HARVARD JOURNAL
on
LEGISLATION

VOLUME 32, NUMBER 1 WINTER 1995

HARVARD JOURNAL ON LEGISLATION
Harvard Law School
Cambridge, MA 02138

Phone: (617) 495-4400
Internet: HJOL@HULAW1.HARVARD.EDU

The HARVARD JOURNAL ON LEGISLATION (ISSN 0017-808x) is published twice each year (winter and summer) by Harvard Law School students. Third-class postage paid at Boston, Massachusetts, and at additional mailing offices.

Publication Policy

The JOURNAL specializes in the analysis of legislation and the legislative process. It focuses on legislative reform and on organizational and procedural factors that affect the efficiency and effectiveness of legislative decision-making. The JOURNAL is especially interested in publishing articles that examine public policy problems of nationwide significance and propose legislation to resolve them.

The JOURNAL welcomes unsolicited manuscripts. Authors interested in being published in the JOURNAL should send their manuscripts and all related correspondence to: Articles Editor, HARVARD JOURNAL ON LEGISLATION, at the above address. Manuscripts should be typed, double-spaced, and submitted in duplicate.

Views expressed in the JOURNAL are those of the authors and do not necessarily reflect the views of the JOURNAL's officers, editors, or other members.

Subscriptions

The JOURNAL's subscription rate for the 1995 volume (Volume 32) is $24.00 domestic rate; $28.00 foreign surface rate ($30.00 in Canada); $42.00 foreign air mail. All correspondence concerning subscriptions should be addressed to the Subscriptions Manager, Publications Center, Harvard Law School, Cambridge, MA 02138.

Unless otherwise requested by the subscriber, the JOURNAL is mailed at bulk rate. For some overseas subscribers, delivery via air mail may be more reliable than standard surface rate service. Claims for missing issues of Volume 32 will be honored only until January 1996.

Individual Issues

Copies of individual issues from the current volume can be obtained directly from the JOURNAL for $12.00 each (plus $2.00 per copy for foreign mailing).

Copies of issues from all past volumes of the JOURNAL are available from Fred B. Rothman & Co., 10368 West Centennial Road, Littleton, CO 80123, (303) 979-5657. Contact Fred B. Rothman & Co. for information on prices for bound volumes.

When purchasing copies of issues from either the JOURNAL or Fred B. Rothman & Co., please send a check for the purchase price when you place your order.

Indexes

Starting with Volume 23, the JOURNAL discontinued publication of volume-by-volume and five-year cumulative indexes. The JOURNAL is indexed in the *Index to Legal Periodicals*.

In Memoriam

Erwin N. Griswold

It is with deep sadness that the *Harvard Journal on Legislation* notes the recent death of Erwin Griswold, dean of the Harvard Law School from 1946 to 1967 and solicitor general of the United States from 1967 to 1973. Dean Griswold will long be celebrated as a giant of the legal profession, for his service both to Harvard Law School and to the nation.

The *Journal* owes its very existence to Dean Griswold, who established it in 1964 as a forum for the publication of students' research on legislation and for the discussion of statutory law.[1] Even after leaving Harvard Law School, Dean Griswold continued his support of the *Journal* by writing introductory essays for the *Journal*'s anniversary issues.[2]

Dean Griswold's support of the *Journal*'s endeavors also included letters critiquing each issue, which the *Journal*'s editors have come to know as "Gris-o-grams." With his critiques of each issue, Dean Griswold provided us a great deal of grandfatherly, candid advice on all aspects of publishing an academic journal. Among his lessons were the following:

- "[Y]our "Soft Drink" Note does deal with a recent statute. I am not at all sure that I agree with [the author's] conclusion, but he does state it clearly and effectively. If I should ever have a soft drink case, which seems unlikely, I am sure that the Note would be very useful."

  —Apr. 22, 1981.

- "P.S.: I believe that my name is duplicated in your mailing list . . . . I am very glad to receive the Journal (of which I am sort of a grandfather), but a single copy is enough."

  —Apr. 22, 1981.

---

[1] Erwin N. Griswold, *Preface*, 1 HARV. J. ON LEGIS. 3 (1964).

[2] Erwin N. Griswold, *The Explosive Growth of Law Through Legislation and the Need for Legislative Scholarship*, 20 HARV. J. ON LEGIS. 267 (1983); Erwin N. Griswold, *Preface*, 30 HARV. J. ON LEGIS. 331 (1993).

- "I hope that you can get your two issues of Volume 20 out more nearly on time. Your work has been well done for the past good many years. It should be timely, too."

—Sept. 30, 1982.

- "[T]he book review is well done, though perhaps a little rough on the Chief Justice."

—Feb. 24, 1989.

- "I particularly like the student work. With a Board the size that you have, why is there not more student work available for publication?"

—Apr. 15, 1991.

- "I first read the [ ] article with much anticipation. I hoped that it might give some real insight into problems of legislative history. I found, though, that it rather quickly took off into the clouds . . . I cannot get over the feeling that it should have been edited considerably."

—Feb. 28, 1992.

The editors of the *Harvard Journal on Legislation* will miss Dean Griswold's advice and his myriad contributions to legal education. We hope that by exemplifying the virtues he taught—commitment to student writing, concision, and a passion for the form and substance of legislation—the *Journal* will continue to honor his memory.

Harvard Journal on Legislation

ARTICLE

DAUBERT'S GUIDE TO THE FEDERAL RULES OF EVIDENCE: A NOT-SO-PLAIN-MEANING JURISPRUDENCE

ANDREW E. TASLITZ*

In Daubert v. Merrell Dow Pharmaceuticals, Inc., *the Supreme Court established a new test for the admissibility of scientific evidence under the Federal Rules of Evidence. At first glance,* Daubert*'s reasoning seems a recent example of what commentators have called a plain-meaning approach to interpreting the Rules. In this Article, Professor Taslitz argues that the Court in* Daubert *and many of the cases preceding it has employed a more complex interpretive method. Taslitz critiques plain-meaning jurisprudence and argues that a flexible approach is the best means of interpreting the Rules. He concludes that the Court is moving toward such an approach.*

In a companion article to be published in the Summer 1995 issue of the Harvard Journal on Legislation, *Professor Taslitz will explore this flexible approach, defining it as "politically realistic hermeneutics." Taslitz will criticize the rationales behind plain-meaning jurisprudence and argue that a flexible, hermeneutical approach best fits the Rules of Evidence.*

In *Daubert v. Merrell Dow Pharmaceuticals, Inc.*,[1] the United States Supreme Court replaced the *Frye v. United States*[2] test, which required that novel scientific evidence be "generally accepted" before being admissible in court, with a "reliability test" embodied in the Federal Rules of Evidence. After *Daubert*, scientific evidence is admissible if it is "relevant and reliable."[3] The trial judge must be persuaded that the preliminary facts of relevancy and reliability have been demonstrated to his satisfaction by a preponderance of the evidence before the jury may hear

---

* Professor, Howard University School of Law; former Assistant District Attorney, Philadelphia, Pa.; B.A., Queens College, 1978; J.D., University of Pennsylvania Law School, 1981.

The author thanks his wife, Patricia V. Sun, Esq., and Professor David Leonard for their numerous helpful comments on earlier drafts of this Article. Appreciation also goes to the author's research assistants, Cheryl Moat, Dalhi Myers, and Michelle Gildea-Beatty for their help in completing this Article, and to the Howard University School of Law for its financial support of this project. Comments about this Article may be addressed via Internet to HJOL@HULAW1.HARVARD.EDU.

[1] 113 S. Ct. 2786 (1993).

[2] 293 F.2d 1013 (D.C. Cir. 1923).

[3] 113 S. Ct. at 2795.

scientific testimony.[4] The judge must therefore examine the science underlying the testimony and base his admissibility decision on a weighing of factors that the Court has, to date, only partly identified.[5] The meaning and wisdom of the *Daubert* decision as a vehicle for improving the judicial system's use of scientific evidence continues to be widely debated.[6] This Article, however, focuses on *Daubert* for a very different purpose: to understand and critique the Supreme Court's approach to interpreting the Federal Rules of Evidence. There has been little but growing scholarly attention paid to how the Court does and should interpret the Rules. That scholarship almost uniformly recognizes that the Supreme Court treats the Rules like any other statute, ignoring their special status as a statute drafted in the first instance by the judicial branch to govern its daily operations.[7] These scholars also argue that the Court has adopted a rigid "plain-meaning" approach to Rule interpretation, allowing exceptions only for highly unusual cases.[8] This approach has been decried as leading to an inflexible, hidebound analysis of the Rules that both prevents them from growing and changing to meet new challenges and unduly limits the discretion needed to individualize justice in particular cases.[9]

But *Daubert* is a fascinating case study that sheds new light on the approach employed by the Court to interpret the Rules. *Daubert* reveals that the Court often feels compelled to cloak much of its analysis in the language of "new textualism," a philosophy which gives virtually controlling weight to the words of the Rules, which invariably are found to be clear when they are not,[10] and claims that further analysis is logically dictated by

---

[4] *See id.* at 2796 & n.10. This Article does not discuss whether "relevance" should be judged by a more permissive standard than "reliability," a possibility that the Court simply did not address. *See infra* text accompanying notes 183–221 (discussing more permissive standard for "conditional relevancy").

[5] *See infra* text accompanying notes 245–254.

[6] *See, e.g.*, Paul C. Giannelli, *Forensic Science:* Frye, Daubert, *and the Federal Rules*, 29 CRIM. L. BULL. 428 (1993); Michael H. Graham, Daubert v. Merrell Dow Pharmaceuticals, Inc.*: No* Frye, *Now What?*, 30 CRIM. L. BULL. 153 (1994).

[7] *See infra* text accompanying notes 273–280; Glen Weissenberger, *The Supreme Court and the Interpretation of the Federal Rules of Evidence*, 53 OHIO ST. L.J. 1307, 1307 (1992) ("Most fundamentally, the Federal Rules of Evidence originated in, and were designed by, the judicial branch.").

[8] *See infra* text accompanying notes 273–280.

[9] *See id.*; Thomas Mengler, *The Theory of Discretion in the Federal Rules of Evidence*, 74 IOWA L. REV. 413, 458 (1989) ("[T]he Advisory Committee intended to give trial courts the maneuverability to do individual justice.").

[10] *Compare infra* text accompanying notes 237–244, 281–287 (noting that the *Daubert* Court purportedly based its decision on Rule 702's "clear" language) *with* James W.

those words rather than by judicial examination of legislative history, judicial discretion, and evidentiary policy. Closer examination of *Daubert* reveals, however, that behind this textualist cloak the Court makes use of a variety of interpretive techniques inconsistent with a "new-textualist" or "plain-meaning" approach to the Rules. Though the Court refuses to acknowledge it, its purported application of plain meaning to the text of the Rules involves the justices in a series of policy-driven choices.

Daubert represents only the latest high-profile example of the Court's approach to the Rules, however, and cannot be fully understood without understanding the other leading, relatively recent cases in which the Court has been called upon to interpret the Rules. Close examination of these cases reveals that they do not display the single-minded devotion to plain meaning that some scholars have claimed. It is true that in several of these cases the text has been given such primacy as to undermine the likely intent of Congress[11] and the wisdom that would be dictated by sound policy. Nevertheless, the Court generally has been reluctant to commit itself expressly to a dramatic new textualism. In a significant number of cases the Court has engaged in a close review of legislative history, common law antecedents, and evidentiary policy, a "grab-bag" approach in which the Court uses numerous databases to select a meaning that will best serve its purposes. A review of the dissents and concurrences from these opinions highlights the internal debate within the Court regarding the proper approach to the Rules.

Adding *Daubert* to the analysis sharpens the debate's complexity. The Court selectively gives primacy—but not sole control—to a plain-meaning approach, sometimes believing it wise to give lip service to such an approach even when it is obvious that the claim of plain meaning is nonsense. But the Court is also drawn by the need for an approach more sensitive to nontextual indications of congressional intent and to the broad policies sought to be served by the Rules, apparently recognizing that mechanical tests generally are chimeras and may lead to undesirable results. This conflict likely has its roots in the Court's

---

McElhaney, *Fixing the Expert Mess*, 20 LITIG. 53, 53–54 (Fall 1993) (observing that Rule 702 is an ambiguous mess, having nothing "to do with how reliable scientific evidence has to be," and offering nothing "to put in . . . [*Frye*'s] place"; the Court was therefore "in the awkward position of making something up . . . .").

[11] *See, e.g.*, the discussion of Bourjaily v. United States, 483 U.S. 171 (1987), *infra* text accompanying notes 45–65.

search for an "objective" standard to justify its decisions that will avoid the appearance of judicial policymaking.

Many flaws of the new textualism are revealed by a hermeneutical approach to statutory interpretation. While "legal hermeneutics" is a term subject to a variety of definitions,[12] here I use it to connote that meaning necessarily depends upon context, history, values, and politics; that the choice of meaning therefore involves a conversation among the relevant members of a community about what that meaning should be; and that an "objective," mechanical interpretation of statutes is therefore impossible.[13] A candid recognition of the teachings of hermeneutics can lead to more rational, effective judicial decisionmaking, bounded to some degree by language and legislative intent (albeit not strictly determined in the way the new textualism assumes).

This Article argues that the Federal Rules of Evidence are particularly appropriate for a more flexible (as opposed to plain-meaning) approach to interpretation; that *Daubert* reveals an implicit, or perhaps explicit, move by the Court toward such a recognition; and that if the Court would candidly accept and acknowledge such an approach, its decisions under the Rules would more persuasively and effectively craft a sensible law of evidence, and better guide the lower courts. This approach would also reduce the prevailing appearance of a result-oriented and inconsistent Court. *Daubert* is, therefore, not the central concern of this Article, but rather a starting point, a way to focus and sharpen our understanding of how best to interpret the Federal Rules of Evidence.

Part I of this Article reviews recent, leading Supreme Court decisions under the Rules to determine the extent to which a plain-meaning philosophy dominates the Court's approach and to examine the Court's internal conflicts concerning the wisdom of the plain-meaning approach. This review focuses on the Court's uncertainty about the wisdom of plain meaning and the many departures from such an approach. Understanding that uncertainty sheds light on what the Court did in *Daubert*.

Part II then places *Daubert* within the Court's evolving tradition of interpreting the Rules, and gauges how, if at all, *Daubert* clarifies or modifies that tradition. After carefully describing

---

[12] *See* Gregory Heyh, *Introduction* to LEGAL HERMENEUTICS: HISTORY, THEORY, AND PRACTICE xii (Gregory Heyh ed. 1992).

[13] *See id.* at xi; *infra* text accompanying notes 328–355.

Daubert, the Article critiques the Court's opinion to illustrate the numerous approaches to interpretation within the opinion and to demonstrate the conflict between what the Court says it is doing and what it is in fact doing. This critique cannot be understood, however, without first understanding the substantive and procedural concepts underlying the precise questions before the Court. Part II thus provides as background a review of the "*Frye* versus relevancy" dispute, which concerns the proper test for admitting scientific evidence and the meaning and importance of Rule 104 as a governing principle in the Court's determination of preliminary facts necessary for admissibility decisions.

Part III examines the Court's single post-*Daubert* Rules decision, *Williamson v. United States*,[14] and the judicial philosophy of the Court's newest member, Justice Stephen Breyer. Part III argues that the Court is well on its way to adopting a more flexible interpretive method than plain meaning, one that adopts a politically realistic view of how legislatures work, while recognizing the importance of what hermeneutics has to teach us about the nature of language.

A companion article to be published in the Summer 1995 edition of the *Harvard Journal on Legislation*[15] will seek to flesh out and defend the wisdom of "politically realistic hermeneutics" as a way of interpreting the Rules, critiquing public-choice theory and demonstrating that its new textualist assumptions do not hold for the Rules of Evidence. Instead, an approach based on the insights of hermeneutics may be far more consistent than plain meaning with the structure of the Rules, the intentions of the drafters, and sound evidentiary policy.

I. BEFORE *DAUBERT*: A NOT-SO-PLAIN-MEANING JURISPRUDENCE

This part reviews in chronological order the most important recent cases in the Supreme Court's statutory-evidence jurisprudence. Close examination of these cases reveals that the Court has not engaged in the kind of pure plain-meaning approach that recent scholarship has suggested.[16] To the contrary, the Court has

---

[14] 114 S. Ct. 2431 (1994).

[15] Andrew E. Taslitz, *Interpretive Method and the Federal Rules of Evidence: A Call for a Politically Realistic Hermeneutics*, 32 HARV. J. ON LEGIS. (forthcoming 1995).

[16] The organizational scheme adopted here tracks that in Randolph Jonakait, *The*

looked to common law roots, legislative history, broad congressional goals, logic, and policy analysis to resolve difficult interpretive problems.

A chronological rather than a thematic review indicates that the Court has not increasingly drifted toward a new textualism. Rather, the Court's emphasis on text, legislative history, and policy has varied with the particular interpretive problem and its consequences. A chronology also illustrates the Court's reliance on a wide variety of data and interpretive techniques (in addition to text) in a single opinion, and thus enables one to contrast concurring and dissenting opinions, which reflect differing interpretive attitudes. Most of these attitudes, however, embody the core of a more dynamic, flexible approach than pure textualism, an approach developed further in *Daubert*.

A. *Common Law Ties*

In *United States v. Abel*,[17] the Court emphasized the Rules' close ties to common law antecedents. The trial court allowed a defense witness to testify on cross-examination that both he and the defendant belonged to a secret prison gang, whose members had sworn to lie on each other's behalf. The Court of Appeals reversed the conviction, holding that the evidence was not relevant to impeach the defendant (who had not taken the stand) and thus served only to prejudice him unfairly for his association with a reviled group.[18] The Supreme Court disagreed, holding that the evidence was admissible to show the witness's pro-defense bias.

Supreme Court, Plain Meaning, and the Changed Rules of Evidence, 68 TEX. L. REV. 745 (1990) (asserting that the Supreme Court follows a "plain meaning" approach to the Rules), to emphasize the very different lessons I draw from those cases. Another leading commentator has recently described the Court's approach as a "moderate version of textualism," one in which text may yield to crystal clear legislative history. Edward J. Imwinkelried, *A Brief Defense of the Supreme Court's Approach to the Interpretation of the Federal Rules of Evidence*, 27 IND. L. REV. 267, 270 (1993). As will become clear shortly, this characterization is also flawed, especially if attention is paid to what the Court actually *does*, not only what it says it is doing. *Cf.* Glen Weissenberger, *Are the Federal Rules of Evidence a Statute?*, 55 OHIO ST. L.J. 393, 401 (1994) (urging scholars to look to "the latest behavior if not the words of the Supreme Court.").

[17] 469 U.S. 45 (1984).

[18] The defendant also argued that Rule 608(b) barred extrinsic evidence of the defendant and defense witness's joint membership in a secret gang, an argument the Court of Appeals apparently accepted. *See* United States v. Abel, 707 F.2d 1013, 1016 (9th Cir. 1983).

The Rules do not expressly state, however, that bias evidence is a permissible form of impeachment. Moreover, several Rules expressly regulate other impeachment techniques, such as character for untruthfulness,[19] conviction of a crime,[20] or certain religious beliefs.[21] Standard rules of statutory construction suggest that where a statute includes certain items, the legislature intended to exclude similar items not expressly mentioned.[22] The Rules' language thus suggest that evidence of bias is not admissible.

In reversing the Court of Appeals, the Supreme Court looked beyond the Rules' text to consider the opinion of the Advisory Committee. The Court saw its role in promulgating the Rules as "merely a conduit" between the Advisory Committee and Congress.[23] Congressional intent could therefore be ascertained in part by understanding the Advisory Committee's intent. That Committee must have known that, before the Rules were promulgated, the Court had held that trial courts must allow some cross-examination to show bias.[24] That holding was in accord with the "overwhelming weight of authority in the state courts" and was reflected in Wigmore's well-known treatise.[25] Furthermore, the Court had also held, shortly before the Rules were adopted, that the Confrontation Clause of the Sixth Amendment to the United States Constitution requires a defendant to have some opportunity to show bias on the part of a prosecution witness.[26] Given this "state of unanimity confronting the drafters" of the Rules, the Court found it unlikely that the drafters intended to scuttle cross-examination for bias entirely.[27]

Moreover, the Court buttressed its reliance on the common law, and its assertion of the absence of any express desire to

---

[19] *See* FED. R. EVID. 608.

[20] *See* FED. R. EVID. 609.

[21] *See* FED. R. EVID. 610.

[22] This is often summarized by the maxim *expressio unius est exclusio alterius*, or the expression of one item means the exclusion of another. BLACK'S LAW DICTIONARY 581 (6th ed. 1990). The parties did not articulate this argument as clearly as I have done here, *see Abel*, 469 U.S. at 49, but the *expressio unius* argument did seem to be one to which the Court was responding. The companion article will argue, however, that standard rules of statutory construction should not control the Rules because of their unique nature as a flexible code. *See* Taslitz, *supra* note 15.

[23] *Abel*, 469 U.S. at 49.

[24] *See id.* at 50 (citing Alford v. United States, 282 U.S. 687 (1931)).

[25] *Id.* at 50 (citing 3 JOHN HENRY WIGMORE, WIGMORE ON EVIDENCE § 1368 (2d ed. 1923)).

[26] *See id.* (citing Davis v. Alaska, 415 U.S. 308 (1974)).

[27] *See id.*

change the common law, by quoting Edward Cleary, reporter for the Advisory Committee:

> In principle, under the Federal Rules no common law of evidence remains. 'All relevant evidence is admissible except as otherwise provided . . . .' In reality, of course, the body of common law knowledge continues to exist, though in the somewhat altered form of a source of guidance in the exercise of delegated powers.[28]

Indeed, the Advisory Committee Notes to Rules 608 and 610, which expressly refer to the common law notion of bias impeachment,[29] seem to recognize the continuing availability of that impeachment technique under the Rules.

While the Court emphasized the Advisory Committee Notes, the opinion also relied upon the text of the Rules. Under Rule 402, all relevant evidence is admissible unless excluded by the Rules, the Constitution, or an act of Congress.[30] The *Abel* Court noted that bias is almost always relevant—that is, almost always affects the probability that a fact of consequence exists[31]—because the jury historically has been entitled to assess all evi-

---

[28] *Id.* at 51–52 (quoting Edward Cleary, *Preliminary Notes on Reading the Rules of Evidence*, 57 NEB. L. REV. 908, 915 (1978) (footnote omitted)).

[29] Rule 608 addresses when character evidence may be used to impeach or rehabilitate the credibility of a witness. FED. R. EVID. 608. Rehabilitation by certain evidence of truthful character is permitted only after a witness's character for truthfulness has been attacked by opinion, reputation, "*or otherwise.*" *Id.* (emphasis added). The Advisory Committee Note declares that "evidence of bias or interest" does not constitute an "or otherwise" attack on truthful character. FED. R. EVID. 608 advisory committee's note. Rehabilitation by evidence of truthful character is not, therefore, a permissible response to bias evidence. In recognizing that some response to bias evidence may be needed, the Note seems to assume the availability of bias impeachment under the Rules.

The Advisory Committee Note to Rule 610 similarly declared that "an inquiry for the purpose of showing interest or bias" is not within the Rule's prohibition against impeachment by evidence of religious beliefs. FED. R. EVID. 610 advisory committee's note.

[30] FED. R. EVID. 402. Professor Imwinkelried has argued that the Court's analysis of Rule 402 was the core of the opinion, while the discussion of common law history was a mere "makeweight." Imwinkelried, *supra* note 16, at 284. He argues that because the Court could have rested its decision on Rule 402, its common law discussion was not necessary and thus not important. *Id.* at 283–84. However, the Court included the common law discussion precisely because it thought the discourse sufficiently necessary and important. Indeed, *Abel* could have been decided based on common law history alone, suggesting that Rule 402 was a mere "makeweight." This is in fact a more credible reading of the opinion, for the common law discussion consumes much relative space, while a one-paragraph 402 analysis is included, apparently as an afterthought. *See* 469 U.S. at 50–52. Moreover, in *Daubert*, the Court chose to distinguish *Abel* rather than disowning *Abel*'s analysis of the common law. *See infra* text accompanying notes 304–308.

[31] Rule 401 defines "relevant evidence" as "evidence having any tendency to make the existence of any fact that is of consequence to the determination of the action more probable or less probable than it would be without the evidence." FED. R. EVID. 401.

dence which might bear on a witness's truthfulness.[32] Therefore, impeachment by bias is almost always permissible under Rule 402. In addition, Rule 607 allows any party to attack credibility[33] (bias being one such means of attack), and Rule 611(b) allows cross-examination on "matters affecting the credibility of the witness."[34] Consequently, the Court in *Abel* concluded that it was permissible to impeach a witness by bias under the Rules.[35]

B. *Plain Meaning (or Not)*

1. The *Bourjaily* Majority

While the *Abel* Court did not exclusively follow a plain-meaning approach to the Rules, the emphasis in later cases and some commentaries[36] has been on *Abel*'s suggestion that no common law of evidence remains. Such a narrowly focused approach, an overtly plain-meaning jurisprudence, was evident three years later in *Bourjaily v. United States*.[37]

Bourjaily—the case identified as giving birth to the new textualism[38]—involved the admissibility of a statement under the co-conspirators' exemption to the hearsay rule. Under that exemption, an out-of-court statement offered against a party is not hearsay if made "by a coconspirator of the party during the course of and in furtherance of the conspiracy."[39] The Court faced two questions. First, must the trial court look solely to evidence independent of the hearsay statement itself to determine that a conspiracy existed? Second, by what standard of proof must the trial judge be convinced that there was a conspiracy?[40]

---

[32] *See Abel*, 469 U.S. at 52.

[33] "The credibility of a witness may be attacked by any party, including the party calling the witness." FED. R. EVID. 607.

