness subsequently becomes unavailable. Thus, stipulations essentially are almost perfect¹⁴⁴ protective devices that parties seldom agree to use. If both parties seek to depose their opponents' expert witnesses, however, the likelihood that they will reach a stipulation agreement is substantially higher than if only one party deposes an expert because both parties may wish to prevent the other's use of the deposition at trial.

3. Cautionary Instructions

Judges can use cautionary instructions to warn juries to scrutinize evidence from expert witness depositions more carefully than live trial testimony because proper cross-examination of the deponents was impossible. Courts theoretically can prevent any prejudice resulting from admission of an expert witness deposition by issuing such instructions.¹⁴⁵ Two problems, however, make cautionary instructions an inadequate protection against prejudice. First, courts probably will not issue cautionary instructions concerning the reliability of deposition evidence because of the hold-

development, however, does not portend necessarily increased judicial sympathy for parties facing the expert witness deposition problem.

143. Very little authority exists concerning availability of this type of stipulation. If parties stipulated that they would use a certain deposition only for discovery purposes, however, the court apparently would have to exclude it from evidence. See *Stevens v. Illinois Cent. R.R.*, 157 Ky. 561, 570, 163 S.W. 747, 751 (1914). Rule 29 of the Federal Rules of Civil Procedure permits stipulations waiving the procedural requirements of depositions. FED. R. CIV. P. 29. Parties frequently use rule 29 to make a deposition admissible despite noncompliance with procedural formalities. See C. WRIGHT & A. MILLER, *FEDERAL PRACTICE AND PROCEDURE* § 2091 (1970).

144. Courts have a discretionary power to set aside stipulations. See FED. R. CIV. P. 29.

145. See Note, *The Limiting Instruction—Its Effectiveness and Effect*, 51 MINN. L. REV. 264, 264 (1966).

ing in *Wright Root Beer*¹⁴⁶ that such cautionary instructions constitute reversible error.¹⁴⁷ Although the Fifth Circuit cited no authority in *Wright Root Beer* for the proposition that the trial court may not issue cautionary instructions concerning admission of an unavailable witness' deposition,¹⁴⁸ no subsequent decisions have limited the *Wright Root Beer* holding. Second, empirical studies cast substantial doubt on the effectiveness of cautionary instructions. Typically, cautionary instructions limit the jury's consideration of evidence to the use for which the evidence is competent.¹⁴⁹ Various studies, however, have suggested that cautionary instructions do not influence juries significantly to limit their consideration of questioned evidence.¹⁵⁰ In fact, one study suggests that a judge's cautionary instructions to a jury actually sensitize the jury to the evidence.¹⁵¹ Furthermore, many "learned jurists" have concluded that limiting instructions are useless because jurors cannot distinguish proper and improper uses of evidence, even if they remember and fully understand the limiting instructions.¹⁵² Thus, parties should not rely on cautionary instructions as the sole protection against use of discovery depositions at trial. Rather, parties probably should request cautionary instructions only as a last resort when they believe a court will not exclude the deposition.

4. A Stricter Test for Unavailability of Expert Witnesses

The ease with which a party can show unavailability of a deposed witness under rule 32(a)(3) and then use the deposition at trial against the deposing party is the primary cause of the expert witness deposition problem.¹⁵³ In *Carter-Wallace, Inc. v. Otte*¹⁵⁴ the United States Court of Appeals for the Second Circuit sug-

146. 414 F.2d 887 (5th Cir. 1969).

147. For a discussion of the *Wright Root Beer* decision see notes 97-108 and accompanying text.

148. *Wright Root Beer*, 414 F.2d at 890-91. See E. DEVITT & C. BLACKMAR, 2 FEDERAL JURY PRACTICE AND INSTRUCTIONS § 72.02 (2d ed. 1970). *Contra* *Treharne v. Callahan*, 426 F.2d 58, 64 (3d Cir. 1970); *Rosenthal v. Peoples Cab Co.*, 26 F.R.D. at 117-18.

149. For example, a court may admit a hearsay statement if the party offers it for a purpose other than proving the truth of the statement itself. See FED. R. EVID. 801(c) advisory committee note.

150. See, e.g., Broeder, *The University of Chicago Jury Project*, 38 NRB. L. REV. 744 (1959); Note, *supra* note 145, at 265-69.