[34] FED. R. EVID. 611(b).

[35] The defendant had also argued that evidence of the witness's gang membership was extrinsic evidence of a specific instance of prior conduct bearing on veracity, which is barred by Rule 608(b). The Court did not reach this question, concluding that "it was enough that the evidence was admissible to show bias." *Abel*, 469 U.S. at 56.

[36] *See, e.g.*, Edward Imwinkelried, *The Meaning of Probative Value and Prejudice in Federal Rule of Evidence 403: Can Rule 403 Be Used to Resurrect the Common Law of Evidence?*, 41 VAND. L. REV. 879, 906 (1988); *see infra* text accompanying notes 233–236, 302–308.

[37] 483 U.S. 171 (1987).

[38] *See* Jonakait, *supra* note 16, at 749.

[39] FED. R. EVID. 801(d)(2)(E).

[40] *Bourjaily* also posed a third question: must the trial court examine the circum-

a. *The "bootstrapping rule."* As to the first question, the Court held that the trial judge may look to both the hearsay statement itself and independent evidence to determine whether a conspiracy existed. The Court acknowledged that two pre-Rules cases, *Glasser v. United States*[41] and *United States v. Nixon*,[42] had been interpreted to establish a "bootstrapping rule," which required trial judges to look solely to independent evidence to find the existence of a conspiracy.

The defendant argued that absent a clear expression of congressional intent to reject this common law rule, the rule should stand. However, the Court determined that Rule 104(a) expressly rejected the common law: "Preliminary questions concerning . . . the admissibility of evidence shall be determined by the court . . . . In making its determination *it is not bound by the rules of evidence except those with respect to privileges.*"[43] As a result, the Court argued, "[i]t would be extraordinary to require legislative history to confirm the plain meaning of Rule 104. The Rule on its face allows the trial judge to consider any evidence whatsoever, bound only by the rules of privilege."[44]

Despite its insistence on the overriding significance of the Rule's plain meaning, the Court relied on legislative history, specifically the Advisory Committee Note to Rule 104.[45] That Note observed that while the authorities were "scattered and inconclusive," sound sense backed the view that the judge in preliminary factfinding should be "empowered to hear any relevant evidence, such as affidavits or other reliable hearsay."[46] This language, the Court concluded, "makes plain the drafters' intent to abolish any kind of bootstrapping rule."[47]

The Court went beyond an analysis of plain meaning and legislative history to address the wisdom of its holding as a

---

stances of each case to determine whether the statement was reliable? The Court's response to this question is implicit in its response to the first two. *See Bourjaily*, 483 U.S. at 173, 179–81.

[41] 315 U.S. 60 (1942).

[42] 418 U.S. 683 (1974).

[43] FED. R. EVID. 104(a) (emphasis added).

[44] *Bourjaily*, 483 U.S. at 178.

[45] *See id.* at 179 n.2.

[46] FED. R. EVID. 104 advisory committee's note.

[47] *Bourjaily*, 483 U.S. at 179 n.2. The Court continued: "Silence [about bootstrapping] is at best ambiguous, and we decline the invitation to rely on speculation to import ambiguity into what is otherwise a clear rule." *Id.* In other words, the Rules' and Advisory Committee's silence about the long-standing common law bootstrapping doctrine do not create a gap, mistake, confusion, or ambiguity in a Rule whose language and policies suggest rejecting the common law. The Court treated silence about another long-standing common law doctrine—the *Frye* rule—similarly in *Daubert*. *See infra* text accompanying notes 302–315.

matter of policy. The defendant argued that hearsay is deemed unreliable absent proof of conspiracy, and should not, therefore, be used to support preliminary factfinding under Rule 104. The Court responded that hearsay is merely *presumed* unreliable, and that presumption is rebuttable. The Court cited Rule 803(24),[48] which permits use of otherwise inadmissible hearsay under certain conditions if circumstantial guarantees of trustworthiness are indicated.[49] Furthermore, a piece of evidence unreliable in isolation may become probative when corroborated by other evidence. Accordingly, trial courts must be free to evaluate even presumptively unreliable statements to determine whether under the particular circumstances they have significant evidentiary worth.

b. *The standard of proof.* The Court adopted a different approach to determine what standard of proof the trial court should apply to decide whether there was a conspiracy. Because the Rules do not expressly address the issue, that question could not be answered by a plain-meaning approach or an examination of legislative history. Consequently, the Court relied entirely on an analysis of evidentiary policy. The Court concluded that the preliminary fact of a conspiracy must be shown by a preponderance of the evidence.

Because preliminary facts do not decide the merits of a party's case, the evidentiary standard of proof may differ from the burden of proof at trial. The preponderance standard ensures that the trial court will find it more likely than not that the technical issues and policy concerns of the Federal Rules of Evidence have been met. Thus, the *Bourjaily* Court found "nothing to suggest that admissibility rulings in this area had been unreliable or otherwise wanting in quality because not based on some higher standard."[50] The Court determined that previous decisions in the area of constitutional law that apply a preponderance standard were controlling in this case.[51]

---

[48] *See Bourjaily*, 483 U.S. at 179.

[49] *See* FED. R. EVID. 803(24).

[50] *Bourjaily*, 483 U.S. at 176.

[51] *See id.* at 176 (citing Colorado v. Connelly, 479 U.S. 157 (1986) (holding that waiver of custodial confessant's rights must be shown by a preponderance of the evidence)); Nix v. Williams, 467 U.S. 431, 444 n.5 (1984) (a finding that illegally obtained evidence inevitably would have been discovered absent the illegality must be shown by a preponderance); United States v. Matlock, 415 U.S. 164 (1974) (stating that voluntariness of consent to search must be shown by a preponderance); Lego v. Twomey, 404 U.S. 477, 488 (1972). The *Bourjaily* Court also rejected the argument

2. Justice Blackmun's Dissent

In dissent, Justice Blackmun, joined by Justices Brennan and Marshall, criticized the simplicity of the Court's purported plain-meaning approach.[52] Justice Blackmun agreed that a Rule's plain meaning, when obvious, should not be ignored or dismissed. He preferred a structural approach, however, cautioning that the apparent plain meaning of a particular Rule should not be so readily accepted without first considering that Rule's "complex interrelations" with other Rules.[53] Justice Blackmun thus considered Rule 104(a)'s interrelation with Rule 801(d)(2)(E), the co-conspirator exemption. Moreover, he concluded that legislative history must always be examined to ensure that it confirms even the apparent plain meaning of a rule.[54] But he found in Rule 801(d)(2)(E)'s legislative history—both in the Advisory Committee Note[55] and in the legislative debates[56]—a clear expression of a congressional intent to retain the common law bootstrapping rule, requiring consideration of only independent evidence of a conspiracy's existence.[57] The apparent conflict between the express language of Rule 104(a), which allows a trial judge to consider *any* unprivileged evidence in preliminary factfinding, and Rule 801(d)(2)(E), which allows only independent evidence, could be reconciled, he concluded, by asking whether a conspir-

---

that admission of the co-conspirators' statements violated the Confrontation Clause of the United States Constitution, concluding that such statements were "firmly rooted" in evidentiary jurisprudence. *See Bourjaily*, 483 U.S. at 181–84.

[52] Justice Stevens also wrote a short concurring opinion. *See Bourjaily*, 483 U.S. at 184. Justice Stevens argued that Glasser v. United States, 315 U.S. 60 (1942), of which the Advisory Committee was surely aware, did not hold that the common law co-conspirator's exception required looking solely to independent evidence of the conspiracy. To the contrary, *Glasser* held only that a co-conspirator's statement may not be considered absent "corroborating" independent evidence of the conspiracy. 315 U.S. at 74. Such a holding, argued Justice Stevens, is consistent with the language of Rule 104(a). Therefore, it was perfectly understandable why the Advisory Committee Note would not address *Glasser*. For our purposes, whether Justice Stevens's interpretation of *Glasser* is correct is not important. What remains important is that Justice Stevens thought that legislative history must be consulted and that silences found there should be explained.

[53] *See Bourjaily*, 483 U.S. at 187 (citing Cleary, *supra* note 28, at 908).

[54] *See Bourjaily*, 483 U.S. at 187.

[55] *See id.* at 188–92 (describing Advisory Committee Note as adopting the common law rationales for the conspiracy exception without change).

[56] *See Hearings on Federal Rules of Evidence Before the Senate Comm. on the Judiciary*, 93 Cong., 2d Sess. 162 (1974) (statement of Richard H. Keatinge and John T. Blanchard).

[57] The details of his extended analysis of Rule 801(d)(2)(E)'s legislative history need not concern us here. The point is that the majority ignored the Rule's legislative history, while Blackmun considered that history central to the analysis.

acy including the declarant and defendant has been shown on the basis of evidence independent of the hearsay declarant's statements.[58]

Justice Blackmun rejected the need for a case-specific inquiry into the reliability of purported co-conspirators' statements. The traditional exemption, he thought, had been "shaped by years of 'real world' experience with the use of co-conspirator statements in trials and by the frank recognition of the possible unreliability of these statements."[59] Such statements would be, at best, the "idle chatter" of a declarant or, at worst, "malicious gossip."[60] Trial judges might nevertheless credit defendant-incriminating statements that the defendant could not explain. Independent evidence of the conspiracy was meant to provide at least some safeguard against these dangers.[61]

3. Discomfort with Plain Meaning

As Justice Blackmun recognized, the *Bourjaily* majority in part followed a pragmatic approach in which the policies underlying the Rules were relevant to answering the interpretive question before the Court. Despite its protests to the contrary, the Court was compelled to examine legislative history, albeit only that of Rule 104. Justice Blackmun demonstrated that how the Court framed the preliminary question of fact could affect the result in *Bourjaily*, even under a plain-meaning approach. The Court itself recognized that the Rules provide no answer regard-

---

[58] *See Bourjaily*, 483 U.S. at 194.

[59] *Id.* at 197. Justice Blackmun recognized that the majority had also addressed these "pragmatic or 'real world'" concerns rather than relying solely on "plain meaning." He took issue, however, with the majority's application of these concerns, which seemed to credit co-conspirators' statements with more reliability than they deserved. *Id.* at 196.

[60] *Id.*

[61] Justice Blackmun rejected the majority's conclusion that the co-conspirators' exemption satisfied the Confrontation Clause because the exemption was "firmly rooted" in practice. "Firmly rooted" exemptions or exceptions are said to be consistent with confrontation values because they have been shown to promote the accuracy of the fact-finding process. *See id.* at 200. However, the co-conspirators' exemption was never primarily based on a demand for accuracy. *See id.* at 200–01. Furthermore, the elimination of the independent-evidence requirement undermines that demand. *See id.* Finally, removal of the independent-evidence requirement undermines the exemption's "firmly rooted" status. *See id.* at 201.

Justice Blackmum was "heartened" to see that the Court reserved the question whether a co-conspirator's statement alone could establish both a conspiracy and the defendant's participation in it because he believed that this could never be the case. *See id.* at 198.

ing the standard of proof. Rather than throwing up their hands, however, the justices crafted their own answer.

While the Court ultimately did place great weight on Rule 104's plain meaning, it was uncomfortable with doing so and sought solace in cramped and inconsistent readings of policy and legislative history.[62] What might explain this approach? One possible answer, supported by empirical research and the results of much of the Court's Federal Rules of Evidence jurisprudence, is that plain meaning helped the prosecution in *Bourjaily*,[63] while common law antecedents achieved that same goal in *Abel*.[64] If this explanation is accurate, the Court's discomfort with plain meaning would come from its recognition that a consistent plain-meaning approach might sometimes undermine a pro-law-enforcement agenda. Of course, the Court also recognized that a plain-meaning approach will not always work where, for instance, the Rules are silent or ambiguous.

We can make several observations from this explanation. One, plain meaning will be emphasized in criminal cases only where the result favors the prosecution. Two, the Court will resist resting a Rules decision solely on plain meaning.[65] Three, to appear consistent in the plain-meaning approach, the Court will increasingly search for plain meaning as a ground for decision, even where none exists. Four, to avoid incoherence, the Court will also rely on other more defensible strategies, particularly where it is obvious that plain meaning does not work. The cases examined in the remainder of this section confirm each of these observations.

---

[62] The Court's reading of policy was "cramped" because it looked to Rule 803(24) (the residual exception) yet ignored 801(d)(2)(E) when only the latter was in issue. Similarly, its review of legislative history was "cramped" because it examined that of Rule 104 while ignoring 801(d)(2)(E) when both were relevant.

[63] *See* JEFFREY A. SEGAL & HAROLD J. SPAETH, THE SUPREME COURT AND THE ATTITUDINAL MODEL 243–260 (1993) (summarizing empirical data). I am not suggesting that the justices consciously so justify their behavior, although that may be true. Rather, I am suggesting that their pro-prosecution values skew their performance. *Cf.* GIRARDEAU A. SPANN, RACE AGAINST THE COURT 5 (1993) (making similar point about Court's race-related jurisprudence).

[64] *See supra* text accompanying notes 17–35 (discussing *Abel*).

[65] Professor Jonakait has also suggested that a pro-prosecution bias may explain the Court's emphasis on plain meaning. *See* Jonakait, *supra* note 16, at 762 n.80. What Professor Jonakait did not note, however, is that the very same bias may explain the Court's resistance to a zealous and consistent plain-meaning jurisprudence.

C. *An Imagined Intent, a Useful Purpose*

1. The *Owens* Majority

In *United States v. Owens*,[66] often incorrectly derided as embodying quintessential plain-meaning jurisprudence at its most absurd, the Court in fact emphasized broad congressional goals to rewrite Rules 801(d)(1)(C) and 804(a)(3). In *Owens*, John Foster, a prison guard, was brutally beaten and hospitalized with a skull fracture. On direct examination, Foster testified that he clearly remembered identifying the defendant as the assailant from his hospital bed. However, he could not remember *why* he made that identification, admitting that he did not recall whether he ever saw his attacker's face. When asked whether he had identified the defendant partly because visitors to his hospital room suggested he do so, Foster responded that he did not recall.

On appeal from the defendant's convictions, the United States Court of Appeals for the Ninth Circuit reversed, holding that Foster's testimony was inadmissible hearsay. The Supreme Court reinstated the conviction, however, concluding that Foster's testimony was admissible under Rule 801(d)(1)(C)'s hearsay exemption for a prior statement identifying a person where the declarant testifies at trial and is "subject to cross-examination" about the statement.[67]

The Court reasoned that the "natural reading" of "subject to cross-examination" is being on the stand, under oath, and willingly responding to questions, all of which occurred in *Owens*.[68] Nothing more was required. Indeed, Professor Jonakait has argued that the *Owens* Court thus applied a plain-meaning approach to the Rule.[69]

Yet the Court did not stop with its "natural reading." Instead, it conceded that limitations on the scope of cross-examination or assertions of privilege may undermine the process such that a "*meaningful* cross-examination," which the Court saw as the

---

[66] 484 U.S. 554 (1988).

[67] *See id.* at 561, 564.

[68] *See id.* at 561. This "natural reading" was not a dictionary definition but rather the meaning "ordinarily" given the words when witnesses are cross-examined. The Court did not say who "ordinarily" assigns such a meaning. Presumably, the interpretive community consists of those who ordinarily put witnesses on the stand, i.e., trial lawyers.

[69] Jonakait, *supra* note 16, at 757.

test intended by the Rule, no longer existed.[70] That effect, the Court noted, was not produced by the witness's assertion of memory loss, which may be "the very result sought to be produced by cross-examination,"[71] and which can be "effective in destroying the force of the prior statement."[72]

Moreover, the Court noted that Congress had considered the possibility of memory loss because under Rule 804(a)(3) a declarant is "unavailable" at trial (a prerequisite for application of Rule 804's hearsay exceptions) if he testifies to a "'lack of memory of the subject matter of . . . [his] statement.'"[73] Thus, despite its awareness of forgetful witnesses, Congress chose not to exclude them from the scope of Rule 801(d)(1)(C)'s hearsay exemption.

The reasons for this congressional choice were revealed by the legislative history. The Advisory Committee believed that out-of-court identifications were more reliable than later, in-court identifications, such that use of the former was to be encouraged.[74] Similarly, the House Judiciary Committee Report noted that, "'[a]s time goes by, a witness's memory will fade and his identification will become less reliable.'"[75] Consequently, admission of out-of-court identifications made closer to the crime is both fairer to defendants and prevents "cases falling through because the witness can no longer recall the identity of the person he saw commit the crime."[76] The Senate Judiciary Committee Report came to similar conclusions.[77] From this history the Court concluded that Congress had in part directed Rule 801(d)(1)(C) to address the problem of when memory loss made it "impossible [for the witness] to provide an in-court identification or testify about details of the events underlying an earlier identification."[78] In other words, Congress created the exemption to enable wrongdoers to be identified at trial even where the eyewitness was by then forgetful—precisely the situation in *Owens*.

This analysis created an internal inconsistency in the Rules, however, as forgetful witnesses are deemed "subject to cross-ex-

---

[70] *Owens*, 484 U.S. at 562.
[71] *Id.*
[72] *Id.*
[73] *Id.* (quoting FED. R. EVID. 804(a)(3)).
[74] *See id.* (citing FED. R. EVID. 801 advisory committee's note).
[75] *Id.* (quoting H.R. REP. No. 355, 94th Cong., 1st Sess. 3 (1975)).
[76] *Id.*
[77] *See id.* (citing S. REP. No. 199, 94th Cong., 1st Sess. 2 (1975)).
[78] *Id.* at 563.

amination" under Rule 801(d)(1)(C) and simultaneously "unavailable" for cross-examination under 804(a)(3). How can a witness be both "subject to cross" and "unavailable" for cross at the same time? The Court saw this as a mere "semantic oddity," an inconsistency that could have been eliminated by renaming Rule 804, "unavailability as a witness, memory loss, and other special circumstances," instead of "unavailability as a witness."

2. Useful Make-Believe

The majority's conclusion—that memory loss of the extreme degree found in *Owens* was within Congress's express contemplation in drafting Rule 801(d)(1)(C)—was a fantasy. Congress concluded that out-of-court identifications closer to the crime were more reliable than in-court identifications. Congress also feared that by the time of trial a witness might be unable to identify his attacker in court or might misidentify the wrongdoer. That is a far cry from saying that Congress thought about the specific situation before the *Owens* Court, where the witness did not remember whether he even saw his attacker in the first place and could not vouch for the truth or accuracy of the earlier identification. Indeed, Congress's assumption that an earlier identification is more reliable than an in-court identification would not hold where the earlier identification was made by someone who had never seen his assailant and who had made the identification because of police suggestion.

The better conclusion is that Congress did not think about the matter. The inquiry into congressional intent, as revealed in statutory language and legislative reports, should be viewed as an attempt to determine what Congress would have intended had it in fact considered the issue in dispute. While such a counterfactual inquiry is make-believe, the inquiry can be useful, helping to cabin judicial discretion while paying homage to furthering legislative intent as one weighty value in statutory interpretation. Even if the Court's inquiry may not be a credible one, this useful make-believe is a more accurate way to characterize the Court's reasoning than its effort to determine actual legislative intent.

The Court did not hesitate to rename Rule 804 to explain a "semantic oddity." This rewriting seems inconsistent with a plain-meaning approach to statutory interpretation. In addition, the Court described without much explanation the intent of the Rule

as admitting out-of-court identifications where there is "meaningful" cross-examination at trial. Nowhere in the Rule or its legislative history, however, does the word "meaningful" appear. Yet adding that word makes sense. As just discussed, a witness's out-of-court identification is presumed to be *more* reliable than an in-court one. This should be apparent to a jury as well. Thus, conviction might turn primarily on the credibility of the out-of-court identification. It follows that the defendant must have some real opportunity to challenge the earlier identification to get a fair trial.[79] If Congress was behaving rationally, therefore, its reason for requiring the witness to be "subject to cross-examination" would have been to allow for a meaningful challenge of the earlier identification.[80]

The Court's error, if there was one, was in its cavalier conclusion that the witness's admission of a total memory loss offered a meaningful opportunity for challenge. The majority might have done better to engage in a more serious inquiry of the likely impact of this cross-examination on a jury, perhaps examining relevant psychological literature.[81] Nevertheless, because the Court inquired into purpose, rejected plain meaning by reading a "meaningfulness" requirement into the Rule, and rewrote the title to Rule 804, *Owens* is a fascinating case.[82]

---

[79] *See* United States v. Wade, 388 U.S. 218, 235 (1967):

> Insofar as the accused's conviction may rest on a courtroom identification in fact the fruit of a suspect pretrial identification which the accused is helpless to subject to effective scrutiny at trial, the accused is deprived of that right of cross-examination which is an essential safeguard to his right to confront the witnesses against him.

See also S. REP. No. 1277, 93d Cong., 2d Sess. 16 (1974) (noting witness at trial "is on the stand and can *explain* an earlier position and be cross-examined [about it] . . . .") (emphasis added).

[80] Asking what goals a rational Congress would have sought is an inquiry into "purpose," a term defined more precisely and distinguished from "intent" in the companion article. *See* Taslitz, *supra* note 15.

[81] *See* Andrew E. Taslitz, *Myself Alone: Individualizing Justice Through Psychological Character Evidence*, 52 MD. L. REV. 1, 110–13 (1993) (illustrating use of psychological data to evaluate jury impact). Justice Brennan, joined by Justice Marshall, did indeed challenge the majority's conclusion that Owens had been offered a meaningful opportunity for challenge, but did so under the rubric of the Confrontation Clause, which is not relevant to our purposes here. 484 U.S. at 564–72. *See* Andrew E. Taslitz, *Does the Cold Nose Know? The Unscientific Myth of the Dog Scent Lineup*, 42 HASTINGS L.J. 15, 117–21 (1990) [hereinafter Taslitz, *Does the Cold Nose Know?*] (discussing the *Owens* Court's confrontation clause analysis).

[82] Although the prosecution still prevailed in *Owens*, as it would have done under a pure plain-meaning approach, such an approach alone would have been a rhetorical failure, contradicting the common sense notion among lawyers that all rules have a purpose. Interestingly, *Owens* used two related but alternative approaches to plain meaning: "imaginative reconstruction" (predicting what *this* Congress would have

D. *Plain Meaning (or Not) Revisited*

Huddleston v. United States[83] also has been simplistically depicted as single-mindedly adopting the new textualism.[84] *Huddleston* decided a question under Rule 404(b), which excludes evidence of other crimes, wrongs, or acts offered to prove a character trait of a person in order to show that the person acted in conformity with that trait on the occasion in question.[85] The Rule declares, however, that prior acts "may . . . be admissible for other purposes, such as proof of motive, opportunity, intent, preparation, plan, knowledge, identity, or absence of mistake or accident."[86]

In *Huddleston*, the defendant argued that the admission of evidence of prior acts was improper under Rule 404(b) unless the trial court first made findings by a preponderance of the evidence that the acts occurred and that the defendant committed those acts.[87] The defendant argued that these preliminary questions were for the judge because prior-acts evidence has a grave potential for improper prejudice. Therefore, the jury should not be exposed to such evidence absent proof convincing the trial court that the acts occurred. The Court disagreed, concluding that it was sufficient that there was evidence under Rule 104(b) from which a reasonable jury could find the existence of these facts by a preponderance of the evidence.[88]

The Court concluded that the defendant's position was "inconsistent with the structure of the Rules of Evidence and the plain meaning of Rule 404(b)."[89] The text of Rule 404(b) contained

---

wanted had it thought about the matter before it) and "purpose" (reading in a "meaningfulness" requirement on the implicit assumption that *any* rational Congress enacting this legislation would want to do so). For a more precise definition of these terms, *see* Taslitz, *supra* note 15.

[83] 485 U.S. 681 (1988).

[84] *See* Jonakait, *supra* note 16, at 755.

[85] *See* FED. R. EVID. 404(b).

[86] *Id.*

[87] *See* FED. R. EVID. 104(a). The prosecution's burden, more precisely stated, was to prove both that prior acts of the defendant's possessing stolen property had occurred and that the defendant *knew* in those instances that the property was stolen. Thus, if on several prior occasions the defendant knowingly possessed stolen property, it was more likely that he possessed the stolen property knowingly in the case before the Court as well. The Court did not state the issue this precisely. *See Huddleston*, 485 U.S. at 684–85. This may be because the Court implicitly concluded that repeated instances of possessing stolen property raise strong suspicion that such property was possessed with knowledge that it was stolen.

[88] *Huddleston*, 485 U.S. at 690.

[89] *Id.* at 687.

no intimation that a showing of a preliminary fact was necessary, and simply declared that prior acts "may be admissible" so long as offered for a non-character purpose.[90] Lacking a requirement for a preliminary-fact showing in Rule 404, the only other strictures are the general ones on relevancy contained in Rules 401, 402, and 403. Under Rule 401, evidence of prior acts is relevant only if there is proof that the acts occurred and the defendant committed them.[91] Rule 402 provides that only relevant evidence is admissible.[92] And Rule 403 allows for the exclusion of relevant evidence if there is a substantial danger of prejudice, confusion, misleading the jury, waste of time, or needless presentation of evidence.[93]

Rule 104(b), however, addresses the actual question of relevance conditioned on fact.[94] Under this Rule, the trial judge neither weighs credibility nor makes a finding of proof of the conditional fact. Instead, his sole role is to examine all the evidence in the case to determine whether a reasonable jury could find sufficient evidence that the preliminary fact existed. In *Huddleston*, the preliminary fact to be determined was that the defendant committed the prior acts. Once that finding has been made, under Rule 403 the trial judge should admit evidence of that preliminary fact unless its probative value is substantially outweighed by prejudice.

In addition to the language and structure of the Rules, the Court also relied upon legislative history. The Advisory Committee Note to Rule 404(b) declined to offer a "mechanical solution"[95] to the problem of whether to admit prior-acts evidence under the Rule. Rather, pragmatic relevancy was the controlling principle: "[t]he determination must be made whether the danger of undue prejudice outweighs the probative value of the evidence in view of the availability of other means of proof and other factors appropriate for making decisions of this kind under Rule 403."[96] The Senate Report agreed. "It is anticipated that with

---

[90] FED. R. EVID. 404(b). This contrasts, for example, with Bourjaily v. United States, 483 U.S. 171, 175 (1987), in which the co-conspirators' exemption expressly required the following facts to be found by a preponderance of the evidence: (1) there was a conspiracy, and (2) the statement was made by a co-conspirator (3) during the course of and (4) in furtherance of that conspiracy. *See* FED. R. EVID. 801(d)(2)(E).

[91] *See* FED. R. EVID. 401 advisory committee's note.

[92] *See* FED. R. EVID. 402.

[93] *See* FED. R. EVID. 403.

[94] *See* FED. R. EVID. 104(b).

[95] Huddleston v. United States, 485 U.S. 681, 688 (1988).

[96] FED. R. EVID. 404(b) advisory committee's note.

respect to permissible uses for such evidence, the trial judge may exclude it *only* on the basis of those considerations set forth in Rule 403, i.e. prejudice, confusion or waste of time."[97]

Moreover, both the House and Senate Reports expressed a concern for avoiding undue restrictions on admission of prior-acts evidence. The House Report explicitly stated that the final version of Rule 404(b) was intended to place "greater emphasis on admissibility than did the final Court version."[98] The Senate Report similarly declared, "[T]he use of the discretionary word 'may' with respect to the admissibility of evidence of crimes, wrongs, or acts is not intended to confer any arbitrary discretion on the trial judge."[99] Therefore, the Court found that the legislative history supported the result dictated by the language and structure of the Rules: no preliminary Rule 104(a) factfinding by the trial judge is necessary.[100]

Thus, the Court once again was able to justify a pro-prosecution result partly by relying on plain meaning. Nevertheless, it also relied, as in the other leading plain-meaning cases, on legislative history and sound policy. These inquiries belie a straightforward concern with plain meaning.

E. *Conflicting Legislative History*

In *Beech Aircraft Corporation v. Rainey*,[101] the Court could have rested its decision on the plain meaning of a Federal Rule. Instead, it relied on an obviously conflicting and inconclusive legislative history and evidentiary policy.

Beech Aircraft involved two Navy pilots who, after banking their airplane sharply to avoid colliding with another airplane, lost altitude, crashed, and were killed. The pilots' spouses brought a products liability suit against the airplane's manufacturer and the company that serviced it. The plaintiffs alleged that the accident was caused by "rollback," a loss of engine power due to a defect in the aircraft's fuel control system. The defendants

---

[97] S. REP. No. 1277, 93d Cong., 2d Sess. 25 (1974) (emphasis added).

[98] H.R. REP. No. 650, 93d Cong., 1st Sess. 7 (1973).

[99] S. REP. No. 1277, *supra* note 97, at 24.

[100] One commentator has suggested that, because the *Huddleston* Court rejected the common law standard, it adopted a new textualist approach. *See* Imwinkelried, *supra* note 16, at 287. However, Imwinkelried ignores the Court's adoption of a preponderance-of-the-evidence standard—a standard consistent with the *Bourjaily* case but nowhere expressly stated in the text of the Rules. *See supra* text accompanying notes 50–51.

[101] 488 U.S. 153 (1988).

countered that the accident was due to pilot error. The key piece of evidence in dispute was the defendant's use of excerpts from an official Judge Advocate General's ("JAG") investigatory report.[102] This report concluded that the most probable cause of the accident was the pilots' failure to maintain a proper distance between the aircraft.[103]

The interpretive question before the Court was whether the JAG Report should have been admitted under the exception to the hearsay rule in Rule 803(8)(C).[104] That Rule excepts from the hearsay bar "reports . . . of public offices or agencies, setting forth . . . in civil actions . . . factual findings resulting from an investigation made pursuant to authority granted by law."[105] Such reports are not admissible, however, if the "sources of information or other circumstances indicate lack of trustworthiness."[106] The plaintiffs argued, and the Court of Appeals concluded, that the term "*factual* findings" excluded opinions like those found in the JAG Report. The Supreme Court reversed the Court of Appeals, basing its decision upon the statutory language, legislative history, and evidentiary policy.

1. The "Plain Language" and Legislative History of Rule 803(8)(C)

Concluding that the Rules were legislative enactments that required reliance on the "'traditional tools of statutory construction,'"[107] the Court began with the language of the Rule itself. The Court noted that Black's Law Dictionary defines a "finding of fact" as "'a *conclusion* by way of reasonable inference from the evidence.'"[108] The Rule itself did not except only "factual findings" but also "*reports* . . . setting forth . . . factual findings."[109] Consequently, the Rule's "plain language" made no distinction

---

[102] *See id.* at 157–58.

[103] *See id.*

[104] There was also a second interpretive question before the Court involving the "rule of completeness" in Rule 106. *See id.* at 170–74.

[105] FED. R. EVID. 803(8)(C).

[106] *Id.*

[107] *Beech Aircraft*, 488 U.S. at 163 (quoting I.N.S. v. Cardoza-Fonseca, 480 U.S. 421, 446 (1987)).

[108] *Id.* at 163–64 (quoting BLACK'S LAW DICTIONARY 569 (5th ed. 1979)) (emphasis added).

[109] FED. R. EVID. 803(8)(C).

between fact and opinion and was consistent with "factual findings," including opinions flowing from factual investigation.[110]

Because the "traditional tools of statutory construction" required inquiry beyond the Rule's language, the Court next turned to the Rule's legislative history. Unfortunately, the two houses of Congress took diametrically opposed views on the meaning of the Rule. The House Judiciary Committee Report stated that "evaluations or opinions contained in public reports shall not be admissible under this Rule."[111] The Senate Judiciary Committee emphatically disagreed, declaring that the House Judiciary Committee's Report did not reflect "an understanding of the intended operation of the rule as explained in the Advisory Committee notes."[112]

The Court had no difficulty resolving this conflict, concluding that the Senate's understanding was "more in accord with the wording of the Rule and with the comments of the Advisory Committee."[113] The Advisory Committee Note was entitled to special weight because Congress did not amend the Committee's draft in any relevant respect, which Congress presumably would have done had it rejected the Advisory Committee's interpretation.[114] But the Advisory Committee Note mentioned no dichotomy between "facts" and "opinions."[115] Indeed, in discussing the admissibility of what it called "evaluative reports," the Committee cited numerous cases and statutes, all of which involved reports that stated conclusions.[116] Consequently, the Court believed that the Committee had assumed the admissibility of such

---

[110] *See Beech Aircraft*, 488 U.S. at 169. Professor Jonakait surprisingly cites this same discussion by the Court as demonstrating that there was no plain meaning—that the Court saw the Rule as ambiguous and, therefore, the plain-meaning rule did not apply. Jonakait, *supra* note 16, at 761 n.78. To the contrary, however, the Court's emphasis on the dictionary definition of the relevant terms as including conclusions reasonably drawn from the evidence, on Rule 803(8)(C)'s failure to mention a distinction between fact and opinion, and on Rule 402's mandate admitting *all* relevant evidence not excluded by some other Rule show that the Court did seek to justify admission of the opinion in *Beech Aircraft* through a new textual analysis. *See also* 488 U.S. 152, 163–64 (1988) (*plain language* of Rule 803(8)(C) rejects the drawing of an arbitrary line between fact and opinion).

[111] H.R. REP. No. 650, *supra* note 94, at 14.

[112] S. REP. No. 1277, *supra* note 93, at 18.

[113] *Beech Aircraft*, 488 U.S. at 165.

[114] *See id.* at 165–66 n.9.

[115] *Id.* at 166.

[116] *Id.* (citing FED. R. EVID. 803(8)(C) advisory committee's note).

reports in the first instance, providing an "escape clause" where a particular report or portion thereof appeared to be untrustworthy.[117] The Court concluded that the Committee wanted to give trial courts the discretion, indeed the obligation, to exclude untrustworthy portions of evaluative reports, whether facts or opinions.[118]

The Court also believed that the Committee's apparent reliance on relevant scholarly commentary was important. The focus of that commentary prior to passage of the Rule was on whether official reports could be admitted "*in view of the fact that they contained the investigator's conclusions.*"[119] Indeed, the Court was especially persuaded by the Committee's express reliance on an article by Professor McCormick taking the position that "[e]valuative reports . . . *though embracing conclusions*, are admissible as evidence of the facts reported."[120]

2. Policies Underlying Evaluative Reports

The structure of the Rules also suggested to the Court a policy justification for its decision. Rule 704 generally permits experts to testify to "ultimate issues,"[121] and Rule 701 permits lay witnesses to offer opinions based on observations when such testimony would be helpful to the factfinder,[122] demonstrating the Rules' "general approach of relaxing the traditional barriers to 'opinion' testimony."[123] Moreover, Rules requiring relevancy and permitting exclusion of unduly prejudicial evidence provided further tools for scrutinizing suspect reports.[124]

---

[117] *See id.* (quoting Senate Committee's explanation of FED. R. EVID. 803(8)(C) advisory committee's note).

[118] *See id.* at 167. The Court emphasized as well that the Committee had devised four non-exclusive factors to guide trial courts: "(1) the timeliness of the investigation; (2) the investigator's skill or experience; (3) whether a hearing was held; and (4) possible bias when reports are prepared with a view to possible litigation." *Id.* at 167 n.11. Although the Court spoke of both the trial court's discretion and obligation, its listing of flexible, non-exclusive factors, not one of which is determinative, is more consistent with a discretionary approach, as is the general structure of the Rules. *See* Taslitz, *supra* note 15 (discussing role of judicial discretion under the Rules).

[119] *Beech Aircraft*, 488 U.S. at 166 n.10.

[120] *Id.* (quoting Charles T. McCormick, *Can the Courts Make Wider Use of Reports of Official Investigations?*, 42 IOWA L. REV. 363, 365 (1957) (emphasis added)).

[121] FED. R. EVID. 704(a).

[122] FED. R. EVID. 701.

[123] *Beech Aircraft*, 488 U.S. at 169. This conclusion played an important role in the *Daubert* case as well. *See infra* text accompanying notes 243–244, 262.

[124] *See Beech Aircraft*, 488 U.S. at 167–68.

There were other policy concerns, ones not expressly derived from the structure of the Rules, that the Court also found relevant. Notably, the line between fact and opinion was a false and difficult one for the Court to draw. The Court viewed the requirement that witnesses give statements of fact rather than opinion as a "best evidence" rule,[125] requiring more specific ("She was uncontrollably screaming and shaking") over more general ("She was in pain") testimony. Furthermore, reports containing conclusions were subject to the "ultimate safeguard"[126]—the opponent's power to present contradictory or impeaching evidence.

Contrary to the suggestion of at least one commentator, the Court could have rested its decision entirely on the plain meaning of the statute.[127] It chose instead to apply a flexible approach, one guided by congressional intent and notions of good evidentiary policy, recognizing that many of the Rules, such as Rule 803(8)(C), were intended to establish a system of guided discretion for trial judges. This is a far cry from a plain-meaning or primarily text-based approach to interpreting the Rules.

F. *Absurd Results and Legislative Compromise*

1. The *Green* Majority

In *Green v. Bock Laundry Machine Co.*,[128] the Court went beyond the flexible but textually rooted approach of *Beech Aircraft* and expressly *rejected* the plain meaning of Rule 609(a) in favor of a meaning more consistent with legislative history and logic. The version of Rule 609 then in effect stated that evidence of a witness's criminal conviction would be admitted to attack credibility only if the crime was "punishable by death or imprisonment in excess of one year . . . and . . . the probative value [of the evidence] . . . outweighs its prejudicial effect *to the defendant* . . . ."[129] The question in *Green* was whether Rule 609 requires a judge to permit the defense to impeach a civil *plaintiff* by evidence of his prior felony convictions without first considering prejudice *to that plaintiff*. The Court concluded that the

---

[125] *Id.* at 169.
[126] *Id.* at 168.
[127] *See* Jonakait, *supra* note 16, at 7; *see also supra* text accompanying notes 107–110.
[128] 490 U.S. 504 (1989).
[129] *Id.* at 509 (emphasis added).

Rule's plain language commanded a weighing of prejudice only to a defendant, whether in a civil or criminal trial.

This literal reading compelled an odd result: civil plaintiffs and civil defendants would be treated differently without justification, likely violating fifth amendment due process. Therefore, the Rule could not mean what it said. Consequently, the Court embarked on a lengthy examination of the legislative history of Rule 609. That history helped to reveal congressional intent.

As the Court explained, the Advisory Committee's first draft of the Rule admitted all *crimen falsi* and felony convictions, eschewing the *Luck* doctrine, which required weighing prejudice against probative value.[130] But the Committee's second draft reversed direction, requiring such a weighing for both felonies and *crimen falsi* convictions.[131] The Committee's primary concern was the prejudice to the witness-accused in a criminal case, for the risk of prejudice was thought minimal for other witnesses.[132]

Senator McClellan and the Justice Department objected to the proposed adoption of the weighing doctrine; therefore, the Advisory Committee again proposed automatic admissibility in its third and final draft.[133] A House Subcommittee of the Judiciary Committee recommended a compromise alternative similar to what Congress finally adopted: automatic admissibility for *crimen falsi* convictions, and a balancing test for other convictions.[134]

The full House Judiciary Committee, on the other hand, proposed impeachment solely by *crimen falsi* convictions without balancing, expressing concern about the deterrent effect of wider admissibility of other convictions upon a criminal defendant who may want to testify.[135] That proposal was adopted by the full House.[136]

---

[130] *See id.* at 514–15. Luck v. United States, 348 F.2d 763 (D.C. Cir. 1965), a pre-Rules case, permitted the introduction of prior convictions of the defendant-witness at the discretion of the judge, who was to weigh factors like these: "(1) the nature of the crime, (2) recency of the prior conviction, (3) similarity between the crime for which there was a prior conviction and the crime charged, (4) the importance of the defendant's testimony, and (5) the centrality of the credibility issue." EDWARD W. CLEARY, ET AL., MCCORMICK ON EVIDENCE § 43, at p. 94 n.9 (3d ed. 1984) (summarizing later cases developing the *Luck* doctrine).

[131] *See id.* at 515 (citing Revised Draft of Proposed Rules of Evidence, 51 F.R.D. 315, 391 (1971)).

[132] *See Green*, 490 U.S. at 515; 51 F.R.D. at 392.

[133] *See Green*, 490 U.S. at 516–17.

[134] *Id.* at 517 (citing H.R. REP. No. 650, *supra* note 98, at 11).

[135] *Id.* at 517–18 (citing H.R. REP. No. 650, *supra* note 98, at 11).

[136] *Id.*

The Senate Judiciary Committee proposed an alternative, intermediate path. For criminal defendants impeachment could only be by *crimen falsi*, but for other witnesses felonies could be used if they passed the balancing test.[137] Nevertheless, the full Senate (prodded by Senator McClellan) reverted to the automatic admissibility of felonies and *crimen falsi* convictions, as the Advisory Committee had recommended in its first and final proposals.[138] The conflict between the House bill, allowing impeachment by *crimen falsi* convictions only, and the Senate bill, embodying automatic admissibility of both felonies and *crimen falsi* convictions, was resolved by the Conference Committee. That Committee's recommendation, ultimately enacted by the full Congress, required balancing for only non-*crimen falsi* felonies.[139]

The Conference Committee's only stated concern was the danger of prejudice to the *criminal* defendant.[140] In the Court's view, the omission of civil litigants was not an oversight. First, the weight of authority preceding the Rules admitted all felonies *without* judicial balancing in both civil and criminal cases. Thus, a party contending that legislative action changed that law had the burden of proof. Second, various drafts of Rule 609 distinguished civil and criminal cases to mitigate prejudice against the *criminal* defendant only. Third, had the conferees meant to protect other witnesses, they could have done so easily. Therefore, the Court concluded that Rule 609's balancing test was not intended to protect civil litigants.

Nor could Rule 403's balancing be permitted because the structure of the Rules demonstrated that Rule 403 was not meant to control. Rule 609 contains weighing language in subsections (a)(1) (protecting criminal defendants against prejudice), (b) (barring use of older convictions), and (d) (pertaining to juvenile convictions). The latter two provisions applied in both civil and criminal cases. The conspicuous absence of balancing in *crimen falsi* convictions and felony convictions of witnesses other than the criminal defendant indicated a legislative intent under Rule 609 to prevent balancing in those cases. Application of Rule 403 would override that intent.

---

[137] *Id.* at 519 (citing S. REP. No. 1277, *supra* note 97, at 14).
[138] *Id.* (citing 120 CONG. REC. 37076, 37083 (1974)).
[139] *Id.* at 519–20 (citing 120 CONG. REC. 40894 (1974)).
[140] *See id.*

2. Justice Scalia's Concurrence

Justice Scalia, in a concurring opinion, agreed that legislative history should be consulted. He thought that the Court should do so primarily to confirm that an absurd and "perhaps unconstitutional" result—the seemingly unthinkable proposition that Rule 609(a)(1) protected civil defendants but not civil plaintiffs—was indeed not considered, thus justifying a departure from the Rule's plain meaning. However, he would have ignored much of the Court's lengthy exploration of legislative history. Scalia was especially troubled by references to committee reports and floor debates, which reflect the views of only a handful of the members of Congress. The meaning of a statute's terms, he argued, should be judged on (1) what is most consistent with context and ordinary usage and thus likely to have been understood by the whole Congress, and (2) what is most compatible with the surrounding body of law into which the provision must be integrated, a compatibility that, by a "benign fiction,"[141] the Court assumes Congress intends. These guidelines are the clearest statement by any justice of a strict plain-meaning approach to the Rules. The majority's lengthy examination of Rule 609's "ideological evolution"[142] is a clear rejection of that approach.

3. Justice Blackmun Dissents Again

Justice Blackmun, joined by Justices Brennan and Marshall, took a different view in dissent. He saw the Conference Committee's Report as the *only* relevant legislative history because its recommendations were enacted. However, he thought the Report was as poorly drafted as the Rule. Consequently, he chose to focus on what he believed to be the Report's "underlying reasoning."[143] That reasoning distinguished between prejudice to witnesses (e.g., damage to their reputations), which was unimportant, and prejudice "improperly influencing the outcome of the trial" (i.e., prejudice to a party). The latter prejudice could affect *any* party, whether in a civil or a criminal case. Justice Blackmun admitted that this interpretation, which would extend balancing to include prejudice to civil plaintiffs, did more vio-

---

[141] *Id.* at 528 (Scalia, J., concurring in judgment).
[142] *Id.* at 529.
[143] *Id.* at 531 (Blackmun, J., dissenting).

lence to the text than did the majority's approach. But the majority's view did more violence to the only logical congressionally articulated rationale for the Rule.

Justice Blackmun conceded that his suggested result was not compelled by either the statutory language or the legislative history, but he saw his result as a "sensible" one, dictated by Rule 102's mandate to construe the Rules to "secure fairness in administration."[144] Furthermore, the Rules applied in both civil and criminal cases, and the dangers of prejudice to a party from prior convictions were present in both such cases.[145]

Justice Blackmun's approach, limiting the Court's examination of legislative history to the Conference Committee's Report, is seriously flawed. First, a coherent notion of legislative intent turns in part on identifying, where reasonably possible, legislative deals and political compromises.[146] Compromises can only be revealed, however, by understanding the competing proposals and the reasons for disagreement.[147] Second, a sensible approach to statutory interpretation seeks to comprehend fully the problem facing the legislature, and a knowledge of intellectual history aids in better understanding that problem.

Professor Weissenberger has argued that the majority's error was not that it went too far, but rather that it did not trace back far enough the intellectual history of the Rules.[148] Specifically, he maintains that the history of the scholarly and political debates that ultimately led to the Rules shows that they were intended to operate as general guidelines for the exercise of wide discretion by the trial judge, based upon experience in the common law tradition.[149] In this sense, the Rules are not really a statute at all. In addition, the intellectual history of Rule 403 revealed that it embodied the Court's traditional common law authority to exclude evidence that might adversely affect the integrity of the factfinding process.[150] There was no indication that Congress meant to limit that inherent authority. In this view,

---

[144] *See* FED. R. EVID. 102.

[145] Justice Blackmun's judgment of good evidentiary policy was indeed vindicated when Rule 609 was amended in 1990 to provide some form of balancing for use of felony impeachment against any witness in any proceeding. *See* FED. R. EVID. 609 advisory committee's note.

[146] *See* Taslitz, *supra* note 15.

[147] *Id.*

[148] *See* Weissenberger, *supra* note 7, at 1337–38.

[149] *See id.* at 1332 (citing EDMUND MORGAN, Foreward to MODEL CODE OF EVIDENCE 12–13 (1942)).

[150] *See id.* at 1337.

Rule 403 balancing should have been available to protect the civil plaintiff in *Green*.[151]

Professor Weissenberger may have gone too far, both in his specific conclusions and in his view of the Rules not being a statute.[152] As to the latter point, Congress did carefully consider, debate, and enact the Rules like any other statute. A better way to explain Weissenberger's insight is that the Rules are a statute of a special kind, one in which Congress meant to defer to the Advisory Committee and to structure trial court discretion, guiding such discretion partly by the Rules' language and underlying policies and partly by common law antecedents.

As to Weissenberger's specific conclusions concerning Rule 609, the focus of the congressional debate was clearly on the opposing interests of the state and the defendant in criminal cases.[153] The Rule protects state interests by admitting *all* felony or *crimen falsi* convictions, and it protects criminal defendants' interests by requiring a balancing test for the *felonies*. The absence of a balancing test for *crimen falsi* convictions reflects Congress's conscious choice that in such cases state interests should predominate. So Weissenberger is wrong to suggest, as he seems to, that 403 balancing governs *crimen falsi* convictions admitted against criminal defendants.

The remainder of Weissenberger's analysis is correct. It is unlikely that Congress had any clear intent to remove 403 protections from civil parties. Absent such a clear indication of intent, retention of common law powers embodied in Rule 403 seems more consistent with Congress's overall intent for the Rules.[154]

G. *Gaps and Ambiguities*

While *Green* noted that plain meaning would sometimes lead to absurd results, *United States v. Zolin*[155] recognized that the

---

[151] *See id.* at 1337–38.

[152] *See id.* at 1309–11. Weissenberger has recently argued that legislative intent should not control and is "fictional at best," nor should judicial intent be a "talisman." Weissenberger, *supra* note 16, at 397, 400. It is unclear whether by the latter comment he means to reject any inquiry into the drafters' intent, whomever the "drafters" may be. If he does mean to reject intent, and he clearly means to reject sole reliance on text, then what should we rely on? His answer may be "policy." If so, then it is policy unconstrained by text or legislative history. This would render evidence purely a creature of common law, and the adoption of a set of rules pointless.

[153] *See supra* text accompanying notes 130–140.

[154] *See* Taslitz, *supra* note 15.

[155] 491 U.S. 554 (1989).

Rules contain gaps and ambiguities for which there is no plain meaning. In such cases, the Court must craft its own rule.

In *Zolin*, the Internal Revenue Service sought tapes under the crime-fraud exception to the attorney-client privilege. The trial court had found the tapes privileged, without reviewing them *in camera*, as the IRS had requested. Two Rules seemed to govern. First, Rule 104(a) declares that the trial court, when faced with preliminary questions of admissibility, is "not bound by the rules of evidence except those with respect to privileges."[156] Second, Rule 1101(c) provides that the "rule with respect to privileges applies at all stages of all actions, cases, and proceedings."[157] These Rules, the Court noted, arguably suggest that a court could not review attorney-client communications in determining whether they were privileged.

The Court believed this reading would lead to an "absurd result,"[158] for exceptions to the attorney-client privilege can generally be proven only by compelling disclosure of the contents of the contested communication. Barring such disclosure would render the crime-fraud exception a dead letter. The Court further found this "Draconian interpretation"[159] inconsistent with the Rule's plain language. The Rule did not say that all materials for which a *claim* of privilege is made must be excluded from consideration. Had Congress meant to require such exclusion, it could easily have said so, as does the analogue of Rule 104(a) in California, which provides that "the presiding officer may not require disclosure of information *claimed to be privileged* . . . in order to rule on the claim of privilege."[160] To read Rule 104 like the California statute, said the Court, would be "counterintuitive."[161] Consequently, since Rule 104(a) does not prohibit *in camera* review, the proper procedure was to be determined as a matter of federal common law, as provided in Rule 501.[162]

Professor Jonakait has described *Zolin* as another illustration of the Court's plain-meaning jurisprudence.[163] Although the Court

---

[156] FED. R. EVID. 104(a).
[157] FED. R. EVID. 1101(c).
[158] *Zolin*, 491 U.S. at 566.
[159] *Id.*
[160] *Id.* (quoting CAL. EVID. CODE § 915(a) (West Supp. 1989)).
[161] *Id.* at 568.
[162] The Rules do not create any privileges. Under Rule 501, claims of privilege involving federal rights in federal courts "shall be governed by the principles of the common law as they may be interpreted by the courts of the United States in the light of reason and experience." FED. R. EVID. 501.
[163] Jonakait, *supra* note 16, at 760.

indeed purported to rely on the plain language of the Rules of Evidence in its opinion,[164] a fair reading of *Zolin* leads to the opposite conclusion. The *Zolin* Court rejected reading Rules 104(a) and 1101(c) as barring *in camera* inspection because such a reading was "absurd"—bad policy that would render the crime-fraud exception meaningless. The discussion of plain meaning was indeed directed not to the Federal Rules, but to California's analogous statute. A comparison with the California statute revealed that the relevant Rules have no plain meaning; they are silent regarding what procedure to follow. Although Rule 104(a) does provide that courts determining preliminary facts, such as the existence of a privilege, are bound by the rules of privilege, how could a court determine whether any document it reviewed under 104(a) was privileged in the first place? The California statute answered this question, but the Federal Rules do not. Indeed, the Court's narrow conclusion was that Rule 104(a) did "not prohibit the *in camera* review sought by the IRS . . . ."[165] This conclusion freed the Court to craft its own rule under Rule 501.