151. Broeder, *supra* note 150, at 754.

152. See Note, *supra* note 145, at 267 n.18 and accompanying text.

153. See *supra* notes 44-86 and accompanying text.

154. 474 F.2d 529 (2d Cir. 1972), *cert. denied*, 412 U.S. 929 (1973).

gested that "there is something unusual about the use of the prior testimony of an expert witness that calls for further scrutiny of his unavailability."¹⁵⁵ The court recognized that when a party previously has hired an expert witness for various purposes, the court should not label that expert "unavailable" because the party chose not to arrange for the expert's presence at trial.¹⁵⁶ The Second Circuit, therefore, was advocating implicitly a stricter test for determining unavailability of an expert witness under rule 32(a)(3).

A stricter unavailability test is sensible for two reasons. First, a stricter test would place a heavier burden of proving unavailability on parties seeking to use an expert witness deposition at trial and, consequently, would reduce the chance of parties abusing the unavailability rule.¹⁵⁷ Under a stricter unavailability test for expert witnesses, courts could require parties to give more compelling reasons than mere distance to justify a witness' absence. Second, by adopting a stricter unavailability test in civil cases, courts would not be dealing with an unfamiliar requirement because a strict unavailability test already is operative for all witnesses in criminal cases.¹⁵⁸

Despite the potential benefits of a stricter unavailability requirement, however, courts are unlikely to adopt this standard because the *Carter-Wallace* decision has had only a limited impact in this area.¹⁵⁹ An amendment to the Federal Rules probably would be a prerequisite to the widespread adoption of a stricter unavailability test for expert witnesses.

5. Local Court Rules

Federal district court judges may adopt by majority action rules governing local practice that are consistent with the Federal Rules of Civil Procedure.¹⁶⁰ Thus, district courts potentially could frame local rules that mitigate in a variety of ways unfair use of

155. *Id.* at 536. The special qualities of expert testimony augment the argument that use of discovery deposition testimony of experts is unfair.

156. *Id.*

157. See *supra* notes 131-33 and accompanying text.

158. See, e.g., *Ohio v. Roberts*, 448 U.S. 56 (1980); *United States v. Sindona*, 636 F.2d 792, 803-04 (2d Cir. 1980), *cert. denied*, 451 U.S. 912 (1981); *United States v. Mann*, 590 F.2d 361, 365-68 (1st Cir. 1978); FED. R. CRIM. P. 15(a).

159. No courts have adopted the stricter standard of unavailability that *Carter-Wallace* implies, perhaps because that court held that the lower court's failure to impose such a standard was harmless error. See *Arrow-Hart, Inc. v. Covert Hills, Inc.*, 71 F.R.D. 346, 348 (E.D. Ky. 1976).

160. FED. R. CIV. P. 83.

expert witness discovery depositions. A rule that required parties to designate the purpose of all expert depositions and that prohibited parties from introducing "discovery" depositions into evidence certainly would curtail the expert witness deposition problem. Similarly, a rule requiring parties to make binding designations at a time before completion of discovery concerning which depositions they would introduce into evidence at least would offer the discovering party a chance to conduct a second deposition with full cross-examination. The district courts also could adopt other strategies for addressing this problem with local rules, such as cautionary instructions or a stricter unavailability standard.

The principle difficulty in solving the expert witness deposition problem with local rules is that they must be consistent with the Federal Rules.¹⁶¹ Arguably, any locally imposed limitation on the operation of rule 32(a) would be invalid as inconsistent with the Federal Rules.¹⁶² District courts could avoid this problem by making enforcement of the local rule discretionary.¹⁶³ A discretionary local rule, however, would provide no more protection than already is available under a protective order.¹⁶⁴ The doubtful success of local rules solving the expert witness deposition problem and the lack of uniformity among local rules that probably would occur throughout the district courts make local rule changes an unsatisfactory solution to this problem.