Zolin involved the Court finding a gap or ambiguity in the Rules' language. This gap or ambiguity resulted as much from what the Rules do say as from what they do not, as well as from intuition and sound policy. Gaps and ambiguities have forced the Court to look beyond the Rules' language in much of its evidentiary jurisprudence,[166] and, as I shall argue, both in discussing *Daubert* and in the companion article, are part of the Rules' sound design.[167]

Commentators' pronouncements that the pre-*Daubert* Court adopted a rigid, plain-meaning jurisprudence are therefore wrong. *Zolin* reaffirmed the flexible interpretive path the Court had embarked on in previous cases. Admittedly, the Court has given undue weight to text and unrealistic policy analyses where nec-

---

[164] *See Zolin*, 491 U.S. at 566.

[165] *Id.* at 568.

[166] *See supra* text accompanying notes 17–35, 50–51, 70–82, and Taslitz, *supra* note 15.

[167] *See id.* There is one remaining recent Supreme Court case interpreting the Rules, United States v. Salerno, 112 S. Ct. 2503 (1992), but it merits little discussion. In *Salerno*, criminal defendants sought to admit under the "former testimony" exception, Rule 804(b)(1), the grand jury testimony of two witnesses who invoked their fifth amendment privilege at trial. The former testimony exception applies, however, only if the party against whom the testimony is now offered—the prosecution—had a "similar motive" to develop the earlier testimony as would be true at trial. FED. R. EVID.

essary to further a pro-prosecution agenda. Nevertheless, it has adhered to the flexible approach clarified by *Daubert*.

II. *DAUBERT*'S GUIDE

In *Daubert v. Merrell Dow Pharmaceuticals, Inc.*,[168] the Court purported to find plain meaning where none exists. Close examination of *Daubert* reveals that (1) textual analysis is rarely so straightforward as the term "plain meaning" suggests, a term masking important policy choices and assumptions, and (2) the Court continues to rely on a wide variety of interpretive techniques, all of which are inconsistent with an objective or mechanical plain-meaning jurisprudence. Indeed, candid acknowledgement of these techniques demonstrates the Court's implicit recognition of a more flexible, politically realistic hermeneutics, an approach to interpreting the Rules in which legislative intent and statutory text are but two factors in a more broad-ranging inquiry.

Part II carefully describes and analyzes *Daubert* in three steps: (1) reviewing the background necessary for understanding *Daubert*, specifically, the *Frye* versus pragmatic relevancy and competency versus conditional relevancy debates; (2) describing the *Daubert* majority and dissenting opinions; and (3) critiquing *Daubert* from the perspective of how the Court searches for meaning in statutory language.

A. *Background*

1. *Frye* versus Relevancy

The classic test for the admissibility of scientific evidence was stated in *Frye v. United States*:

804(b)(1). The defendants argued that "adversarial fairness" required the Court to ignore the "similar motive" requirement clearly stated in the Rule. One need not be committed to a plain-meaning jurisprudence to expect the Court's rejection of this argument. *But see* Eileen A. Scallen, *Classical Rhetoric, Practical Reasoning and the Law of Evidence* 44 AM. U. L. REV. (forthcoming May 1995) (challenging the wisdom of a plain-meaning evidence jurisprudence by exploring what the teaching of classical rhetoric brings to evidence law). Any analysis that would jettison entirely the express "similar motive" requirement in Rule 804(b)(1) gives both text and legislative intent less weight than they warrant. *See Taslitz, supra* note 15.

[168] 113 S. Ct. 2786 (1993).

> Just when a scientific principle or discovery crosses the line between the experimental and demonstrable stages is difficult to define. Somewhere in this twilight zone the evidential force of the principle must be recognized, and while courts will go a long way in admitting expert testimony deduced from a well-recognized scientific principle or discovery, the thing from which the deduction is made must be sufficiently established to have gained general acceptance in the particular field in which it belongs.[169]

a. Frye*'s flaws.* *Frye*, properly understood, provides an independent safeguard for the trustworthiness of scientific testimony, i.e., peer review by an independent group of disinterested scientists who have examined the experimental data and determined that the conclusions should indeed be accepted.[170] The problems with *Frye*, however, are myriad. *Frye* applies only to "novel," "scientific" testimony, though the courts have been unclear about what constitutes a novel technique and often have chosen whether or not to apply *Frye* without any coherent explanation of why one expert opinion is presumably "scientific" and another is not.[171] There are also problems in defining the relevant scientific community among whose experts the principle must be generally accepted, in determining whether general acceptance applies only to principles or to techniques applying those principles as well, and in deciding whether those with an obvious bias should be excluded from the relevant community.[172] Finally, critics have derided *Frye* for being too conservative, as it screens out novel but reliable techniques.[173]

b. *The relevancy alternative.* Before *Daubert*, the primary alternative to *Frye* was the "relevancy test." This test, based upon Rules 403 and 702, essentially treats scientific evidence as any other evidence, balancing probative value against the dangers of prejudice, jury confusion, and undue delay.[174] The court must, as a result, carefully examine the science itself to determine how probative the technique is, and inquire into the possibility that the jury will misuse or misunderstand the technique.[175] Numerous variations of the relevancy test have

---

[169] 293 F.2d 1013, 1014 (D.C. Cir. 1923).
[170] *See* Taslitz, *Does the Cold Nose Know?*, *supra* note 81, at 55–56.
[171] *See id.* at 54, 61.
[172] *See* Paul C. Giannelli, *The Admissibility of Novel Scientific Evidence:* Frye v. United States, *a Half-Century Later*, 80 COLUM. L. REV. 1197, 1208–15 (1980).
[173] *See id.* at 1223–24.
[174] *See* Taslitz, *Does the Cold Nose Know?*, *supra* note 81, at 84–85.
[175] *See id.*

been suggested, all of which involve a listing of factors to be considered by the trial court. One well-known version of this test outlines seven factors:

(1) the technique's general acceptance in the field;
(2) the expert's qualifications and stature;
(3) the use which has been made of the technique;
(4) the potential rate of error;
(5) the existence of specialized literature;
(6) the novelty of the invention; and
(7) the extent to which the technique relies on the subjective interpretation of the expert.[176]

The relevancy approach has been lauded because it makes room for reliable techniques that are too novel to be generally accepted and it directly addresses the real question of concern—whether we can trust this scientific technique.[177] The disadvantage is that it requires judges to understand and critique the underlying science, a task that some maintain is beyond their reach.[178] Properly understood, however, *Frye* also requires the court to examine the underlying experimental data to ensure that there is general acceptance that adequate experimental support exists for a technique. These requirements indicate that the two approaches may not be as sharply different as they are sometimes portrayed.[179]

c. *The interpretive problem.* The interpretive problem before the Court in *Daubert* was this: neither the text of the Rules nor their legislative history mentioned either *Frye* or the words "general acceptance."[180] Nor was any special test for the admissibility of scientific evidence articulated.[181] This silence was remarkable in light of *Frye*'s wide acceptance and its implicit recognition of a need to screen unreliable scientific evidence from the jury.

Moreover, the recognition of a need for any special screening test—*Frye*, relevancy, or something else—necessarily involves the Court in preliminary factfinding of either the single fact of

---

[176] JACK B. WEINSTEIN & MARGARET BERGER, WEINSTEIN'S EVIDENCE ¶ 702[03], at 702-18 to 702-19 (1987). *See also* Mark McCormick, *Scientific Evidence: Defining a New Approach to Admissibility*, 67 IOWA L. REV. 879, 911–12 (1982) (advocating 11-factor test).

[177] *See* Mark McCormick, *supra* note 176, at 911–16.

[178] *See* Giannelli, *supra* note 172, at 1208–28 (summarizing arguments for and against *Frye*).

[179] *See* Taslitz, *Does the Cold Nose Know?*, *supra* note 81, at 64.

[180] *See infra* text accompanying notes 226–229.

[181] *See id.*

"general acceptance" or the multiple facts required by a relevancy approach. Such involvement raises an additional question: what standard of proof should guide the Court in finding such facts? That question was also faced by the *Daubert* Court, albeit in a truncated fashion.[182] Addressing that question in turn requires an understanding of the differences between "competency" and "conditional relevancy," differences alluded to in this Article's earlier discussion of some of the Court's evidence cases and which now require a more detailed explanation.

2. Competency versus Conditional Relevancy

Under Rule 104 of the Federal Rules of Evidence, there are two types of preliminary-fact questions that a judge may decide—competency questions and conditional relevancy questions.[183] Competency questions are "[p]reliminary questions concerning the qualification of a person to be a witness, the existence of a privilege, or the admissibility of evidence."[184] Such questions are "determined by the court,"[185] that is, the court hears evidence on both sides, makes any necessary credibility judgments, and decides whether the preliminary facts exist.[186] The existence or non-existence of those facts then determines admissibility. Thus, if there is a dispute over whether an expert has certain qualifications, the judge will decide whom to believe. If the judge believes that the witness is indeed qualified, the expert's testimony (subject to surviving other rules of evidence) will be admitted, with the jury free to re-evaluate how impressive it believes the expert's credentials to be. If, on the other hand, the judge concludes that the expert is not sufficiently qualified, the jury will never hear the expert's testimony.

Conditional relevancy questions are those for which the "relevancy of evidence depends upon the fulfillment of a condition

---

[182] *See infra* text accompanying notes 278–298.

[183] *See* CLEARY, ET AL., *supra* note 130, at 135–39.

[184] FED. R. EVID. 104(a).

[185] *Id.*

[186] Edward J. Imwinkelried, *Judge Versus Jury: Who Should Decide Questions of Preliminary Facts Conditioning the Admissibility of Scientific Evidence?*, 25 WM. & MARY L. REV. 577, 592 (1984) [hereinafter Imwinkelried, *Judge Versus Jury*]; *see* FED. R. EVID. 104 advisory committee's note. Professor Imwinkelried's suggestion to the contrary in a later article was apparently a typographical error. *See* Edward J. Imwinkelried, *The Meaning of Probative Value and Prejudice in Federal Rule of Evidence 403: Can Rule 403 Be Used to Resurrect the Common Law of Evidence?*, 41 VAND. L. REV. 879, 887 (1988).

of fact."[187] With such questions, the judge need only be convinced that there is sufficient evidence from which a reasonable jury may find that the "condition of fact" exists.[188] Consequently, the judge will not resolve credibility disputes and, indeed, may hear only the evidence tending to show that the preliminary fact exists, because if some reasonable jury could believe the evidence on that one side the judge's task is at an end.[189]

a. *The problem.* The problem for scientific evidence is that the Rules do not clearly parse out which preliminary-fact questions (other than the qualification of an expert to be a witness, which is clearly a competency question) are competency questions and which are conditional relevancy questions.

Perhaps the most important preliminary fact for which this categorization is critical is whether a scientific principle is valid,[190] for there may very well be a dispute between experts on this central question. Should the judge or the jury make this decision? As a matter of policy, the decision certainly should be treated as one of competency to be made by the judge. This is so for a simple reason: jurors who hear evidence regarding validity and conclude that a technique is not sufficiently valid nevertheless may be unable to ignore the evidence that they have heard.[191] Thus, if a jury concludes that a technique works only forty-nine percent of the time, arguably the technique should be found invalid, yet jurors will still know that the technique works often and may be so impressed by the particular expert as to believe, at least subconsciously, that surely this is one of the times when the technique actually works. Moreover, hours will have been spent discussing the validity of the technique. Thus, jurors are more likely to have processed the information thoroughly, so that memory of that information will be stronger, and it may simply be impossible to purge the memory.[192] If our goal, therefore, is to prevent jurors from making decisions based upon

---

[187] FED. R. EVID. 104(b). Whether the concept of conditional relevance makes any sense has been challenged by several commentators. *See, e.g.*, Ronald J. Allen, *The Myth of Conditional Relevancy*, 25 LOY. L.A. L. REV. 871 (1992); Dale A. Nance, *Conditional Relevance Reinterpreted*, 70 B.U. L. REV. 447 (1990); Vaughn C. Ball, *The Myth of Conditional Relevancy*, 14 GA. L. REV. 435, 436, 438 (1980).

[188] *See* Huddleston v. United States, 485 U.S. 681 (1988).

[189] *See id.* at 690.

[190] *See infra* text accompanying notes 283–285.

[191] *See* Imwinkelried, *Judge Versus Jury*, *supra* note 186, at 604–06 (noting that sound policy, unlike the Federal Rules, requires the judge, not the jury, to decide).

[192] *Id.*

invalid scientific techniques, we should leave the question of validity to the trial judge.

b. *Imwinkelried's solution.* Professor Imwinkelried has argued, however, that under the Federal Rules of Evidence the validity question is indeed left to the jury as a conditional relevancy question.[193] He bases his argument primarily on the plain text of Rule 901, which reads in relevant part:

> (A) GENERAL PROVISION
>
> The requirement of authentication or identification as a condition precedent to admissibility is satisfied by evidence sufficient to support a finding that the matter in question is what its proponent claims.
>
> (B) ILLUSTRATIONS
>
> By way of illustration only, and not by way of limitation, the following are examples of authentication or identification conforming with the requirements of this rule: . . .
>
> (9) PROCESS OR SYSTEM.
>
> Evidence describing a process or system used to produce a result and showing that the process or system produces an accurate result.[194]

Rule 901 authentication questions are, in the view of the Advisory Committee, ones of conditional relevancy controlled by Rule 104(b).[195]

Example (9)'s extension of these conditional relevancy questions to the validity of scientific techniques is partly revealed by the Advisory Committee Note, which states that Example (9) is "designed for situations in which the accuracy of a result is dependent upon a process or system which produces it. X rays afford a familiar instance. Among more recent developments is the computer . . . ."[196] The phrase "process or system" is broad enough to encompass scientific instruments, and the examples cited in the Advisory Committee Note—X-rays and computers—

---

[193] *See id.*

[194] FED. R. EVID. 901(a), (b)(9).

[195] "This requirement of showing authenticity or identity falls in the category of relevancy dependent upon fulfillment of a condition of fact and is governed by the procedure set forth in Rule 104(b)." FED. R. EVID. 901(a) advisory committee's note.

[196] *Id.*

are clearly illustrative scientific techniques. Therefore, according to Professor Imwinkelried, the intention of the Note must have been that Example (9) govern all scientific techniques.[197]

Furthermore, Professor Imwinkelried argues that the example must be read in context. Other Rule 901 examples include conditional relevance problems, like the authentication of writings.[198] In addition, "Rule 901 requires proof that proffered evidence is 'what its proponent claims,'"[199] in keeping with the general test for the conditional relevancy of questions of authenticity.[200] Construing Example (9) in context, therefore, requires reading it as governing the preliminary conditionally relevant fact whether scientific evidence is what its proponents claim, and what they claim is that the evidence is based upon valid scientific principles.[201] In addition, Example (9) cannot be seen as setting a floor of the proof necessary to admit scientific evidence because Rule 901 clearly states that "[t]he requirement of authentication . . . as a condition precedent to admissibility *is satisfied by*" the quantum of evidence specified in the Rule.[202]

c. *Imwinkelried's critics.* There are two broad responses to Professor Imwinkelried's argument. First, there is a difference between "authentication of a process generally and a showing that a particular machine works as intended."[203] Rule 901(b) arguably speaks only to the latter question.

> The emphasis of Rule 901 is upon showing that the offered item of evidence is what it is claimed to be, i.e. that it is genuine, and the rule as applied to computer evidence seems directed more to the point that the printout is a correct reflection of what is in the machine rather than that what is in the machine is correct.[204]

The two examples given in the Advisory Committee Note further support this reasoning. Both X-rays and computers are based upon well understood and accepted scientific principles. Reli-

---

[197] *See* Imwinkelried, *Judge Versus Jury*, *supra* note 186, at 609–10.

[198] *See* FED. R. EVID. 901(b)(2),(3).

[199] Imwinkelried, *Judge Versus Jury*, *supra* note 186, at 609 (quoting FED. R. EVID. 901(a)).

[200] *See id.* at 609–10.

[201] *See id.*

[202] FED. R. EVID. 901(a) (emphasis added); *see* Imwinkelried, *Judge Versus Jury*, *supra* note 186, at 610–12.

[203] 2 STEPHEN A. SALTZBURG & MICHAEL M. MARTIN, FEDERAL RULES OF EVIDENCE MANUAL 482 (5th ed. 1990).

[204] CLEARY, ET AL., *supra* note 130, at 885 n.6.

ability questions regarding those techniques, therefore, arise only in connection with their application in a particular instance. Moreover, for the reasons noted above, novel scientific evidence raises concerns over trustworthiness and reliability. The courts have recognized that where such concerns are raised the courts may properly exercise a screening function,[205] treating preliminary-fact questions as competency questions.

Second, Rules 104 and 901 probably were not intended to change common law rules, which likely treated the validity of scientific techniques as a competency question. Rule 104 does not on its face appear to modify common law rules, which made the same distinction between questions of competency and those of conditional relevancy.[206] Rule 901, in turn, does not on its face modify Rule 104.[207] Indeed, "there is no evidence that anyone who played a major part in the drafting of [Rule 901] explicitly stated that the common law would be drastically modified and the supervising role of the Trial Judge drastically reduced."[208] Thus, the common law "chain of custody" requirement demands more than proof of mere relevancy because of fears concerning the dangers of tampering with or falsifying the identity of real evidence.[209] Tape recordings often have required more proof than mere logical relevancy, for example, that the recorder was in good working order and no one tampered with the tapes.[210]

Rule 901, however, also contains examples that imply more than mere conditional logical relevancy. Rule 901(b)(2) offers as an example of proper authentication, "[n]onexpert opinion as to the genuineness of handwriting, based upon familiarity not acquired for purposes of litigation."[211] This phrasing suggests that a non-expert who has acquired familiarity with handwriting for the purposes of litigation could not authenticate a handwriting exemplar. The likely justification for that bar is that a non-expert

---

[205] *See* United States v. Downing, 753 F.2d 1224, 1240 n.21 (3rd Cir. 1985). The *Downing* court accepted similar arguments to those made here, arguments expressly identified by the court as responsive to Professor Imwinkelried's reading of the Rules.

[206] *See* 2 SALTZBURG & MARTIN, *supra* note 203, at 480.

[207] *See id.*

[208] *Id.*

[209] *See id.*

[210] *See id.* at 479–80. *See also* United States v. McMillan, 508 F.2d 101 (8th Cir. 1974), *cert. denied*, 421 U.S. 916 (1975) (basing the decision in the common law and imposing more onerous proof requirements than are shown on the face of Rule 901); United States v. Hassell, 547 F.2d 1048 (8th Cir. 1977), *cert. denied*, 430 U.S. 919, 1054 n.12 (1977) (applying more rigorous test for authenticating a recording of a telephone conversation than would be required on the face of Rule 901).

[211] FED. R. EVID. 901(b)(2).

with an interest in litigation might be biased.[212] But if mere conditional relevancy were the test, any witness with knowledge, no matter how obtained, would be able to testify, with bias being left to affect the weight of the testimony before the jury.[213]

Similarly, Rule 901(b)(6), in relevant part, states that telephone conversations may be authenticated by evidence that a call was made to the number assigned by the telephone company to a certain person and "*circumstances*, including self-identification, show the person answering to be the one called . . . ."[214] The "circumstances" have generally been interpreted to require more than mere self-identification.[215] Self-identification would seem to be logically relevant to whether the person answering is the one whom you meant to call, but because the courts recognize how easy it would be for one person to say he is someone else, more is required.[216] Some case law has suggested that federal courts will follow pre-existing common law rules on authentication even where they require more than simple logical relevancy.[217]

The majority common law view immediately prior to adoption of the federal rules was that the question of the validity of a scientific technique was a competency question for the trial judge.[218] Courts heard evidence on both sides of the validity question, made a credibility judgment, and then ruled on the objection.[219] If the rules did not change the common law in this respect, then the question of the validity of a scientific principle should be decided in the first instance by the judge.

d. *Interpretive significance.* The competency versus conditional relevancy debate between Professor Imwinkelried and his critics has interpretive significance. Professor Imwinkelried's argument is primarily textual. Although he refers to the Advisory Committee Note to Rule 901, that reference is made primarily to show that legislative history

---

[212] *See* 2 SALTZBURG & MARTIN, *supra* note 203, at 479.

[213] *See id.*

[214] FED. R. EVID. 901(b)(6)(A) (emphasis added).

[215] *See* 2 SALTZBURG & MARTIN, *supra* note 203, at 479.

[216] *See id.*

[217] *See id.* at 480 (summarizing cases).

[218] *See* Imwinkelried, *Judge Versus Jury*, *supra* note 186, at 598.

[219] *See, e.g.*, People v. Kelly, 549 P.2d 1240 (Cal. 1976); Witherspoon v. Superior Court of County of Los Angeles, 183 Cal. Rptr. 615 (Cal. Ct. App. 1982). *But see* State v. Kersting, 623 P.2d 1095, 1099 (Or. Ct. App. 1981) (using a relevancy approach, rather than *Frye*, the court described the trial judge's only task as determining whether "a competent expert [has] testifie[d] that the scientific process in question is reliable.").

is consistent with the text, a limited and supplementary use of legislative history with which some plain-meaning theorists would agree.[220]

Professor Imwinkelried's critics also rely on text. To the extent that they do so, however, they reveal the bankruptcy of a plain-meaning approach, for where there are two reasonable ways to read language—Professor Imwinkelried's and his critics'—there is no single plain meaning. More significantly, however, his critics move beyond text to consider sound evidentiary policy and common law history, considerations clearly inconsistent with a plain-meaning approach. Therefore, by siding with the critics and treating preliminary questions concerning the validity of scientific evidence as competency questions, *Daubert* squarely rejected a plain-meaning or new textual approach to the Rules.[221]

B. Daubert *Described*

1. Background

A detailed review of the facts and procedural history of *Daubert* is not particularly important for our purposes and has been done thoroughly elsewhere.[222] In short, Joyce Daubert gave birth to a deformed child and sued Merrell Dow Pharmaceuticals, Inc., the manufacturer of Bendectin, an anti-nausea drug that Ms. Daubert took while pregnant. The trial court granted the defendant's motion for summary judgment on the ground that the Dauberts could not prove that Bendectin caused the birth defect.[223] In doing so, the trial court gave no weight to the plaintiff's experts' affidavits seeking to establish causation, concluding that those experts' opinions were not based on generally accepted scientific theories.[224] The Court of Appeals for the Ninth Circuit affirmed on similar grounds,[225] squarely questioning whether *Frye* survived the Federal Rules.

---

[220] Jonakait, *supra* note 16, at 746 ("Today, the [plain meaning] standard requires courts to enforce a statute's literal language unless the legislative history of a provision explicitly indicates that the legislators intended another meaning."). *See also* Imwinkelried, *supra* note 16, at 271 (describing Court's purportedly similar approach as "moderate textualism.").

[221] *See infra* text accompanying notes 281–301.

[222] *See* sources cited *supra* note 6.

[223] *See* Daubert v. Merrell Dow Pharmaceuticals, Inc., 727 F. Supp. 570, 572 (S.D. Cal. 1989).

[224] *See id.* at 572.

[225] *See* Daubert v. Merrell Dow Pharmaceuticals, Inc., 951 F.2d 1128, 1128, 1130 (9th Cir. 1991).

2. United States Supreme Court Opinion

The Supreme Court held that *Frye* did not survive the adoption of the Federal Rules, and that in its stead the Rules established a "relevancy and reliability" test.[226]

a. *Rejecting general acceptance.* The Court's rejection of *Frye* was straightforward and cursory. The Court reasoned that Rule 702 governed and nowhere did it mention "general acceptance." Moreover, there was no legislative history suggesting that Congress intended to incorporate general acceptance into Rule 702 or the federal rules as a whole. To the contrary, the drafting history made no mention of *Frye*. The "austere"[227] *Frye* standard, which made it difficult to admit scientific evidence, would be contrary to the "liberal thrust"[228] of the Rules, which took the "'general approach of relaxing the traditional barriers,'" like *Frye*, "to 'opinion' testimony."[229]

Furthermore, the Court held, Rule 402 provided that *all* relevant evidence was admissible unless expressly excluded by the Constitution, the Rules, or other rules or acts of Congress.[230] There was no such express exclusion for scientific evidence. Such evidence might be relevant, and thus admissible, even though not generally accepted, because of the Rules' "liberal"[231] definition of relevance as *any* tendency to make a matter of consequence more or less probable.[232]

Of course, Rule 702 might bar evidence not generally accepted if that Rule incorporated the common law *Frye* doctrine. *Frye* was, after all, part of the common law for over fifty years prior to enactment of the Rules. If Congress had meant to reject such a well-established common law doctrine, it arguably would have said so.

---

[226] *See* Daubert v. Merrell Dow Pharmaceuticals, Inc., 113 S. Ct. 2786, 2795, 2799 (1993). Unlike *Frye*, the Court's new test applies to all scientific techniques and theories, not just those that are "novel." *Id.* at 2796 n.11. The Court noted, however, that sufficiently well-established propositions are less likely to be challenged and may even be subject to judicial notice. *See id.*

[227] *Id.* at 2794.

[228] *Id.*

[229] *Id.* (quoting Beech Aircraft Corp. v. Rainey, 488 U.S. 153, 169 (1988) and citing Jack B. Weinstein, *Rule 702 of the Federal Rules of Evidence is Sound; It Should Not Be Amended*, 138 F.R.D. 631, 631 (1991) ("The Rules were designed to depend primarily upon lawyer-adversaries and sensible triers of fact to evaluate conflicts.")).

[230] *See Daubert*, 113 S. Ct. at 2793–94 (citing FED. R. EVID. 402).

[231] *Id.* at 2794.

[232] *See* FED. R. EVID. 401.

The Court acknowledged that it had found that the common law filled a gap in the Rules in *Abel*.[233] There, it should be recalled, the Court held that the Rules permitted impeachment for bias, even though bias was nowhere mentioned in the Rules.[234] The *Daubert* Court, however, distinguished *Abel* on two grounds. First, *Abel* was consistent with Rule 402 because *Abel* admitted, rather than excluded, relevant evidence. Second, as discussed earlier, it was unlikely that the Rules' drafters had intended to change *Abel*'s common law bias rule. In contrast, given the Rules' liberal treatment of expert testimony, it is likely the drafters did intend to displace *Frye*. The *Frye* question was thus more similar to *Bourjaily*,[235] where the Court "was unable to find a particular common-law doctrine in the Rules and so held it superseded."[236]

b. *A new standard.* The far more extensive portion of the Court's opinion sought to justify the new "relevancy and reliability" standard. The Court grounded this standard primarily in the language of Rule 702, which, according to the Court, "clearly contemplates some degree of regulation of the subjects and theories about which an expert may testify."[237]

Rule 702's "clear contemplation" of regulation of expert testimony was obvious, in the Court's view, from the Rule's reference to "scientific . . . knowledge."[238] "'[S]cientific' implies a grounding in the methods and procedures of science . . . . '[K]nowledge' connotes more than subjective belief or unsupported speculation."[239] The something "more" that defined knowledge was, according to Webster's, "'any body of known facts or . . . any body of ideas inferred from such facts or accepted as . . . truths

---

[233] United States v. Abel, 469 U.S. 45 (1984). *See supra* text accompanying notes 17–35 (discussing *Abel*).

[234] *See id.* The Court in *Daubert* described *Abel*, however, as concluding that the Rules "occupy the field," although the Court acknowledged that under *Abel* the common law could nevertheless sometimes serve as an aid in interpretation. *See Daubert*, 113 S. Ct. at 2794.

[235] *See* United States v. Bourjaily, 483 U.S. 171 (1987). *See supra* text accompanying notes 37–49 for a discussion of *Bourjaily*.

[236] *Daubert*, 113 S. Ct. at 2794.

[237] *Id.* at 2795. *But see* Paul C. Giannelli, Daubert*: Interpreting the Federal Rules of Evidence*, 15 CARDOZO L. REV. 1999, 2016–17 (1994) (arguing that the Court was wrong to rely on Rule 702 because the Advisory Committee Note to that rule addressed only the distinction between demonstrable expert knowledge and lay knowledge, *not* the distinction between speculative or uncertain or experimental expert knowledge—the *Frye* and *Daubert* questions—and more demonstrable expert knowledge).

[238] *Daubert*, 113 S. Ct. at 2795.

[239] *Id.*

on good grounds.'"[240] Scientists, however, viewed themselves as dealing with the "scientific method," a process, not a guarantee of certainty.[241] "Scientific knowledge" was thus derived from the scientific method and based on "good grounds." Significantly, the requirement of "good grounds" was a clear standard of evidentiary reliability.[242]

Rule 702 also requires that the evidence offered "assist the trier of fact to understand the evidence or to determine a fact in issue."[243] The Court viewed this requirement as going primarily to relevance, for evidence that is not relevant cannot be helpful to the jury. One aspect of relevance is "fit,"[244] the notion that scientific validity for one purpose is not necessarily validity for other, unrelated purposes. Thus, knowledge of the phases of the moon, noted the Court, might be relevant to how dark it was on a particular night, but not to how likely it was that a particular individual behaved irrationally.

The Court also found that Rules 701, 703, and 602 shed light on why Rule 702 embodied the relevancy and reliability requirements. Rule 701 barred lay witnesses from giving opinions unless "rationally based on the perception of the witness,"[245] a restatement of the "familiar requirement of first-hand knowledge or observation."[246] Unlike Rule 701 lay witnesses, however, experts have wide latitude to offer opinions not based on first-hand knowledge. For example, under Rule 703 experts may opine based on hearsay if it would be "reasonably relied upon by experts in the particular field."[247] Because the first-hand knowledge requirement reflects a "'pervasive manifestation' of the common law insistence upon the most reliable sources of information,"[248] the Court stated that the Rules must have eliminated

---

[240] *Id.*

[241] *Id.* Interestingly, the Court's *authority* for this last proposition was the briefs of selected *amici*. *See id.* (citing Brief for Nicolaas Bloembergen *et al.* as *Amici Curiae* at 9, *Daubert*, 113 S. Ct. 2786 (No. 92-102); Brief for American Association for the Advancement of Sciences as *Amicus Curiae* at 7–8, *Daubert*, 113 S. Ct. 2786 (No. 92-102)).

[242] *Id.* at 2795. Scientists, the Court noted, used the word "reliability" differently from the sense in which the Court used it. *Id.* at 2795 n.9. "Scientific reliability" is shown where a principle's application produces consistent results. *Id.* "Scientific validity" is shown where a principle supports what it purports to show. *Id. Evidentiary* reliability is shown by *scientific* validity. *See id.*

[243] FED. R. EVID. 702.

[244] *See Daubert*, 113 S. Ct. at 2795–96.

[245] FED. R. EVID. 701.

[246] FED. R. EVID. 701 advisory committee's note.

[247] FED. R. EVID. 703.

[248] *Daubert*, 113 S. Ct. at 2796 (quoting FED. R. EVID. 602 advisory committee's note).

the first-hand knowledge requirement for experts only because it was assumed that the knowledge and experience of their disciplines provided a reliable basis for their opinions.

c. *Guiding factors.* Having established relevancy and reliability requirements in Rule 702, the Court then sought to give trial judges guidance regarding the application of that new standard. The Court stated in *Daubert*, "[w]e are confident that federal judges possess the capacity to undertake this review."[249] The Court concluded, without citation or analysis, that relevancy and reliability were to be treated as competency questions under Rule 104(a).[250] The first factor to be considered by the trial judge, concluded the Court, was whether the theory or technique could be and had been tested. The Court derived this factor from the nature of the scientific method, which requires the generation and testing of hypotheses,[251] and cited works on the philosophy of science to support this understanding.[252]

A second factor was whether the theory or technique had been subjected to peer review and publication. The Court conceded that publication is merely one aspect of peer review and did not necessarily correlate with reliability. Indeed, the Court noted that sometimes well-grounded and innovative theories have not yet been published, and some propositions are too new or of too

---

[249] *Id.*

[250] *See id.* at 2796–97 & n.10.

[251] *See id.* at 2796.

[252] *See id.* at 2796–97 (citing Michael G. Green, *Expert Witnesses and Sufficiency of Evidence in Toxic Substances Litigation: The Legacy of Agent Orange and Bendectin Litigation*, 86 NW. U. L. REV. 643, 645 (1992) ("Scientific methodology today is based on generating hypotheses and testing them to see if they can be falsified; indeed, this methodology is what distinguishes science from other fields of human inquiry."); CARL G. HEMPEL, PHILOSOPHY OF NATURAL SCIENCE 49 (1966) ("[T]he statements constituting a scientific explanation must be capable of empirical test."); KARL R. POPPER, CONJECTURES AND REFUTATIONS: THE GROWTH OF SCIENTIFIC KNOWLEDGE 37 (5th ed. 1989) ("[T]he criterion of the scientific status of a theory is its falsifiability or refutability, or testability.")).

The Court's citation of literature on the philosophy of science was selective, for there are many philosophers of science who reject falsificationism as the grand criterion that makes science superior to other methods of inquiry. *See, e.g.*, ALAN F. CHALMERS, WHAT IS THIS THING CALLED SCIENCE?: AN ASSESSMENT OF THE NATURE AND STATUS OF SCIENCE AND ITS METHODS 66 (2d ed. 1982) ("An embarrassing historical fact for falsificationists is that if their methodology had been strictly adhered to by scientists then those theories generally regarded as being among the best examples of scientific theories would never have been developed because they would have been rejected in their infancy."); Donald Campbell, *Foreword* to ROBERT K. YIN, CASE STUDY RESEARCH: DESIGN AND METHODS 7 (rev. ed. 1984) ("More and more I have come to the conclusion that the core of the scientific method is not experimentation per se, but rather the strategy connoted by the phrase 'plausible rival hypotheses.'"); PAUL DIESING, HOW DOES SOCIAL SCIENCE WORK? REFLECTIONS ON PRACTICE (1991) (summarizing numerous theories that reject a falsificationist approach to science).

limited interest to be published. Nevertheless, the Court considered the fact of (or lack of) publication as a relevant factor in assessing scientific validity because submission to the scrutiny of the scientific community is part of good science and increases the likelihood of uncovering methodological flaws.[253] Again, to support this analysis the Court cited articles by scientists emphasizing the role of peer review and publication.[254]

This time citing both case law applying a relevancy approach to scientific evidence and a well-known evidence treatise,[255] the Court found three additional factors to be important: (1) "the known or potential rate of error";[256](2) "the existence and maintenance of standards controlling the technique's operation";[257] and (3) whether the principle or technique had "widespread acceptance," for a "known technique that has been able to attract only minimal support within the community may properly be viewed with skepticism."[258] In a footnote the Court recognized that a number of "authorities," which included a law journal article on the subject,[259] have each presented a slightly different set of factors to guide the reliability determination. The Court simply noted that all of these versions may have merit, but expressed no opinion regarding their particular details.[260]

The Court additionally emphasized that the judge's focus is to "be solely on principles and methodology, not on the conclusions that they generate."[261] The Court reminded judges to keep Rules 703, 706, and 403 in mind, in addition to Rule 702. Other than reciting the content of these Rules, the Court offered no commentary on why or how the Rules were to be used, except

---

[253] *Daubert*, 113 S. Ct. at 2796.

[254] *See id.* The Court cited SHEILA JASANOFF, THE FIFTH BRANCH: SCIENCE ADVISORS AS POLICYMAKERS 61–76 (1990); David F. Horrobin, *The Philosophical Basis of Peer Review and the Suppression of Innovation*, 263 JAMA 1438 (1990); JOHN M. ZIMAN, RELIABLE KNOWLEDGE: AN EXPLORATION OF THE GROUNDS FOR BELIEF IN SCIENCE 130–33 (1978); Arnold S. Relman & Marcia Angell, *How Good is Peer Review?*, 321 NEW ENG. J. MED. 827 (1989).

[255] *See id.*

[256] *See id.* (citing United States v. Smith, 869 F.2d 348, 353–54 (7th Cir. 1989) (addressing "error rate of spectrographic voice identification")).

[257] *See id.* (citing United States v. Williams, 583 F.2d 1194, 1198 (2d Cir. 1978) (considering as a factor in a reliability analysis a "professional organization's standards governing spectrographic analysis")).

[258] *Id.* at 2797 (quoting United States v. Downing, 753 F.2d 1224, 1238 (1985)). The Court also cited WEINSTEIN & BERGER, *supra* note 176, ¶ 702[03], at 702-41.

[259] *See Daubert*, 113 S. Ct. at 2797 n.12.

[260] *Id.*

[261] *Id.* at 2797.

for emphasizing that expert evidence can be powerful and misleading and therefore requires close judicial control.

Finally, the Court addressed two concerns about the wisdom of its new approach. First, the Court rejected fears that

> abandonment of "general acceptance" as the exclusive requirement for admission will result in a "free-for-all" in which befuddled juries are confounded by absurd and irrational pseudoscientific assertions. In this regard respondent seems to us overly pessimistic about the capabilities of the jury, and of the adversary system generally. Vigorous cross-examination, presentation of contrary evidence, and careful instruction on the burden of proof are the traditional and appropriate means of attacking shaky but admissible evidence.[262]

Second, the Court deemed irrelevant claims that a screening role for the judge would sanction scientific orthodoxy and thus would be inimical to research for the truth. The claims were irrelevant because law does not search for cosmic truth. Rather, law seeks to resolve disputes finally and quickly. Conjectures that are probably wrong are unlikely to serve these goals. While a gatekeeping role for the judge may prevent the jury from learning of some authentic insights and innovations, that is the balance struck by the Rules and the need for particularized resolution of legal disputes.

3. Chief Justice Rehnquist's Dissent

Chief Justice Rehnquist, joined by Justice Stevens, concurred in part and dissented in part. Chief Justice Rehnquist agreed that *Frye* had not survived enactment of the Rules.[263] He disagreed, however, with the Court's adoption of a "reliability and relevancy" test and the Court's "general observations" about what that test meant.[264] The Chief Justice argued that, having rejected *Frye*, the Court had no need to detail what test would take its place. Rather, the Court should have left further development of the law to future cases. More importantly for purposes of this

---

[262] *Id.* at 2798. The Court also noted that trial courts remain free to direct judgments or to grant summary judgment where a reasonable juror could not find it more likely than not that a particular position was true. These devices were to be preferred to wholesale exclusion under the general acceptance test. *Id.*

[263] *Id.* at 2799.

[264] *Id.* at 2799–800.

Article, Rehnquist challenged the majority's method of statutory interpretation.

Chief Justice Rehnquist first took issue with the Court's use of authority. Twenty-two *amicus* briefs were filed in the case, and the majority's opinion contained thirty-seven citations to those briefs and other secondary sources. Those briefs were, however, "markedly different from typical briefs,"[265] in the Chief Justice's view, because they did not deal with case law or statutory language. Instead, they dealt with definitions of scientific knowledge, scientific validity, and peer review, matters beyond the expertise of judges. While Rehnquist acknowledged that such materials may be useful or even necessary in deciding how to apply Rule 703, he argued that their unusual subject matter should have caused the Court to proceed with caution instead of deciding more than necessary.[266]

Moreover, while he had "no doubt" that Rule 702 gives judges some "gatekeeping responsibility," Rehnquist did not think that the Rule imposed on them either the obligation or the authority to become "amateur scientists."[267] He expressed dismay at what he saw as the Court's confidence in the capacities of federal judges to handle matters outside their expertise. "I defer to no one in my confidence in federal judges; but I am at a loss to know what is meant when it is said that the scientific status of a theory depends on its 'falsifiability,' and I suspect some of them will be, too."[268]

The Chief Justice further questioned whether the Rules authorized creation of the new test. While the Rules expressly referred to relevance, he noted that "there is no similar reference in the Rule[s] to 'reliability.'"[269] He noted with some sarcasm that the Court constructed an argument for the reliability requirement by parsing and defining the language "[i]f scientific, technical, or other specialized knowledge will assist the trier of fact to understand the evidence or to determine a fact in issue . . . ."[270] But, he queried, did Rule 702 actually contemplate parsing this language into numerous subspecies of expertise, or did its authors merely choose general descriptive language covering the types

---

[265] *Id.* at 2799.
[266] *Id.*
[267] *Id.* at 2800.
[268] *Id.*
[269] *Id.*
[270] *Id.*

of expert testimony that courts have customarily received?[271] If such a parsing was intended, then what is the difference between "scientific" and "technical" knowledge? In addition, are the Court's new requirements to apply to both types of knowledge as well as to "other specialized knowledge"?[272]

In short, Chief Justice Rehnquist apparently viewed the majority decision as departing from a plain-meaning approach to the Rules, from a common-sense reading of Rule 702's language, and from the dictates of wise policy and a respect for the limited abilities of generalist judges.

C. *The Search for Meaning*

The Court's modern rendition of the plain-meaning approach dictates application of a statute's literal meaning unless contrary legislative history clearly demonstrates that another meaning was intended.[273] At least one commentator has argued that the Court has followed a plain-meaning approach under the Rules, except where, as in *Green*, such an approach would lead to an absurd and probably unconstitutional result.[274] This commentator derides the plain-meaning approach because it "squelches evidence law's historic dynamism and abolishes common-law methods of resolving evidentiary disputes."[275]

Other commentators have taken a more measured view, describing the Court as having an "affinity"[276] for plain-meaning analysis, but recognizing that the approach is not "monolithic."[277] Instead, the Court sometimes relies on legislative history, structure, and the overall policies of the Rules for problems simply not addressed by the Rules or where a Rule's language is not susceptible to a plain meaning.[278] Nevertheless, these commentators worry that the Court's preference for the plain-meaning approach is disingenuous in cases which Congress purposely left

---

[271] *Id.*
[272] *Id.*
[273] Jonakait, *supra* note 16, at 746; Imwinkelried, *supra* note 16, at 271.
[274] Jonakait, *supra* note 16, at 761.
[275] *Id.* at 784.
[276] Hon. Edward R. Becker & Aviva Orenstein, *The Federal Rules of Evidence After Sixteen Years—The Effect of "Plain Meaning" Jurisprudence, the Need for an Advisory Committee on the Rules of Evidence, and Suggestions for Selective Revision of the Rules*, 60 GEO. WASH. L. REV. 857, 866 (1992).
[277] *Id.*
[278] *Id.*

a point open or vague, ignores the unstated common law assumptions that coexist with the Rules, and will therefore lead to interpretations of the Rules that defeat congressional intent.[279] Many commentators agree, however, that the Court is committed to a primarily plain-meaning approach. As a result, these commentators have explored the implications of this approach for the future growth of the law of evidence, urging congressional and advisory committee oversight as a means of avoiding these horrors.[280]

Part I of this Article challenged these conclusions. *Daubert* fuels this challenge. While *Daubert* does not follow any single, clear approach to statutory interpretation, the case does suggest that, however the Court's approach is characterized, the approach is capable of at least sometimes encouraging powerful, logical change in evidence law. Moreover, *Daubert*'s approach may suggest a growing recognition by the Court that the Rules often function purely as guidelines for the exercise of judicial discretion. That recognition plants the seeds for a very different approach to interpreting the Rules.

1. Rule 104

For example, the *Daubert* Court concluded, with virtually no discussion, that whether scientific testimony meets Rule 702's requirements is a Rule 104(a) competency question to be resolved by the trial judge.[281] This conclusion, however, constitutes a rejection of a plain-meaning approach. As discussed above, Professor Imwinkelried has articulated a persuasive primarily plain-meaning argument that the question of the validity of scientific evidence is a question of conditional relevancy, *not* competency.[282]

Indeed, in addition to the arguments raised by Professor Imwinkelried, *Daubert* itself implicitly suggested a plain-meaning argument that scientific validity is a conditional relevancy question. *Daubert* properly recognized that a scientific technique "assists" the jury only if the technique is relevant.[283] A scientific

---

[279] *Id.* at 868.

[280] *See* sources cited *supra* notes 16, 276.

[281] *See Daubert*, 113 S. Ct. at 2796 n.10.

[282] *See supra* text accompanying notes 193–202, 220–221. Professor Imwinkelried's argument is persuasive *if* one accepts a plain-meaning approach to the Rules, which this Article concludes the Court has wisely rejected.

[283] *See Daubert*, 113 S. Ct. at 2795–96.

technique that is not reliable is not relevant, as it has no tendency to make a fact of consequence more or less probable.[284] Thus, reliability is a conditional relevancy problem.[285] "Conditional relevancy" is controlled by Rule 104(b) and requires sufficient proof to enable a reasonable jury to believe by a preponderance of the evidence that the preliminary fact exists.[286] This is a very low standard of admissibility that would essentially leave the question of reliability to the jury.[287] The Court, however, expressly rejected this low standard of admissibility, and, in doing so, rejected both Professor Imwinkelried's plain-meaning arguments and those suggested by *Daubert* itself.

The Court's rejection of such arguments is further suggested by the fact that Professor Imwinkelried's plain-meaning interpretation was raised (however cursorily) in the *Daubert* briefs.[288] Those briefs attacked Imwinkelried's view primarily on grounds of policy, rather than textual meaning.[289] Indeed, the scholarly assault on Professor Imwinkelried, while addressing text, has given much greater weight to evidentiary policy and common law history, breaking sharply from a text-based approach.[290]

Of course, text-based or purportedly plain-meaning support for the Court's conclusion that validity is a competency question can be found. Thus, a scientific technique cannot "assist the jury" under Rule 702 if unfair prejudice substantially outweighs probative value.[291] Rule 702 analysis is similar to that under Rule 403:[292] when courts engage in such balancing they must first determine as preliminary facts how much unfair prejudice is likely and what amount of probative value inheres in the evidence. The latter inquiry necessarily requires a *judicial* judgment of the precise *degree* of reliability (which requires assessing

---

[284] *See* FED. R. EVID. 103; Jonakait, *supra* note 16, at 767.

[285] *See supra* text accompanying notes 193–219.

[286] *See supra* text accompanying notes 183–190.

[287] *See* Jonakait, *supra* note 16, at 767. *Cf.* Giannelli, *supra* note 237, at 1999, 2011, 2014–15 (1994) (arguing, without discussing Rule 104, that *Daubert*'s requirement of relevancy *and* reliability is a more demanding standard than the pre-*Daubert* relevancy approach, even though both approaches involve a multi-factor balancing test).

[288] Brief for a Group of American Law Professors as *Amicus Curiae* In Support of Neither Part at 22 n.16, *Daubert*, 113 S. Ct. 2786 (No. 92-102) [hereinafter Law Professors].

[289] *See id.* at 22 n.16.

[290] *See supra* text accompanying notes 203–219.

[291] Jonakait, *supra* note 16, at 768.

[292] *Id.*

evidence on *both* sides of the reliability question) since greater reliability means greater probative value.[293]

There are, however, several problems with this plain-meaning approach. First, as Professor Jonakait points out, the approach is inconsistent with *Huddleston*,[294] in which the Court held that proof of the preliminary fact whether the defendant committed prior acts under Rule 404(b) was a conditional relevancy question. One argument that the defendant raised against this conclusion was that, under Rule 403, prejudice automatically substantially outweighs probative value absent a preliminary judicial finding that the defendant committed the prior crime.[295] The Court rejected this argument, holding that "Rule 403 admits of no such gloss."[296] A case-by-case balancing approach under Rule 403, not a fixed standard for admissibility requiring a preliminary judicial finding that the prior act occurred, prevailed in *Huddleston*.

Rule 702's "assists the jury" balancing approach also arguably admits of "no such gloss" as a fixed procedural or substantive standard requiring a preliminary judicial finding of scientific validity. Indeed, the balancing invited by the vague "assists the jury" test, combined with the permissive language that a qualified expert "may" testify where helpful, suggests that an ad hoc, case-by-case approach to admissibility is required under Rule 702, similar to that required under *Huddleston*'s interpretation of Rule 403.[297] Such balancing vests substantial discretion in the trial judge.[298] If trial judges have such discretion, however, there arguably will be a range of decisions in which some courts admit a particular scientific technique and some exclude the same technique. If the trial judge is granted the extreme deference on appeal that is common for discretionary evidentiary decisions,[299] appellate courts are likely to find an abuse of discretion only

---

[293] *Id.* at 769.

[294] *Id.*

[295] Huddleston v. United States, 485 U.S. 681, 689 n.6 (1988), discussed *supra* text accompanying notes 83–100.

[296] *Id.*

[297] Jonakait, *supra* note 16, at 769. Professor Jonakait did not mention the permissive "may" language in Rule 702, but his conclusion—that the plain meaning of Rule 702 points to an ad hoc, case-by-case approach to the reliability of scientific techniques—is further supported by such language.

[298] This point will be explained further in the companion article. *See* Taslitz, *supra* note 15.

[299] David P. Leonard, *Appellate Review of Evidentiary Rulings*, 70 N.C. L. REV. 1155, 1228 (1992).

where the technique is so unreliable that the evidence is not even relevant. But if the courts are merely inquiring whether there is sufficient evidence of the preliminary fact of scientific validity to render the technique relevant, the inquiry is one of conditional relevancy, not competency.[300] The plain-meaning argument in support of the Court's conclusion that validity is a competency question thus falls of its own weight.

The point of this discussion is that the most persuasive arguments supporting the Court's conclusion regarding Rule 104 competency rely more heavily on non-text-based arguments than do counterarguments that scientific validity is a conditional relevancy question. The Court's rejection of the conditional relevancy conclusion is thus a rejection of a plain-meaning approach. Indeed, Professor Jonakait predicted prior to *Daubert*, that the Court would hold, consistent with its plain-meaning approach, that scientific validity is a conditional relevancy question.[301] That Professor Jonakait's prediction was wrong demonstrates that the Court does not have the single-minded commitment to plain meaning that he and others have suggested.

2. The Silence of *Frye* (and the Common Law)

After years of extensive debate over the wisdom of *Frye*,[302] an extended analysis of the *Frye* question by the Court might have been expected. The Court instead analyzed the question summarily, avoiding any inquiry into the evidentiary policy wisdom of *Frye*. The Court purported to rely on the plain meaning of Rule 702 and on an abbreviated structural and historical analysis of the expert evidence rules to divine a legislative intent inconsistent with *Frye*.

Nowhere do the Rules or the legislative history expressly address *Frye*. Neither declares that "general acceptance of scientific techniques shall be required before evidence of such techniques may be admitted," nor do they state the opposite. Partly because of this absence of a reference to *Frye*, the Court was

---

[300] Again, Professor Jonakait did not spell out the argument that review of Rule 702 decisions is likely to be made under an abuse of discretion standard and, therefore, reversal is likely only if the conditional relevancy test is failed. This argument, however, logically follows from his point that the ad hoc approach to Rule 702 will lead to inconsistent decisions and few meaningful precedents on the admissibility of scientific evidence. *See* Jonakait, *supra* note 16, at 769–70.

[301] *Id.* at 768–70.