6. Amendment to the Federal Rules of Civil Procedure

An amendment to the Federal Rules offers the most uniform and effective approach for resolving the expert witness deposition problem. Again, several possible approaches might work. The Advisory Committee could recommend an amendment to rule 32 that required courts to distinguish between discovery and evidentiary depositions. The amendment could exclude from the current admissibility provisions of rule 32 any deposition that a party takes expressly for discovery purposes. Alternatively, a revised rule 32 could require courts to treat admissibility of expert witness and nonexpert witness depositions differently, perhaps by providing that courts never can consider expert witnesses unavailable for

161. *Id.*

162. The language of rule 32 does not imply a limitation concerning the type of deposition admissible under its provisions. See *supra* notes 44-86.

163. Cf. Sherman & Kinnard, *supra* note 142, at 264 n.76 and accompanying text (discussing the validity of local rules limiting the total number of interrogatories).

164. See *supra* notes 137-42 and accompanying text.

trial under rule 32(a). This amendment effectively would allow parties to introduce expert witness depositions only by agreement.

The best solution to the expert witness deposition problem is to amend rule 26(b)(4), which already governs discovery of expert witnesses. Rule 26(b)(4) recognizes that parties need an adequate opportunity for discovery of expert witnesses before trial to cross-examine and rebut their testimony at trial.¹⁶⁵ To achieve this end, rule 26(b)(4)(A) permits parties to discover any expert witness that a party-opponent expects to call as a witness at trial. An amendment to rule 26(b)(4) could make the deposition process fair for deposing and deposed parties by specifying that if a deposed expert whom a party had designated to be a trial witness subsequently became unavailable to testify at trial, any deposition which a party-opponent previously took of the expert would be inadmissible upon motion pursuant to rule 26(b)(4)(A)(ii) or by agreement with the other party. This amendment simply would make binding the designation of an expert as a trial witness and would prevent deposed parties from using the discovery efforts of deposing parties as evidence. An amendment of this type would not prejudice parties who want to call expert witnesses to testify at trial because they still would retain the option of preserving the expert's testimony through direct examination during depositions as protection against subsequent unavailability of the witness. If a party should choose to preserve expert testimony in this manner, the opponent would at least be on notice of the deposition's intended use and would have an opportunity to prepare for cross-examination of the expert through discovery.

This proposed change in rule 26 would reorder priorities by allowing the risk of an expert witness' becoming unavailable to fall upon the proponent of the expert's testimony. Under the current Federal Rules, an expert's absence probably is more harmful to the opponent of the expert's testimony, who inadvertently may have created unfavorable evidence by deposing the expert in preparation for trial.

VI. CONCLUSION

The degree to which courts inhibit discovery and prejudice parties by admitting expert witness depositions into evidence under rule 32(a)(3) when the expert becomes unavailable to testify at trial is difficult to measure. Because the subject matter of con-

165. See *supra* note 6.

temporary civil litigation has become highly technical and complex, parties increasingly have used expert witnesses at trial for a variety of purposes. This phenomenon suggests that the expert witness deposition problem should be a common problem in civil cases in federal courts. Although only a few cases have addressed this problem, their holdings probably discourage parties from arguing against admission of expert witness depositions regardless of how much the circumstances of a case favor inadmissibility.

The expert witness deposition problem presents a strong possibility of denying discovering parties an adequate opportunity to cross-examine unavailable expert witnesses. During expert witness depositions, parties usually seek information that will help them prepare an effective strategy for cross-examining the expert at trial. Thus, requiring deposing parties to cross-examine experts during depositions, before they have acquired the information necessary to conduct effective cross-examination, inevitably discourages parties from deposing expert witnesses. The expert witness deposition problem controverts one of the basic purposes underlying the liberal system of discovery in federal courts—to “make a trial less a game of blind man’s buff [sic] and more a fair contest with the basic issues and facts disclosed to the fullest practicable extent.”¹⁶⁶ In addition, courts, by perpetuating this problem, erroneously have eliminated the process of deposing a party-opponent’s designated trial witnesses as a pretrial discovery technique by effectively treating the process as an extension of formal trial proceedings that presume the deposing party has prepared fully to cross-examine the expert. Thus, courts and rulemakers should take active steps to remedy the expert witness deposition problem because it threatens to undermine the process of deposing witnesses before trial, one of the most important pretrial discovery techniques available to parties in federal courts.

STEVEN D. PARMAN

166. *United States v. Proctor & Gamble Co.*, 356 U.S. 677, 682 (1958).