[302] *See* Giannelli, *supra* note 172, at 1208–28.

persuaded that *Frye* was not part of the Rules. The Court recognized that *Frye* had a lengthy common law history before the Rules, yet it ignored its assertion in *Green* that a party claiming that a statute has changed settled law has the burden of proof.[303] Instead, by relying on mere *silence* as evidence that the settled rule was rejected, the Court shifted the burden of proof, without explaining or justifying this shift, to the party maintaining that no major change in the law had been wrought.

The Court did acknowledge that in *Abel* it had looked to the common law. The Court pointed out, however, that *Abel* involved complete silence on the question whether the use of bias survived adoption of the Rules, whereas in *Daubert* Rule 702 specifically addressed expert testimony.[304] This distinction does not work, though, for at the time of *Abel* other specific rules governed impeachment generally, and those rules were silent as to bias.[305] Similarly, in *Daubert*, a specific rule governed the admissibility of expert testimony,[306] yet, according to the Court, that Rule was silent about *Frye*. Thus, the *Abel* Court inferred from silence that no change in the law was intended, while the *Daubert* Court inferred from silence that Congress intended a change in the law. To draw different inferences from such silence is wholly unjustified.

The Court also distinguished *Abel* on the grounds that its holding in *Abel* was consistent with Rule 402's mandate that all relevant evidence be admitted absent a specific statutory or constitutional rule to the contrary.[307] Since no Rule barred bias evidence, which is always relevant, this evidence had to be admitted in *Abel*. On the other hand, the *Frye* rule would exclude relevant evidence without a clear rule prohibiting admission of the evidence. Again, this is mere sophistry since the focus in *Abel* was on whether the common law was embodied in or rejected by the Federal Rules of Evidence.[308] If the common law had excluded bias evidence, and the impeachment Rules' drafters had intended not to change the common law, then the Rules also would have excluded bias, despite Rule 402. Similarly, in *Daubert*, the ques-

---

[303] Green v. Bock Laundry Machine Co., 490 U.S. 504, 521–22 (1989). *See supra* text accompanying notes 128–140.

[304] Daubert v. Merrell Dow Pharmaceuticals, Inc., 113 S. Ct. 2786, 2794 (1993). *See supra* text accompanying notes 174–179, 227–236.

[305] *See* FED. R. EVID. 610, 611.

[306] FED. R. EVID. 702.

[307] *See Daubert*, 113 S. Ct. at 2793–94.

[308] *See supra* text accompanying notes 16–35.

tion was whether the relevant Rules adopted or rejected the common law *Frye* rule. If the drafters intended the expert evidence Rules to include *Frye*, then the Rules would bar admission of some relevant evidence, and Rule 402 would be no bar to such a holding.

On the other hand, the Court simply might have been saying that Rule 402 expresses a general preference for rules of admissibility rather than rules of exclusion. Despite its repeated protests to the contrary, however, the rule that the *Daubert* Court ultimately adopted was a rule of exclusion of unreliable scientific evidence, even where such evidence is indeed relevant. Moreover, the new rule created the possibility that well-accepted techniques would nevertheless be rejected because they were insufficiently reliable. A simple-minded preference for inclusion over exclusion simply does not end the analysis.

The Court, again in a cursory fashion, also seemed to use statutory language in a structural fashion when it examined the relationship among various rules to divine an overall purpose. Thus, the Court described the expert evidence rules as having a "liberal thrust" and a "general approach of relaxing the traditional barriers to 'opinion' testimony."[309] Presumably, the Court referred to the abandonment of the "beyond the ken" of laymen test, the elimination of the requirement of a hypothetical question for expert opinions not based entirely on first-hand knowledge, the express allowance of opinions based on hearsay, and the general end to the bar on expression of opinions on ultimate issues.[310] Knowledge of the common law enables one to understand that the Rules changed what came before. The Court's argument was inductive: if some changes were made to liberalize admission of expert evidence compared to the common law, then the Rule's drafters must have intended to liberalize all common law expert evidence rules, even those on which the Rules are silent. Therefore, the "austere standard" of *Frye* must have been intended to be rejected by the Rules.[311]

---

[309] *Daubert*, 113 S. Ct. at 2794.

[310] PAUL R. RICE, EVIDENCE: COMMON LAW AND FEDERAL RULES OF EVIDENCE 1031–32, 1050–51, 1063–67, 1073–74, 1105–09, 1123–24 (2d ed. 1990). Rule 704(b) still prohibits opinions as to the ultimate issue whether a criminal had a mental state or condition constituting an element of the crime or a defense. FED. R. EVID. 704(b) advisory committee's note.

[311] *Daubert*, 113 S. Ct. at 2794.

The Court's analysis of the *Frye* question reflects an affinity for something close to plain meaning in the sense that the Court's emphasis was on the language of the Rules. The Court made some brief reference to common law standards, however, as necessary to understand the meaning of certain language, since the expert evidence rules often abandoned common law exclusionary rules. Nevertheless, it is curious that the Court relied so heavily on silence. No affirmative analysis was made of the actual language of the expert evidence rules to examine whether they could be read to adopt the *Frye* standard. A cursory analysis of the words used was made to demonstrate changes from the common law on issues other than *Frye*,[312] but no analysis was made of the language relevant to *Frye* itself. This is especially curious because, having "established" that *Frye* did not survive adoption of the Rules, the Court then analyzed precisely the language that would have been relevant to whether *Frye* survived as a way of demonstrating the existence of an alternative test.[313]

This is significant for several reasons. First, silence could have been used to support a conclusion contrary to that reached by the Court. The Rules expressly include language and legislative history to emphasize the rejection of certain common law rules. For example, with one exception, the Rules expressly declare that an expert opinion is not barred simply because it concerns an ultimate issue.[314] Indeed, as just discussed, the Rules expressly reject numerous common law hurdles to the admissibility of expert testimony. The drafters were thus obviously aware of and consciously considering whether to continue or modify the common law rules concerning expert testimony. Therefore, if they expressly chose to change four common law rules, for example, but they did not expressly modify the *Frye* rule, it may logically be argued that *Frye* was one common law rule that the drafters did not want to change. Had they wanted to do so, they would have said so, as they did for numerous other common law expert evidentiary doctrines. Likewise, since the *Frye* rule was not to be changed, there was no reason to mention that rule in legislative history. Indeed, reasoning of this kind likely supported the Court's assertion in *Green* that one contending that a well-settled rule has been changed by a statute bears the burden

---

[312] *See id.*
[313] *See id.* at 2795–98.
[314] FED. R. EVID. 704.

of proof.[315] The Rules' silence as to *Frye* thus undermines the Court's analysis as much as it supports it.

Second, an argument that the language of the Rules, if not necessarily their plain meaning, embodies *Frye* was indeed crafted by the parties, yet ignored by the Court. The parties making these arguments tracked the type of analysis used in the "relevancy and reliability" portion of the *Daubert* opinion defining "scientific knowledge." One party, citing a well-known work on the philosophy of science, argued that the term "'scientific knowledge' naturally refers to grounds that are deemed good by the relevant scientific . . . community, that is, to claims that are validated, or derived, according to the accepted standards in the relevant field."[316] Moreover, since "knowledge" is obviously more than mere personal opinion, the need for referral to the opinions of a relevant community is again suggested.[317]

Additionally, methods other than witness testimony, such as judicial notice and reference to learned treatises, are sometimes available for proving "scientific knowledge."[318] Such methods are available only when there is ample proof of validation by other communities. Thus, judicial notice applies where a principle is so well accepted that it is not even subject to reasonable dispute.[319] Learned treatises may be used as substantive evidence only if they are "established as a reliable authority,"[320] arguably requiring recognition by the scientific community.[321] Expert testimony about recognition is thus needed where such recognition cannot easily be proved by other means.

Furthermore, Rule 703 permits experts to testify about matters not within their personal knowledge and based upon hearsay if such bases are "reasonably relied upon" by experts in the field, a phrase which the Advisory Committee Note describes as intending to "bring the judicial practice into line with the practice

---

[315] *See supra* text accompanying note 140.

[316] Brief for Respondent at 14, *Daubert*, 113 S. Ct. 2786 (No. 92-102) (citing J. KOURANY, SCIENTIFIC KNOWLEDGE: BASIC ISSUES IN THE PHILOSOPHY OF SCIENCE 112 (1987) ("[H]ypotheses 'must prove their mettle' to become part of 'scientific knowledge[.]'")). Indeed, it was in this brief that a party first argued that the word "knowledge," in isolation, in ordinary usage means "good grounds" for the belief that something is known. This was the precise definition used by the *Daubert* Court, but to reach a very different conclusion.

[317] *See id.* at 14–16.

[318] *Id.* at 16 (citing FED. R. EVID. 201, 803(18)).

[319] *See id.*

[320] FED. R. EVID. 803(18).

[321] *See* Brief for Respondent, *supra* note 316, at 16.

of the experts themselves when not in court."[322] Again, this comment arguably suggests general acceptance by the relevant community of experts, and it would be odd indeed to require such acceptance only for some opinions but not for others.

Similarly, the "will assist" the jury requirement of Rule 702 and the weighing of prejudice against probative value in Rule 403 obviously require some standard for ensuring that scientific evidence will help the jury more than harm it, and given the "aura of infallibility"[323] that may surround scientific evidence, *Frye* would provide a significant guarantee that only reliable evidence comes before the jury.[324] As argued above in reference to the Court's treatment of silence, the *Frye* test was an accepted test for ensuring such reliability when the Rules were adopted. Therefore, it is equally plausible that *Frye* was intended to meet this requirement of reliability embodied in Rules 702 and 403 as it is that some other test applied.

The point of this discussion is not to suggest that Merrell Dow's arguments regarding the most "natural" reading of Rule 702's language are more convincing. Rather, the point is that there were reasonable arguments that the language was consistent with and indeed embodied *Frye*—arguments supplemented further by policy analyses and legislative history.[325] Leading commentators in the evidence field have found many of these arguments convincing.[326]

Rather than rebutting such arguments, the Court chose to focus on what the Rules did not say rather than what they did, and to portray the text of the Rules as being so clear on the question of *Frye* that little discussion was needed. The Court thus moved on to analyze only one view of what the words meant—that the words created a flexible "relevancy and reliability" standard. While that analysis may or may not have been a good one, the Court's choice created the false impression of clarity of language. In fact, Rule 702 vaguely asks only whether the evidence "will assist the trier of fact," an ambiguity not clarified by its

---

[322] FED. R. EVID. 703 advisory committee's note.

[323] *See* United States v. Addison, 498 F.2d 741, 744 (D.C. Cir. 1974) ("Scientific proof may in some instances assume a posture of mystic infallibility in the eyes of a jury of laymen . . . ."; United States v. Amaral, 488 F.2d 1148, 1152 (9th Cir. 1973) ("aura of special reliability and trustworthiness" surrounds scientific testimony).

[324] Brief for Respondent, *supra* note 316, at 14–16.

[325] *See supra* text accompanying notes 309–324.

[326] *See, e.g.*, DAVID LOUISELL & CHRISTOPHER MUELLER, FEDERAL EVIDENCE § 105, at 818 (1977) (concluding that *Frye* survived the federal rules); United States v.

skimpy legislative history. There is thus no way to resolve the otherwise indeterminate question of what that test means except by resorting to a policy analysis.

But judicial policy analysis is often both unavoidable and desirable. Indeed, as will be discussed shortly and as other commentators have recently argued, the most sensible way to read the Rules is as a general guide designed to structure trial judge discretion and thus to permit the development in many areas of judge-made rules.[327] The Court's error was in seeking to mask this reality by reference to a non-existent clarity of language.

3. Relevancy and Reliability

a. *Interpretive communities.* Having established that *Frye* did not apply, the Court then asked itself what test should apply. The Court looked to the community of ordinary English speakers to define "knowledge" as "any body of known facts or . . . any body of ideas inferred from such facts or accepted as truths on *good grounds*."[328] This strategy of looking to the dictionary definition of a word is consistent with a plain-meaning approach to statutory construction.

The Court also sought to define "science," this time turning to scientists' own understandings rather than ordinary meanings.[329] Some *amici* briefs proposed a different definition than that offered by the proponents of *Frye*, emphasizing not general acceptance but rather science as a "*process* for proposing and refining theoretical explanations about the world that are subject

---

Downing, 753 F.2d 1224, 1235 (3d Cir. 1985) (recognizing that *Frye* is one method by which courts can assess the validity of novel scientific information and safeguard the jury against potentially flawed expert testimony).

[327] The companion article will develop this point in great detail. *See* Taslitz, *supra* note 15. Professor Giannelli, without developing a general theory for how to interpret the Rules, does seem to recognize this point in devoting an entire article to defending primarily one thesis: the Rules and their legislative history were silent about *Frye* because at the time the Rules were adopted most scientific techniques used in the federal courts were indeed generally accepted. Giannelli, *supra* note 237, at 1999, 2009 (1994). Thus, Edward Cleary, the reporter for the Rules, acknowledged in 1981 that "'the Advisory Committee did not specifically consider *Frye*, nor did the congressional groups to my knowledge.'" *Id.* at 2014 (quoting letter from Edward W. Cleary to Paul C. Giannelli (Feb. 12, 1981) (on file with Professor Giannelli)). Therefore, concludes Professor Giannelli, the *Daubert* Court should simply have admitted that the drafters never even thought about *Frye* and that the Court had no option but to decide the matter before it as a question of policy.

[328] Daubert v. Merrell Dow Pharmaceuticals, Inc., 113 S. Ct. 2786, 2795 (1993) (quoting WEBSTER'S THIRD INTERNATIONAL DICTIONARY 1252 (1986)) (emphasis added).

[329] *See supra* text accompanying notes 238–242.

to further testing and refinement."[330] "Scientific" knowledge, therefore, must be based on good grounds rooted in the scientific method, a standard the Court described as one of evidentiary reliability.[331]

The definitional problem facing the Court was thus solved by the choice of an interpretive community. The term "scientific knowledge" was not to be defined by lay understandings, or even by a somewhat more precise dictionary definition, but rather by the community of scientists, and within that community, a specific sub-community represented by *amici*. It is highly unlikely that Congress, members of its subcommittees, or anyone involved in the drafting of Rule 702 had this specific definition in mind. The turn to a specialized community apart from the ordinary citizen and apart from the community of lawyers is not a common use of a plain-meaning approach.[332]

The Court followed a third definitional strategy by giving the phrase "assist the trier of fact" the meaning assigned to it by elite lawyers, emphasizing that what is not relevant cannot "assist the trier of fact."[333] The Court also cited *United States v. Downing* for the proposition that relevance includes "fit," i.e., a close enough link between the testimony and the facts to aid the jury.[334] This concept of "fit" was linked to science by citing a law review article on the subject.[335]

Ultimately, the Court looked to three very different interpretive communities: (1) the ordinary speaker, as represented by a common dictionary (defining "knowledge"); (2) a sub-community of scientists (defining "science"); and (3) lawyers (defining "assisting the trier of fact"). Not once did the Court explicitly recognize or justify this shift.

Equally important, the Court insisted upon this definitional strategy to create the illusion of merely finding a plain meaning

---

[330] *See Daubert*, 113 S. Ct. at 2795 (quoting Brief for American Association for the Advancement of Sciences as *Amici Curiae*) (emphasis added).

[331] *See id.* at 2795.

[332] *See supra* text accompanying notes 66–73, 107–110 (Court has looked to dictionaries or lawyer usages in finding purported plain meanings).

[333] *Daubert*, 113 S. Ct. at 2795–96 (quoting WEINSTEIN & BERGER, *supra* note 176, ¶ 702[02], at 702-18).

[334] *See id.* at 2795–96 (citing United States v. Downing, 753 F.2d 1224, 1242 (3d Cir. 1985)).

[335] *See id.* at 2796 (citing James Starrs, Frye v. United States *Restructured and Revitalized: A Proposal to Amend Federal Evidence Rule 702*, 26 JURIMETRICS J. 249, 258 (1986)).

embodied in Rule 702. The Court seemingly followed the will of Congress. Consequently, it did not consider as a matter of policy whether the newly declared "relevancy and reliability" approach made more sense than *Frye*. Deciding which community's meaning should control, however, is itself a policy choice. As will be discussed in the companion article, such judicial policymaking is fully consistent with congressional intent.[336]

b. *Truncated policies from varied data.* The Court did offer some policy justifications for its new standard. Relying on a combination of text, structural arguments, legislative history, and the common law, the Court noted that Rules 702 and 703 remove the first-hand knowledge requirement. That requirement, however, represented "a 'most pervasive manifestation' of the common law insistence upon 'the most reliable sources of information,'" according to the Advisory Committee Note to Rule 602.[337] The Court reasoned that deleting the first-hand knowledge requirement for experts could thus be justified only if there were some alternative way to assure the most reliable source of information. To accomplish this, the Court read a reliability requirement into Rule 702.

Of course, the Court again ignored the question whether *Frye* might be a more effective way of assuring reliability. Moreover, because the Rule 602 Advisory Committee Note relies upon the common law, the Note suggests that the Rules' drafters expected references to be made to the common law to guide the quest for reliable sources of data. If this is true, then how could *Frye*, the predominant common law test for screening scientific evidence when the Rules were drafted, be so clearly superseded by Rule 702? The Court was silent on this issue.

c. *Judicial competence.* The Court next considered whether judges are capable of assessing whether their reasoning or methodology is scientifically valid. The Court, without citation, simply asserted its confidence that "federal judges possess the capacity to undertake this review."[338] This is quite an assumption to leave undefended, and fears about the inability of judges to make precisely these kinds of decisions were among the defenses offered for *Frye*.[339] Chief Justice Rehnquist

---

[336] *See* Taslitz, *supra* note 15.
[337] *Daubert*, 113 S. Ct. at 2796.
[338] *Id.*
[339] *See* Giannelli, *supra* note 172, at 1229.

in his dissent indeed raised some doubts about this assumption by his candid admission that he was "at a loss to know what is meant when it is said that the scientific status of a theory is its 'falsifiability'."[340] Moreover, if the Court is correct that the reliability inquiry is "embodied" in Rule 702, then is it not clear that Congress had confidence in the judicial ability to undertake this task? Rather than inquiring into congressional intent regarding the abilities of trial judges, however, the Court engaged in its independent assessment that judges have such abilities. If an independent judicial assessment is appropriate, then the Court implicitly recognized that there is a role for judicial policymaking under the Rules.

d. *Jury competence.* The Court's final justification for its newly articulated rule was that juries are fully capable of handling expert evidence, especially given the availability of the adversary system safeguards, such as vigorous cross-examination, presentation of contrary evidence, careful instruction on the burden of proof, and in the unusual case where no reasonable jury could reach a verdict in favor of a particular party, the availability of a directed verdict.[341] Again, no citation is offered to support these assertions, while there is indeed conflicting evidence about the capabilities of juries for handling such complex matters.[342] Similarly, it is questionable whether effective adversarial safeguards are always in place (e.g., whether trial lawyers are always competent).[343] All these considerations would have been relevant to whether *Frye* or some other rule made sense. The Court's unsupported assertions thus stand more as an article of faith than a defensible premise for its opinion.

e. *Philosophers, scholars, and judges.* The Court next sought to flesh out its test by articulating factors to guide trial judges. When articulating these factors the Court cited works on the philosophy and practice of science, case law that had carefully investigated the value of particular scientific techniques, and leading scholarly commentators

---

[340] *Daubert*, 113 S. Ct. at 2800.

[341] *Id.* at 2798.

[342] *Compare* Taslitz, *Does the Cold Nose Know?*, *supra* note 81, at 17–28, 59–61 (illustrating juror inability fairly to evaluate expert scent lineup evidence) *with* Edward J. Imwinkelried, *The Standard for Admitting Scientific Evidence: A Critique from the Perspective of Jury Psychology*, 28 VILL. L. REV. 554, 564, 566 (1983) (suggesting jurors are not overawed by scientific evidence).

[343] *See, e.g.*, Taslitz, *Does the Cold Nose Know?*, *supra* note 81, at 52–134 (reviewing the many arguments trial lawyers uniformly failed to raise to exclude an unquestionably unreliable form of scientific evidence).

on the law.[344] In short, the Court engaged in an investigation, however abbreviated, of the hallmarks of good scientific methodology and the ways in which courts can most simply and effectively identify and use such hallmarks. Such a wide-ranging, interdisciplinary, and policy-driven inquiry can be justified only if Congress, in adopting the Rules, intended to delegate to the courts the authority to develop the meaning of the term "assist the trier of fact." If so, then the Rules are not and should not be bound by rigid notions of plain meaning and other purportedly mechanical rules of statutory interpretation.

f. *Guided discretion.* Of equal importance, however, is what the Court did not do. The *amici* argued that the Court should adopt a *de novo* standard of review for scientific information that "transcends a particular dispute."[345] Such facts are more like legislative facts, which the Court has traditionally reviewed under a *de novo* standard.[346] Furthermore, there is no justification for permitting different courts to reach different decisions on such matters. Thus, if a plaintiff alleges that smoking caused his lung cancer, there should be *de novo* review of the transcendent scientific notion that relevant experts are capable of drawing such conclusions, and at the same time review under the more traditional abuse of discretion standard for the case-specific inquiry whether the particular expert in the case was so qualified.[347] The Court never expressly addressed what the standard of review from a *Daubert* ruling should be, but it did give some significant hints.

First, the *amici* conceded that if the traditional abuse of discretion standard applied, then Rule 104 governed. The Court, presumably after reviewing the briefs, concluded that Rule 104 indeed controlled.[348] Second, the Court emphasized that it was articulating factors to guide trial judge decisionmaking.[349] However, it made clear that these factors were not exhaustive, and that it expected trial judges to craft new factors as dictated by the needs of experience.[350] Furthermore, the Court emphasized its confidence in the ability of trial judges to make these kinds

---

[344] *See supra* text accompanying notes 249–262.

[345] *See* Law Professors, *supra* note 288, at 22.

[346] *Id. See also* Giannelli, *supra* note 237, at 2011–12; David L. Faigman, Elise Porter, & Michael J. Saks, *Check Your Crystal Ball at the Courthouse Door, Please: Exploring the Past, Understanding the Present, and Worrying About the Future of Scientific Evidence*, 15 CARDOZO L. REV. 1799, 1820–22 (1994).

[347] *See id.*

[348] Daubert v. Merrell Dow Pharmaceuticals, Inc., 113 S. Ct. 2786, 2796 n.10 (1993).

[349] *See id.* at 2796–97.

[350] *See id.* at 2796 ("Many factors will bear on the inquiry, and we do not presume to set out a definitive checklist or test.").

of decisions and that these decisions should be "flexible."[351] Additionally, trial judges were instructed to be "mindful" of other applicable rules, including the need to be sensitive to the potential impact of the evidence on a jury under Rule 403,[352] a matter traditionally recognized to be within the sound discretion of the trial judge.[353]

In short, the Court's emphasis on the flexible nature of the inquiry, on balancing, on the need for sound judgment, and on the abilities of trial judges to perform these tasks suggests that the Court was establishing a set of flexible guidelines for the exercise of trial judge discretion, not a rigid set of rules. Furthermore, this emphasis also suggests that discretion can exist only if there is some deference to trial judge decisionmaking, and, in particular, the standard of review called abuse of discretion.[354]

As *amici* suggested, this approach may create inconsistencies, and there will be practical problems deciding when discretion has been abused. Nevertheless, the allocation of significant discretion to a trial judge in making decisions about the admissibility of scientific evidence is consistent with and probably mandated by the Rules.[355] Yet the Court's purported plain-meaning jurisprudence and its alleged reliance on seemingly mechanical tools for statutory interpretation is inconsistent with the exercise of trial judge discretion and the related case-by-case growth of evidence law by common law methods that such discretion entails.

In conclusion, ambiguous phrases like "assist the trier of fact" simply do not have a plain meaning. Moreover, assigning meaning to a text necessarily turns on the choice of interpretive community, which is itself a policy-driven decision. *Daubert* thus demonstrates the emptiness of a text-centered jurisprudence.

On the other hand, beneath the new textual rhetoric, *Daubert* reveals a sensitivity to the need for judicial value choices and an awareness of the role of judicial discretion within the broad

---

[351] *Id.* at 2797.

[352] *Id.* at 2798.

[353] *See* 2 SALTZBURG & MARTIN, *supra* note 203, at 164–75 (summarizing Rule 403 cases reviewed under abuse of discretion standard).

[354] *See* Leonard, *supra* note 299, at 1227. The Court's reliance on United States v. Downing, 753 F.2d 1224, 1240 (3d Cir. 1985) ("Therefore, we will review district court decisions to admit or exclude novel scientific evidence by an abuse of discretion standard.") further supports this conclusion. *See Daubert*, 113 S. Ct. at 2197.

[355] *See* Taslitz, *supra* note 15.

limits set by text and by whatever other indicators of congressional intent are available. This sensitivity is consistent with the Court's continued adherence to a more flexible, pragmatic approach to reading the Rules that is embodied in portions of other decisions. While new textual rhetoric continues to distort sound decisionmaking, an awareness of the more pragmatic approach embodied in the Court's decisions lays the groundwork for a more sensible evidentiary jurisprudence.

III. AFTER *DAUBERT*: *WILLIAMSON V. UNITED STATES* AND THE APPOINTMENT OF JUSTICE BREYER

Since *Daubert*, two events have raised hopes of the Court's moving toward a politically realistic hermeneutics under the Rules. First, the Court has decided *Williamson v. United States*,[356] which hints that several justices may accept a policy-oriented approach to the Rules. Second, Associate Justice Stephen Breyer has replaced retiring Associate Justice Harry Blackmun.[357]

A. Williamson v. United States

Rule 804(b)(3) excepts from the hearsay rule "a *statement* which . . . at the time of its making . . . so far tended to subject the declarant to . . . criminal liability . . . that a reasonable person in the declarant's position would not have made the statement unless believing it to be true."[358] *Williamson* raised the question whether a criminal defendant's confessions were "statements" within the meaning of this Rule.

Police officers stopped Reginald Harris in a rental car while he was driving on a highway. Harris consented to a search of the car, which revealed a quantity of cocaine in two suitcases in the trunk. At first Harris claimed that an unnamed Cuban had given him the cocaine to place in a dumpster for a pickup. After the arresting officer suggested making a controlled buy, however, Harris changed his story, saying that he had in fact been trans-

---

[356] 114 S. Ct. 2431 (1994).

[357] Joan Biskupic, *Breyer Takes Court Oath At Chief Justice's Cottage*, WASH. POST, Aug. 4, 1994, at A8.

[358] FED. R. EVID. 804(b)(3) (emphasis added). Where such statements are offered to exculpate the accused, the Rule further requires that "corroborating circumstances clearly indicate the trustworthiness of the statement." *Id.*

porting the cocaine to Atlanta for Fredel Williamson, who had witnessed Harris's arrest and therefore would not be caught in a phony buy.

Despite these admissions, which were buttressed by physical evidence linking Williamson to Harris's car, and despite Harris's grant of immunity, Harris refused to sign a written statement or testify at Williamson's trial.[359] Harris's confession was admitted at trial against Williamson, who was subsequently convicted of possessing cocaine with intent to distribute and related charges. The lower court apparently ruled that (1) every portion of the confession was against Harris's penal interest, (2) the corroborating circumstances sufficiently ensured the trustworthiness of the testimony, and (3) given Harris's unavailability, his confessions were therefore admissible.

1. Justice O'Connor's Opinion

In an opinion authored by Justice O'Connor, the first two portions of which were joined by Justices Blackmun, Stevens, Scalia, Souter, and Ginsburg, the Court held that Rule 803(b)(3) does not allow for the admission of non-self-inculpatory statements, even when they are made as part of a broader, generally self-inculpatory narrative.[360] Accordingly, the Court vacated the judgment of the Court of Appeals and remanded the case for further proceedings consistent with its opinion, a difficult task because there was no majority concerning the proper procedure on remand.

The Court agreed that there was no plain meaning of the term "statement," as the dictionary offered two conflicting definitions: (1) a "report or narrative," supporting the trial court's conclusion, and (2) a "single declaration or remark," rejecting the trial court's approach.[361] The Court held that the "principle behind the Rule, so far as it is discernible from the text, points clearly to the narrower reading."[362] The Court reasoned that reasonable people, including dishonest ones, tend not to make self-inculpatory statements unless they believe them to be true. "One of the most effective ways to lie is to mix falsehood with truth, espe-

---

[359] 114 S. Ct. at 2434.
[360] *Id.* at 2435.
[361] *Id.* at 2434–35.
[362] *Id.* at 2435.

cially truth that seems particularly persuasive because of its self-inculpatory nature."[363] Consequently, limiting the exception to only *individual* statements likely to be trustworthy guards against this danger. With this textual answer the Court's analysis—were it a new textualist one—would have been complete.

By analyzing the Rule's legislative history,[364] however, the Court departed from the new textualism. The Court's analysis of legislative history, specifically of the Advisory Committee Note, was framed as a response to Justice Kennedy's partial dissent, which argued for the presumptive admission of most portions of narrative confessions. However, the Court nevertheless left the door open to a radical new textualism, stating that it was not deciding how much weight to give the Advisory Committee Note in this particular situation because the policy behind the Rule clearly outweighed the Note.[365]

Nevertheless, the Court did examine the Advisory Committee Note.[366] The Note expresses what Justice Kennedy interpreted as a clear preference for admitting non-self-inculpatory statements when they are related to self-inculpatory ones: "[T]he third-party confession . . . may include statements implicating [the accused], and under the general theory of declarations against interest *they would be admissible* as related statements."[367] Justice O'Connor concluded, however, that the Note was not quite so clear, for it went on to declare that whether such statements are in fact against the defendant's interest must be determined from the *circumstances of each case*.[368] Those circumstances might sometimes render the statement inadmissible when, for example, made to "curry favor" with the authorities.[369] On the other hand, a statement implicating another might be admissible when made

---

[363] *Id.*

[364] *See id.* at 2435–36.

[365] *Id.* at 2436. This statement is a curious one. If the Note actually reflects congressional intent, then how can a contrary reading of policy show appropriate deference to the legislature? Two readings are possible. First, only text, not congressional "intent," matters to this Court. If so, this reading is disingenuous, for the text simply does not have the clarity regarding how the policy stated in the rule applied to this situation. A second, alternative reading is that the Court recognized that both text and legislative history were ambiguous—as the Court indeed stated—and that the Court therefore had no choice but to apply the general policy in the way the Court decided wisest for the particular problem before it. This would be a more candid reading, but the Court chose instead to adhere to a new textualist rhetoric, apparently and incorrectly concluding that this would promote the Court's legitimacy.

[366] *Id.* at 2435–36.

[367] *Id.* (quoting FED. R. EVID. 804(b)(3) advisory committee's note) (emphasis added).

[368] *See id.* at 2436–37.

[369] *Id.* at 2436.

to a third party from whom the declarant cannot expect any reward.[370] Similarly, the Note cites Professor McCormick, who favored a narrow reading of the word "statement." "[A]dmit[ting] the disserving parts of the declaration, and exclud[ing] the self-serving parts . . . seems the most realistic method of adjusting admissibility to trustworthiness, where the serving and disserving parts can be severed."[371]

In the last part of her opinion, joined only by Justice Scalia,[372] Justice O'Connor emphasized that some portions of Harris's statement were neutral or might be thought by a reasonable person to be exculpatory because "[s]mall fish in a big conspiracy often get shorter sentences than people who are running the whole show."[373] Nevertheless, the inquiry was a fact-sensitive one, requiring a remand for careful examination of the circumstances. This emphasis on the need for fact-gathering and an initial determination of trustworthiness by the trial court implicitly recognized that some trial court discretion is required.[374]

2. Justice Kennedy's Opinion

Justice Kennedy, joined by Chief Justice Rehnquist and Justice Thomas,[375] agreed with Justices O'Connor and Scalia that a preliminary-fact determination by the trial judge under Rule 104(a) was necessary because the question involved a "difficult, factbound determination."[376] Justice Kennedy more explicitly emphasized the need for trial court discretion. Justice Kennedy also agreed that the policy behind the Rule was to admit statements likely to be trustworthy because they were made against the declarant's penal interest.[377] In his view, however, the Rule was silent about the admission of statements collateral to trustworthy, self-incul-

---

[370] *See id.* at 2436–37.

[371] *Id.* at 2436 (quoting CHARLES MCCORMICK, LAW OF EVIDENCE § 256, at 551–53 (1954)).

[372] *Id.* at 2437–38.

[373] *Id.* at 2437.

[374] Justice O'Connor's emphasis on discretion contrasted sharply with the approach of Justice Ginsburg, who, in an opinion joined by Justices Blackmun, Stevens, and Souter, saw no role for the trial judge. *Id.* at 2439. Justice Ginsburg would have excluded the entire confession, remanding solely to consider whether the error was harmless because, in her view, the entire confession was simply an effort by someone caught in the act to shift blame to Williamson. *See id.* at 2438–40.

[375] *Id.* at 2440–45.

[376] *Id.* at 2445.

[377] *Id.* at 2440.

patory statements.[378] Indeed, at the time the Rules were adopted the commentators widely debated this very question.[379]

The Advisory Committee Note clearly indicated that at least some collateral statements, i.e., statements not in themselves inculpatory, were admissible and seemed to favor their admissibility.[380] Moreover, the citation to Professor McCormick further demonstrated that collateral neutral statements particularly were to be favored.[381] In Justice Kennedy's view, the majority had ignored the Court's practice of paying attention to the Advisory Committee Note where a Rule's text is silent.[382] Moreover, the majority erred in ignoring the Court's previously stated presumption that, absent contrary indications, the common law rule was meant to be adopted by Congress; and at the time the Rules were adopted, the common law rule was to admit some collateral statements.[383] Indeed, the Note apparently adopted this common law view by accepting the "general theory" that collateral statements were admissible.

Justice Kennedy's most heated words were reserved for the majority's policy analysis. It was disingenuous, he suggested, for the majority to argue that a Rule silent on the treatment of collateral statements expressed clear policy support for anything.[384] More importantly, as a practical matter, most self-inculpatory statements referring to one or more co-conspirators—the only kind that bring the penal exception into play—will not be entirely against the declarant's self-interest, presumably because some responsibility, however indirectly, is foisted on another. To exclude all such statements would render the exception meaningless.

According to Justice Kennedy, therefore, the only sensible rule was to admit all of a statement containing a fact against penal interest, subject to two limits: (1) excluding collateral statements that are *so* self-serving that they are unreliable, and (2) excluding an entire statement made under circumstances where it was likely that the declarant had a significant motive to obtain favor-

---

[378] *Id.* at 2440–42.

[379] *See id.* at 2440–41 (summarizing the debate).

[380] *Id.* at 2442–43.

[381] McCormick favored admitting collateral neutral statements ("John and I robbed the bank"), but excluding self-serving collateral statements ("John, not I, shot the bank teller"). *Id.* at 2441.

[382] *Id.* at 2442.

[383] *Id.* at 2442–43.

[384] *Id.* at 2441–42.

able treatment, such as a promise of leniency.[385] The first limit requires the trial court to balance self-serving against disserving aspects of the declarant's statement, at least where such aspects are severable.[386] That balancing would likely require admitting collateral neutral statements and excluding self-serving ones.

3. Justice Scalia's Concurrence

Justice Scalia responded to Justice Kennedy by denying that the majority had adopted the "simplistic" reading that he suggested.[387] While the majority did reject admissibility of an extended narrative simply because it included a self-inculpatory statement, "a declarant's statement is not magically transformed from a statement against penal interest into one that is inadmissible merely because the declarant names another person or implicates a possible co-defendant."[388] Justice Scalia argued, for example, that if a lieutenant in an organized crime operation named other actors and described the inner workings of an extortion racket, such statements would be admissible under the majority's opinion.[389] If the context showed that the declarant minimized his own culpability, such minimization would affect but not necessarily determine admissibility.[390] The admissibility question should be decided on the basis of the Rule's policy, not some arbitrary distinction between "collateral neutral" and "collateral self-serving" statements.[391]

4. Interpretive Implications

Williamson is the first major Rules opinion in which the prosecution did not prevail.[392] A careful reading, however, indicates that the decision likely did not favor the defense. First, as Justice Kennedy pointed out, the majority's view may sometimes work against the defense by excluding some trustworthy but exculpatory statements. Moreover, both Kennedy's view and Scalia's

---

[385] *Id.* at 2445.
[386] *See id.* at 2444–45 (citing MCCORMICK, *supra* note 367, § 256, at 553).
[387] *Id.* at 2438.
[388] *Id.*
[389] *Id.*
[390] *See id.*
[391] *Id.*
[392] Of course, the only result was a remand, leaving the possibility that the prosecution will yet prevail.

version of the majority opinion allow significant trial judge discretion under the penal interest exception, thus opening the door for potential prosecution victories under a law-and-order trial judiciary. Second, under any plausible reading of the Rule, *some* collateral statements will be excluded, as Professor McCormick, the Advisory Committee, and the logic of the Rule itself suggest. Language is rarely, if ever, "plain," but neither is it infinitely malleable within a given social framework. The Court could not, therefore, allow every statement that implicates another person to be automatically admissible. Some prosecution losses are unavoidable unless the Court openly declares a pro-prosecution bias, a move that would seriously endanger its legitimacy. Given that some limits on the prosecution are unavoidable, some members of the *Williamson* Court did what they could to limit the damage.

Justice Kennedy's approach would probably help the prosecution the most.[393] He would in effect favor a presumption of admissibility of collateral statements in confessions implicating or exculpating third parties, and such confession work to the prosecution's advantage more often than they do the defense's. Justices O'Connor and Scalia's approach expresses no such presumption, yet Scalia's concurrence seems to recognize that the prosecution will still be able to use many powerful statements. Most importantly, under either Kennedy's or O'Connor's approach, trial judges have great discretion, and, as just discussed, given that an absolute rule favoring the prosecution is not viable, a discretionary approach is the prosecution's next best bet.

The point is not that the justices consciously or subconsciously favor the prosecution. Rather, even after *Williamson*, their behavior is consistent with a pro-prosecution bias, one that may help to explain their behavior but does not necessarily lead them to an incorrect result. To the contrary, this Article has argued that a discretionary approach is most consistent with the overall design of the Rules. Therefore, either a neutral approach or a pro-prosecution bias should lead the Court to continue moving toward such a discretionary approach.

While the Court's attitude toward legislative history is still uncertain, in light of its other decisions construing the Federal Rules of Evidence, it seems likely that such history will continue

---

[393] This is so despite Justice Kennedy's protestations that his approach would be more favorable to the defense than the majority's approach. *See* 114 S. Ct. at 2440–45.

to have an important role in many decisions interpreting the Rules. Indeed, every justice in *Williamson* has signed onto an opinion that considered legislative history. Moreover, the focus on sound policy and trial judge discretion is fully consistent with a more flexible, politically realistic hermeneutics, though the Court has yet to apply a self-conscious, hierarchical version of this approach.

B. *Justice Breyer's Appointment*

The arrival of Associate Justice Stephen Breyer should move the Court even closer to a hermeneutic approach. In his scholarly writings Justice Breyer has expressed strong support for consulting legislative history, albeit with a more thorough and politically astute eye.[394] His experience as chief counsel for the Senate Judiciary Committee suggests he is well-accustomed to the practical politics of lawmaking.[395] Moreover, he has recognized the unavoidable role of policy analysis in statutory interpretation.[396] In his recent Senate confirmation hearings he described himself as a pragmatist,[397] and the media saw him as committed to a case-by-case approach sensitive to real world concerns, rather than bound by a single overarching theory of statutory interpretation.[398] Indeed, in those hearings he emphasized the need to rely on a wide range of sources in interpreting statutes.[399] Perhaps most importantly, he has stressed the need to interpret a statute's purpose—what the legislature was trying to achieve—and then consider "how the interpretation is likely, in light of that purpose, to work out in the world."[400]

---

[394] *See* Stephen Breyer, *On the Uses of Legislative History in Interpreting Statutes*, 65 S. CAL. L. REV. 845, 864–65 (1992).

[395] Linda Greenhouse, *Portrait of a Pragmatist*, N.Y. TIMES, July 14, 1994, at D22.

[396] *See id.*; Breyer, *supra* note 394, at 872.

[397] Greenhouse, *supra* note 395, at A1, D22.

[398] *See id.* at A1.

[399] *See* Joan Biskupic, *Breyer Gives View of How He 'Judges'*, WASH. POST, July 14, 1994, at A6 (In his testimony before the Senate Judiciary Committee, "Breyer . . . often cited—as tools for judging the text of a law—its history, legal precedent, the conditions of life in the past, the present and 'a little bit of projection into the future.'") (quoting Breyer).

[400] *See* Greenhouse, *supra* note 395, at D22; *see also* Biskupic, *supra* note 399, at A6 ("'The present and past traditions of our people are important because they can show how past language reflecting past values . . . apply in present circumstances.'") (quoting Breyer).

Williamson revealed a strong movement by Justices Kennedy and Thomas, and Chief Justice Rehnquist, toward a politically realistic hermeneutics. Moreover, Justice Stevens, although signing onto the majority opinion in *Williamson*, wrote *Green*, which both rejected plain meaning and insisted on an extended examination of a Rule's intellectual history.[401] Because of his emphasis on the importance of pragmatic, case-specific reasoning and his willingness to examine legislative history and goals, Justice Breyer will likely be sensitive to trial judge discretion on evidentiary rulings and thus give adherents to this approach a working majority of five.[402] Moreover, Justice Breyer's reported skill at building consensus could sway others to his position. Thus, the future of a politically realistic hermeneutics as a guide to interpreting the Rules looks bright.

IV. CONCLUSION

Contrary to the prevailing wisdom, the Court's evidence decisions do not reflect a rigid plain-meaning jurisprudence. Instead, the cases reveal a flexible, pragmatic philosophy, which consid-

---

[401] *See supra* text accompanying notes 128–154.

[402] A preliminary test of this prediction will be made this term when the Court decides two pending evidence cases: United States v. Mezzanato, 998 F.2d 1452 (9th Cir. 1993), *cert. granted* 114 S. Ct. 1536 (1994), and United States v. Tome, 3 F.3d 342 (10th Cir. 1993), *cert. granted* 114 S. Ct. 1048 (1994).

Mezzanato raises the question whether a criminal defendant may waive the protections of Rule 410, which generally prohibits the admissibility of a criminal defendant's statements made during plea negotiations. *See Mezzanato*, 998 F.2d at 1454 ("The legislative history of these Rules is quite clear that plea negotiation statements are not admissible to impeach a defendant . . . . Commentators offer no disagreement."). Both the majority and dissenting opinions agreed that this was an "issue to which Congress did not speak." *Id.* at 1454, 1459. One version of textualism concludes that if the Rule does not mention waiver, it is therefore prohibited. *See* Frank Easterbrook, *Statutes' Domain*, 50 U. CHI. L. REV. 533, 540–48 (1983). This is a decidedly pro-prosecution result. A more candid recognition of the need for pragmatic judicial gap-filling, however, will require the Court to engage in judicial policymaking consistent with Congress's broad goals.

Tome raises the question whether Rule 801(d)(1)(B)'s hearsay exception for statements "consistent with the declarant's testimony and . . . offered to rebut an express or implied charge . . . of recent fabrication or improper influence or motive" applies only to statements made *prior* to the time a motive to fabricate arose. *See* FED. R. EVID. 801(d)(1)(B); *Tome*, 3 F.3d at 349–51. The text of the Rule, fairly read, does not resolve this issue. A textualist reading of the Rule is that no pre-motive requirement may be read into the Rule because the Rule does not expressly require one. In response, one might argue that a more natural reading of the Rule requires that the prior consistent statement precede the motive to fabricate, for only then can the earlier statement "rebut" a charge of "recent" fabrication. Given these two plausible readings of the Rule, the Court should turn to the common law, the Advisory Committee Note, and sound policy to craft a fair interpretation.

ers a wide range of data, including text, legislative and intellectual history, broad purposes, and practical implications. The companion article will develop the theoretical justifications for this pragmatic interpretive approach and suggest ways to structure and organize the interpretive enterprise.[403]

[403] *See* Taslitz, *supra* note 15.

ARTICLE

BACK TO THE FUTURE: APPRAISAL RIGHTS IN THE TWENTY-FIRST CENTURY

MARY SIEGEL*

Though the roots of the appraisal remedy extend back more than a century, recent trends in corporate practices and recent developments in corporate law have intensified scholarly debate on the subject. In this Article, Professor Siegel examines the remedy's nineteenth-century origins and the purposes for which the remedy was created. The author then analyzes whether these purposes remain viable rationales for the appraisal remedy in the twentieth century. She concludes with a series of recommendations intended to tailor the appraisal remedy to dual modern purposes: to provide a fair cash exit for shareholders and to monitor the actions of management in certain conflict-of-interest transactions.

There is a resurgence of interest in appraisal rights,[1] a remedy that provides shareholders who are dissatisfied with certain corporate transactions the option of having their shares repurchased by their corporation for cash.[2] Due in part to the high cost of

---

*Professor of Law, Washington College of Law, The American University. A.B., Vassar College, 1972; J.D., Yale University, 1975.

The research for this Article was supported by research funds from the Washington College of Law.

The author gratefully acknowledges the invaluable research assistance of David Cohen, J.D. 1995, Washington College of Law, and Thomas J. Rebull, J.D. 1995, Washington College of Law. The author also thanks Louis J. Trotter, Jr. for his editorial assistance, and Dennis J. Block, Stephen W. Carr, Patrick Garrett, Michael Goldman, James J. Hanks, Jr., William Prickett, John Subak, and Peter Walsh for their helpful comments. Finally, the author thanks the previous and current members of the Corporate Laws' Appraisal Rights Task Force for their numerous insights. Comments about this Article may be addressed via Internet to HJOL@HULAW1.HARVARD.EDU.

[1] *See* ERNEST L. FOLK III ET AL., FOLK ON THE DELAWARE GENERAL CORPORATION LAW § 262.1 (3d ed. 1992) (stating that appraisal rights are flourishing, contrary to the author's observation 20 years before that appraisal rights had decreased to the point whereby elimination was possible).

[2] *See* Hideki Kanda & Saul Levmore, *The Appraisal Remedy and the Goals of Corporate Law*, 32 UCLA L. REV. 429, 429 (1985) (defining the appraisal remedy). Often, the acquiring or merged entity pays the dissenting stockholder. *See, e.g.*, CONN. GEN. STAT. ANN. § 33-369(e) (West 1987) ("The surviving or new corporation shall . . . be responsible and liable for all the liabilities, obligations and penalties, including liability to dissenting shareholders, of each of the merging or consolidating corporations"); DEL. CODE ANN. tit. 8, § 262(i) (1974) ("The Court shall direct payment of the fair value of the shares . . . by the surviving or resulting corporation to the stockholders entitled thereto"); *cf.* MODEL BUSINESS CORP. ACT ANN. § 13.01 cmt. 1 (P-H Supp. 1986) [hereinafter MBCA ANN.] (noting that in a merger or share exchange, an acquiring or successor corporation must assume the obligations of the acquired or disappearing corporation).

exercising appraisal rights, the remedy has long been viewed as useless except to shareholders with a large number of shares.[3] The increase in the holdings of institutional investors, however, has produced a stable of such large-bloc shareholders.[4] Moreover, a decline in the number of tender offers,[5] for which no corporate statute grants appraisal rights, and the growth in the number of mergers,[6] for which every corporate statute grants appraisal rights,[7] has vastly increased the frequency of appraisal-triggering transactions. Finally, the U.S. Supreme Court recently suggested that shareholders without sufficient voting power to alter the outcome of a vote should seek protection in state law

---

[3] *See, e.g.*, Green v. Santa Fe Indus., Inc., 533 F.2d 1283, 1297–98 n.4 (2d Cir. 1976) (Mansfield, J., concurring) (observing that in Delaware, only minority shareholders owning large blocs of shares would find appraisal financially beneficial), *rev'd on other grounds*, 430 U.S. 462 (1977); Pellman v. Cinerama, Inc., 503 F. Supp. 107, 110 (S.D.N.Y. 1980) (noting that individual minority shareholders with small claims would find appraisal economically unattractive, given costs for discovery and expert witnesses); *see also* WILLIAM FLETCHER, FLETCHER CYCLOPEDIA OF THE LAW OF PRIVATE CORPORATIONS § 5906.160 (1993) (citing the expense of an appraisal proceeding as a deterrent to individual shareholders); Roberta Romano, *The Political Economy of Takeover Statutes*, 73 VA. L. REV. 111, 175 (1987) (noting that small shareholders are unlikely to pursue appraisal rights, as they cannot spread their costs over a large number of shares); *cf.* Bayless Manning, *The Shareholders' Appraisal Remedy: An Essay for Frank Coker*, 72 YALE L.J. 223, 232–33 (1962) ("[I]t is hard to see that the average shareholder . . . can hope to gain anything from the [appraisal] statutes"); Joel Seligman, *Reappraising the Appraisal Remedy*, 52 GEO. WASH. L. REV. 829, 860 (1984) (observing that cost provisions of appraisal statutes constitute a substantial barrier to initiation of an appraisal); Robert B. Heglar, Note, *Rejecting the Minority Discount*, 1989 DUKE L.J. 258, 271 n.2 (1989) (describing the appraisal remedy as both unwieldy and expensive, partially due to high costs for attorneys and expert witnesses).

[4] The percentage of U.S. equities outstanding held by U.S. institutional investors rose from 6.1% of total equities outstanding in 1950, to 43.6% of total equities outstanding as of the third quarter of 1993. NYSE, FACT BOOK FOR THE YEAR OF 1993 89 (1994); *see also* Carolyn Kay Brancato, *Institutional Investors and Corporate America: Conflicts and Resolutions, reprinted in The Impact of Institutional Investors on Corporate Governance, Takeovers, and the Capital Markets: Hearing Before the Subcommittee on Securities of the Senate Committee on Banking, Housing, and Urban Affairs*, 101st Cong., 1st Sess. 3, 52 (1990) (observing that in 1988 institutional investors owned, on average, 52% of the outstanding shares of the top 50 U.S. corporations); Robert D. Rosenbaum & Michael E. Korens, *Trends in Institutional Shareholder Activism: What the Institutions Are Doing Today, in* INSTITUTIONAL INVESTORS: PASSIVE FIDUCIARIES TO ACTIVIST OWNERS 47 (1990) (observing that during the 1980s, institutional shareholders became much more active in corporate affairs); Clifford L. Whitehill, *Institutional Ownership, in* INSTITUTIONAL INVESTORS: PASSIVE FIDUCIARIES TO ACTIVIST OWNERS 75, 79–80 (1990) (declaring that "institutions are the market").

[5] MERGERSTAT REVIEW 1993, at 82 (1993). The annual number of tender offers for publicly traded companies has fallen from a high of 217 in 1988 to 32 in 1993, with a low of 18 in 1992. *Id.*

[6] *Id.* at 2. After falling from a high of 3336 in 1986, the year before the October 1987 stock market crash, the number of merger and acquisition announcements rose from 2032 in 1987 to 2663 in 1993. *Id.*

[7] MBCA ANN., *supra* note 2, § 13.02 statutory comparison, at 1371–73 (Supp. 1992 & Supp. 1993) (stating that all 52 jurisdictions grant appraisal rights in some merger situations).

remedies instead of in the proxy rules of the federal securities laws.[8] The appraisal remedy will thus assume increased importance as these shareholders consider their remaining options.

Increased numbers of institutional investors and appraisal-triggering transactions and the decreased availability of the federal securities laws contribute to a renewed interest in appraisal rights. That interest highlights the significant inconsistencies in appraisal rights in the fifty corporate statutes, the Model Business Corporation Act ("MBCA" or "Model Act"), and the American Law Institute's *Principles of Corporate Governance* ("*Principles*").[9] These statutes differ more in the appraisal area, perhaps, than in any other area in corporate law. For example, no consensus exists on which transactions should trigger appraisal rights, which classes of shareholders should have these rights, what procedures should be used, how costs should be allocated, how fair value should be computed, or whether and when appraisal rights should be a shareholder's exclusive remedy.[10] This increased attention to appraisal rights not only has refocused the corporate world on these differences and the lack of an identifiable cause

---

[8] Virginia Bankshares, Inc. v. Sandberg, 501 U.S. 1083 (1991). Because the Virginia statute denies appraisal rights to bank shareholders in a merger, VA. CODE ANN. § 13.1-718, the Court in *Sandberg* did not have to decide whether the federal proxy rules provide a remedy for shareholders who forego their appraisal remedy due to a material misrepresentation or omission in the proxy materials. The four dissenting Justices interpreted the majority opinion as denying a federal remedy to such shareholders. 501 U.S. at 1112 (Kennedy, J., dissenting). *But see* Wilson v. Great Am. Indus., 979 F.2d 924 (2d Cir. 1992) (finding that a minority shareholder who lost his appraisal rights due to misrepresentation could nevertheless maintain a cause of action under § 14(a) of the Securities Exchange Act); *cf.* Howing Co. v. Nationwide Corp., 972 F.2d 700, 708–09 (6th Cir. 1992) (holding that the loss of a state law remedy creates a federal action under § 13(e) of the Securities Exchange Act), *cert. denied*, 113 S. Ct. 1645 (1993).

[9] *See* 2 PRINCIPLES OF CORPORATE GOVERNANCE: ANALYSIS AND RECOMMENDATIONS pt. 7, ch. 4 (1994) [hereinafter PRINCIPLES] (report of the ALI, outlining model standards for the appraisal remedy).

[10] For example, the Model Act grants appraisal rights for many triggering events, including mergers, share exchanges, a sale or exchange of substantially all corporate property, and certain amendments to the articles of incorporation; but the Model Act permits only voting shareholders to exercise appraisal rights. MODEL BUSINESS CORP. ACT § 13.02 (1984) [hereinafter MBCA]. Delaware, in contrast, grants appraisal rights only in the event of a merger or consolidation but grants these rights to both voting and non-voting stockholders. DEL. CODE ANN. tit. 8, § 262(a),(b) (1991). The Model Act and Delaware law also differ in their prepayment provisions. *Compare* MBCA § 13.25(a) (1984) (directing the corporation to pay its estimate of fair value when the corporate action is taken or upon receipt of the payment demand) *with* DEL. CODE ANN. tit. 8, § 262(d)–(f) (1991) (not requiring any prepayment). Additionally, whereas the Model Act generally assesses costs against the corporation, Delaware allows the court to allocate costs as it deems equitable. MBCA § 13.31(a) (1984); DEL. CODE ANN. tit. 8, § 262(j) (1991).

for such diverse results, but also has created interest in their resolution.

The extant literature has raised but failed to resolve most of these issues. The first round of scholarly assault on the remedy noted the paucity of appraisal cases and railed against the remedy's cumbersome, technical and expensive process.[11] Scholars argued that the many steps required to perfect appraisal rights, as well as high costs and expenses for attorneys and experts, greatly deterred utilization of the remedy.[12] Fix these problems, it was argued, and the remedy would be resurrected. In response to these concerns, in 1978 the Committee on Corporate Laws substantially revised the MBCA's appraisal provisions. The revisions simplified the appraisal process,[13] encouraged settlement,[14]

---

[11] *See, e.g.*, Manning, *supra* note 3, at 231 (noting that the appraisal procedure "has grown long and expensive," and concluding that "the dissenting shareholder faces an unattractive and complex procedural obstacle course"); MELVIN A. EISENBERG, THE STRUCTURE OF THE CORPORATION: A LEGAL ANALYSIS 70 (1976) (reaffirming Manning's observations that the appraisal process is "highly technical, drawn-out, and expensive"); Richard M. Buxbaum, *The Dissenter's Appraisal Remedy*, 23 UCLA L. REV. 1229, 1253–54 (1976) (declaring that California created an appraisal remedy which is both technical and difficult to use); *see also* Morris v. Peoples Indep. Bancshares, Inc., 632 So. 2d 1340 (Ala. 1994) (denying shareholders appraisal rights because they failed to submit their stock certificates for notation as required by ALA. CODE § 10-2A-163(h), even though the corporation was not prejudiced by such failure and corporation knew shareholders were demanding their appraisal rights); Schneyer v. Shenandoah Oil Corp., 316 A.2d 570 (Del. Ch. 1974) (denying appraisal where the shareholder submitted his complaint to the court three days late).

[12] Professor Eisenberg characterized appraisal rights as "a remedy of desperation" for minority shareholders. EISENBERG, *supra* note 11, at 83 (stating that appraisal right is technical, expensive, uncertain in result, and ultimate award is taxable); *see* Manning, *supra* note 3, at 233 ("The only things certain [for a dissenting shareholder pursuing the appraisal remedy] are the uncertainty, the delay, and the expense"). Dean Manning observes that appraisal statutes are, at best, "of modest and infrequent help to the dissenting shareholder . . . ." *Id.* at 238; *see also* Michael Phillips, *Weinberger to Rabkin: Fine Tuning the Doctrine of Corporate Mergers*, 11 DEL. J. CORP. L. 839, 842–43 (1986) (observing that even under Delaware's "liberalized" appraisal remedy, shareholders would, in many instances, forego the burdens of appraisal because the costs outweigh the expected gains); Seligman, *supra* note 3, at 860 (arguing that cost provisions of appraisal statutes are a "substantial barrier" to exercising appraisal rights). Seligman asserts that these costs constitute a greater barrier to initiating appraisal than do procedural requirements. *Id.*

[13] *See, e.g.*, MBCA §§ 13.20–13.23 (1984). The 1978 revision of the MBCA implemented numerous substantive improvements to the appraisal process. *See, e.g.*, MBCA § 81(b) (1978) (instituting requirements for advance notice by the corporation to their shareholders of their right to dissent); MBCA § 81(d) (1978) (instituting guidelines for how to dissent).

[14] The Model Act requires the corporation, "as soon as the proposed corporate action is taken," to pay each dissenter the amount the corporation estimates to be the fair value of the shares, plus interest. MBCA § 81(f)(3) (1978). By decreasing the amount in dispute, the Model Act seeks to encourage settlement. By contrast, with the exception of New York, large commercial jurisdictions such as California and Delaware do not provide for early payment of the undisputed amount. *Compare* N.Y. BUS. CORP. LAW

and shifted the cost of the process onto the corporation, absent unusual circumstances.[15] Many states followed the Model Act's lead.[16] Despite these efforts to make appraisal a viable shareholder remedy, the paucity of appraisal cases—even in Model

---

§ 623(g) (McKinney 1986) (requiring advance payment by the corporation of 80% of the corporation's offer, to those shareholders who submitted certificates to the corporation, within 15 days of the expiration of the time for filing notices of dissent) *with* CAL. CORP. CODE § 1303(b) (West 1990) (requiring payment of the agreed-upon fair market value of the dissenter's shares, within 30 days *after* an agreement is reached); DEL. CODE ANN. tit. 8, § 262(d)–(f) (1991) (requiring only that the resulting or surviving corporation notify qualified stockholders of their rights to appraisal).

[15] MBCA § 81(i) (1978). Ordinarily, the corporation assumes the costs of the appraisal proceeding, including the reasonable compensation and expenses of appraisers appointed by the court. If, however, the dissenting shareholder's demand for additional payment was made arbitrarily, vexatiously or not in good faith, the court may assess some or all of the costs against the dissenting shareholders. MBCA § 13.31(a) (1984). The court, however, may assess as it deems equitable the fees and expenses of counsel and experts for the parties against the corporation if the corporation fails to substantially comply with the appraisal provisions or against either party if that party acts arbitrarily, vexatiously or not in good faith. MBCA § 13.31(b). Finally, if one dissenter's counsel was a substantial benefit to other dissenters, the court may award attorneys' fees from the award of all dissenters who so benefited. MBCA § 13.31(c).

In contrast to the Model Act, California specifically requires only that the corporation pay all costs of an action where the eventual appraised value of the shares exceeds 125% of the corporation's valuation of the shares. CAL. CORP. CODE § 1305(e) (West 1990). In all other cases, the statute directs the court to assess or apportion costs, including both appraisal costs and counsel and court fees, as it considers equitable. *Id.*

The New York statute states that each party "shall bear its own costs and expenses," unless the court chooses otherwise. N.Y. BUS. CORP. LAW § 623(h)(7) (McKinney 1986). Additionally, New York allows the court, in its discretion, to assess any or all costs and fees, both procedural and legal, against the corporation if the fair value of the dissenters' shares "materially exceeds the amount which the corporation offered to pay." *Id.*

The Delaware statute specifies as follows: (1) the court may apportion costs as it deems equitable; and (2) upon application of a shareholder, the court may charge attorneys' fees against all benefiting shares on a pro rata basis. DEL. CODE ANN. tit. 8, § 262(j) (1991).

[16] Twenty-nine jurisdictions have adopted provisions substantially equivalent to §§ 13.25 (requiring prepayment) and 13.31 (assessing costs against the corporation) of the Model Act. *See* ARK. CODE ANN. §§ 4-27-1325, -1331 (Michie 1991); COLO. REV. STAT. § 7-4-124(6), (9) (1986); FLA. STAT. ANN. ch. 607.1320 (Harrison Supp. 1992); GA. CODE ANN. §§ 14-2-1325, -1331 (1989); HAW. REV. STAT. § 415-81(f), (i) (1985); IDAHO CODE § 30-1-81(f), (i) (1980); ILL. ANN. STAT. ch. 805, para. 11.70(c), (i) (Smith-Hurd 1993); IND. CODE ANN. §§ 23-1-44-15, -20 (Burns 1989); IOWA CODE ANN. §§ 490.1325, .1331 (West 1991 & Supp. 1994); KY. REV. STAT. ANN. §§ 271B.13-250, -310 (Baldwin 1988); MICH. STAT. ANN. §§ 21.200(769), (774) (Callaghan 1990 & Supp. 1994); MINN. STAT. ANN. § 302A.473(5), (8) (West 1985 & Supp. 1994); MISS. CODE ANN. §§ 79-4-13.25, .31 (1989); MONT. CODE ANN. §§ 35-1-834, -839 (1993); NEV. REV. STAT. §§ 78.497, .502 (1993); N.H. REV. STAT. ANN. §§ 293-A:13.25, :13.31 (Supp. 1993); N.C. GEN. STAT. §§ 55-13-25, -31 (1990); OR. REV. STAT. §§ 60.577, .594 (1993); 15 PA. CONS. STAT. ANN. §§ 1577, 1580 (Pamph. 1994); S.C. CODE ANN. §§ 33-13-250, -310 (Law. Co-op 1990); S.D. CODIFIED LAWS ANN. §§ 47-6-46, -49 (1991); TENN. CODE ANN. §§ 48-23-206, -302 (1988); UTAH CODE ANN. §§ 16-10a-1325, -1331 (1994); VT. STAT. ANN. tit. IIA, §§ 13.25, 13.31 (1993); VA. CODE ANN. §§ 13.1-737, -741 (Michie 1993); WASH. REV. CODE ANN. §§ 23B.13.250, .310 (West Supp. 1994); WIS. STAT. ANN. §§ 180.1325, .1331 (West 1992); WYO. STAT.

Act jurisdictions—persists.[17] Thus, while reformation of the appraisal remedy's complex and expensive procedures may have provided a more workable remedy, its work remained undefined.

The second scholarly assault shifted its focus. These scholars argued that the remedy was useful even though little used. They argued that although the remedy's original functions were outdated,[18] appraisal rights were useful in policing corporate management's negotiations in appraisal-triggering transactions. These scholars explained that the potential for sufficiently unhappy shareholders to demand appraisal rights serves to deter management from self-dealing.[19]

---

§§ 17-16-1325, -1331 (1989); *cf.* N.Y. BUS. CORP. LAW § 623(g), (i) (McKinney 1986) (providing for prepayment by the corporation of 80% of its estimate of fair value).

[17] There is no way to document the number of appraisal settlements. In the 16 years since the substantive revision of the MBCA's appraisal procedures in 1978, however, only 25 appraisal cases have been litigated in the 31 jurisdictions which have adopted a prepayment provision substantially equivalent to § 13.25 of the Model Act. *See* Foy v. Klapmeier, 992 F.2d 774 (8th Cir. 1993); Foy v. Klapmeier, No. 3-90 CIV 292, 1991 U.S. Dist. LEXIS 21071 (D. Minn. Feb. 21, 1991); Summers Hardware & Supply Co. v. Brockman, No. CIV-2-92-128, 1992 U.S. Dist. LEXIS 21378 (E.D. Tenn. Dec. 11, 1992); Santa's Workshop v. A.B. Hirschfeld Press, Inc., 851 P.2d 264 (Colo. Ct. App. 1993); Breniman v. Agricultural Consultants, Inc., 829 P.2d 493 (Colo. Ct. App. 1992); Egret Energy Corp. v. Peierls, 796 P.2d 25 (Colo. Ct. App. 1990); Walter S. Cheesman Realty Co. v. Moore, 770 P.2d 1308 (Colo. Ct. App. 1988); Pioneer Bancorp., Inc. v. Waters, 765 P.2d 597 (Colo. Ct. App. 1988); Waters v. Double L, Inc., 769 P.2d 582 (Idaho 1989); Institutional Equip. & Interiors, Inc. v. Hughes, 562 N.E.2d 662 (Ill. Ct. App. 1993); Independence Tube Corp. v. Levine, 535 N.E.2d 927 (Ill. Ct. App. 1988); Sieg Co. v. Kelly, 512 N.W.2d 275 (Iowa 1984); Whetstone v. Hossfeld Mfg. Co., 457 N.W.2d 380 (Minn. 1990); Spinnaker Software Corp. v. Nicholson, 495 N.W.2d 441 (Minn. Ct. App. 1993); MT Properties, Inc. v. CMC Real Estate Corp., 481 N.W.2d 383 (Minn. Ct. App. 1992); National Computer Systems, Inc. v. Bardonaro, No. C9-89-1370, 1990 Minn. App. LEXIS 198 (Minn. Ct. App. Feb. 20, 1990) (unpublished opinion); Rigel Corp. v. Cutchall, 511 N.W.2d 519 (Neb. 1994); *In re* Mohasco Corp., 591 N.Y.S.2d 399 (N.Y. App. Div. 1992); Cawley v. SCM Corp., 530 N.E.2d 1264 (N.Y. Ct. App. 1988); Chrome Data Systems, Inc. v. Stringer, 820 P.2d 831 (Or. Ct. App. 1991); Columbia Management Co. v. Wyss, 765 P.2d 207 (Or. Ct. App. 1991); Genesco, Inc. v. Scolaro, 871 S.W.2d 487 (Tenn. Ct. App. 1993); Wachtel v. Shoney's, Inc., 830 S.W.2d 905 (Tenn. Ct. App. 1991); Eastern Tennessee Transp., Inc. v. Ketron, No. 295, 1991 Tenn. App. LEXIS 167 (Tenn. Ct. App. Mar. 7, 1991); Robblee v. Robblee, 841 P.2d 1289 (Wash. Ct. App. 1992).

[18] *See* Daniel R. Fischel, *The Appraisal Remedy in Corporate Law*, 1983 AM. B. FOUND. RES. J. 875, 879, 901–02 (1983) (suggesting that the appraisal remedy today is effectively "an implied contractual term that increases the value of all shares" by discouraging opportunistic behavior by parties); *see* Kanda & Levmore, *supra* note 2, at 434 (noting that the conventional view of the appraisal remedy's purposes "hardly needs another dismembering").

[19] *See* Kanda & Levmore, *supra* note 2, at 433 (asserting that appraisal should allow shareholders to uncover managerial misbehavior); Seligman, *supra* note 3, at 841 (arguing that appraisal should be utilized to discourage self-dealing by management); *cf.* Fischel, *supra* note 18, at 876 (positing that appraisal serves as protection "in situations where certain groups are more likely to attempt to appropriate wealth from other groups than to maximize the value of the firm").

While both attacks on the appraisal remedy have provided insight, the reasons for the under-utilization of the remedy have, to date, not been identified. The Achilles heel in the appraisal literature has been its myopic search within the remedy itself. Considering the remedy in isolation, the extant literature has failed to resolve not only why the remedy is little used, but also what the proper function of the remedy is. Without such an understanding to guide the revisions of the remedy, reformation has alternated between retreat and expansion.[20]

This Article analyzes the functions ascribed to the appraisal remedy, both historically and currently, and makes recommendations that will provide optimal use of the remedy in the future. Part I describes the evolution of the remedy and emphasizes certain facts, repeatedly omitted or de-emphasized by the appraisal literature, that are critical to an understanding of the remedy's function. This history explains why the function of the remedy was uncertain even at its origin. In fact, scholars have posited three distinct original purposes for the remedy. Part II explains why two of these three purposes are outmoded today, leaving a cash exit at fair value as the only surviving viable purpose. Part III illustrates how the monitoring theory advanced by current scholars is largely a repackaging of the remedy's historical function of providing a cash exit at fair value, leaving little theory about the remedy's modern role or utility. Parts II and III together conclude that the appraisal remedy can serve two functions in modern corporate statutes: to provide a fair cash exit and to monitor conflict-of-interest transactions. Part IV recommends private parties contract to expand the remedy in non-market corporations[21] beyond a solely transaction-oriented

---

[20] *See* Kanda & Levmore, *supra* note 2, at 432 (discussing the evolution of appraisal rights). The authors note that the evolution of appraisal rights has involved an expansion of the triggering transactions but a restriction of the application of appraisal rights through the "market out" exception. *Id.*; *see infra* notes 82–84 and accompanying text (discussing the market out exception). This ambivalence about the appraisal remedy has generated diametrically opposed views. *Compare, e.g.*, Manning, *supra* note 3, at 238 (decrying the remedy's uselessness) *with, e.g.*, Stringer v. Car Data Systems, 841 P.2d 1183, 1184 (Or. 1992) (reasoning that the remedy serves as nothing less than the linchpin that protects minority shareholders). Similarly, some scholars characterize the remedy as a weapon of the minority. *See, e.g.*, PRINCIPLES, *supra* note 9, pt. 7, ch. 4 introductory note. One eminent scholar, however, has characterized the remedy as a disguised tool of the majority. *See* Manning, *supra* note 3, at 227; *see also infra* notes 86–114 and accompanying text (analyzing whether the appraisal remedy is designed to protect majority shareholders).

[21] As used hereafter, "non-market corporation" means any stock corporation all the shares of which are neither listed on a national stock exchange nor regularly quoted in

base, but recommends statutorily limiting the remedy in market corporations to conflict-of-interest transactions. Part IV further recommends increasing the remedy's availability beyond voting shares and provides guidance on determining fair value.

I. EVOLUTION OF THE REMEDY

The appraisal remedy in corporate law evolved from a confluence of contract principles and business exigencies. In the nineteenth century, courts uniformly recognized that the corporate charter gave each shareholder a vested contract right with both the corporation and the state.[22] Each shareholder, as a contracting party, was required to consent to amend the corporate charter. Thus, each shareholder had the power to block any proposed change. As recently as the late nineteenth century, courts protected each shareholder's contract right to veto asset sales, charter amendments and consolidations.[23]

When Justice Story suggested, in the landmark case of *Trustees of Dartmouth College v. Woodward*,[24] that a state could reserve the power to amend its contract with the corporation,[25] the requirement of unanimity for charter amendments began to erode. Beginning with charter amendments in public utilities and railroads, the standard changed from unanimity to majority consent.[26] For private corporations, the notion of majority control seemed equally attractive. Building on Justice Story's theory,

---

an over-the-counter securities market by one or more members of a national or affiliated securities association. This definition is patterned on the definition of close corporations contained in J.A.C. Hetherington & Michael P. Dooley, *Illiquidity and Exploitation: A Proposed Statutory Solution to the Remaining Close Corporation Problems*, 63 VA. L. REV. 1, 52 (1977).

[22] *See* William J. Carney, *Fundamental Corporate Changes, Minority Shareholders, and Business Purposes*, 1980 AM. B. FOUND. RES. J. 69, 78 (1980) (observing that the rule of unanimous consent was based on the conception that the corporate charter is a contract among all the shareholders, as well as between the corporation and the state).

[23] *See, e.g.*, Mason v. Pewabic Mining Co., 133 U.S. 50, 58 (1890) (declaring the injustice of compelling minority shareholders to become members of a new corporation through an asset sale for stock); *see generally* Carney, *supra* note 22, at 79 n.34 (listing decisions granting shareholders injunctive relief against corporations that attempted to alter the corporation without unanimous shareholder consent).

[24] 17 U.S. (4 Wheat.) 518 (1819).

[25] *Id.* at 675, 712 (Story, J., concurring) (noting that a state can, in granting a corporate charter, reserve authority to amend the nature of the charter); *cf.* Carney, *supra* note 22, at 84 (observing that state corporation statutes adopted in the latter 19th century reserved for states the authority to amend corporate charters).

[26] *See* Carney, *supra* note 22, at 85 (observing that, for public utilities, corporate charters often permitted amendments without requiring unanimous shareholder consent). In most cases concerning public corporations, states found a public purpose

private corporations added provisions in their charters authorizing amendments by less than unanimous consent.[27] Such attempts to skirt the unanimity requirement but still remain within the parameters of contract law coincided with the emergence of courts and legislatures more sympathetic to the corporations' perspective. Judges and legislators had begun to consider issues beyond the individual shareholder's contract right, such as the wisdom of allowing one shareholder to block change thought desirable by the majority.[28]

As was the case with charter amendments, the evolution from unanimity to majority control for other major corporate transactions proceeded within the confines of contract law. Mergers, consolidations and sales of all corporate assets originally required unanimous consent.[29] If, however, the corporation was insolvent with no prospect for profit, courts permitted the majority to sell all corporate assets for cash, to wind up the corporation, and to distribute the remaining cash to the shareholders.[30] Such action is consistent with contract law, which permits nonperformance of a contract where the purpose of the agreement has been frustrated.[31] This application of contract law, however,

---

sufficient to warrant a fundamental corporate change where a simple majority of the shareholders believed the merger or consolidation "made sense." *Id.*

[27] *See, e.g.*, Sprague v. Illinois River R.R. Co., 19 Ill. 174 (1857) (affirming the validity of a provision in the corporate charter allowing amendment of the charter upon approval by a majority of the board of directors without requiring unanimous consent of the shareholders). In *Sprague*, Chief Justice Caton asserted that "no instance can be found where the unanimous consent of all the shareholders of the corporation is made necessary to an acceptance of such amendments . . . ." *Id.* at 180; *cf.* VICTOR MORAWETZ, A TREATISE ON THE LAW OF PRIVATE CORPORATIONS § 951 (2d ed. 1886) (recognizing that a corporation can, through its corporate charter, provide for consolidation upon a majority approval, against the desires of minority shareholders); 5 SEYMOUR D. THOMPSON, COMMENTARIES ON THE LAW OF PRIVATE CORPORATIONS § 6057 (Joseph W. Thompson ed., 2d ed. 1910) (recognizing that majority approval of a consolidation is valid, provided "the charter at the time of the incorporation expressly permits" such a consolidation by the majority).

[28] *See* Carney, *supra* note 22, at 80 & n.38 (noting the injustice of allowing a single shareholder to block a corporate change desired by all other shareholders).

[29] *See* MORAWETZ, *supra* note 27, § 951 (stating that unless the corporate charter provides otherwise, a corporation may consolidate with another corporation only with the consent of all its shareholders); 5 THOMPSON, *supra* note 27, § 6056 (stating that "the consent of every stockholder is absolutely essential to a consolidation," in the absence of specific statutory authority to the contrary). *See generally* Carney, *supra* note 22, at 77–81 (observing that at common law, dissenting shareholders historically could block any fundamental change in a corporation); Elliott J. Weiss, *The Law of Take-Out Mergers: A Historical Perspective*, 56 N.Y.U. L. REV. 624, 627 (1981) (lamenting that under the standards described by Carney, "[m]ajority shareholders were powerless to overcome a dissenter's opposition, except by buying him off").

[30] Carney, *supra* note 22, at 87.

[31] *Id.*

quickly gave way to business needs. By the late nineteenth century, courts permitted such asset sales by the majority for marketable stock[32]—and later less marketable stock[33]—when no cash buyers were available. Thus, when the seller dissolved and distributed to its shareholders the buyer's stock, the seller's shareholders became shareholders of the buyer. In short, when the corporation was insolvent, majority rule replaced vested rights and stock replaced cash as acceptable consideration. In effect, these two changes forced minority shareholders to become shareholders of a new corporation.

Courts next relaxed the requirement of insolvency to allow majority rule in situations in which the corporation was not yet insolvent.[34] Soon after, a corporation suffering merely poor prospects could effectuate these asset sales.[35] Eventually, courts permitted what has become today's norm: a corporate sale of all assets for either cash or stock whenever the majority determines that such a sale is in the best interests of the corporation.[36]

With the majority able to effectuate both charter amendments and asset sales for stock or cash, the requirements that mergers and consolidations could be effectuated only for stock consideration[37] and only by unanimous consent[38] became ripe for change. It had not gone unnoticed that the permitted asset sales for cash

---

[32] *Id.* at 88. Courts justified an asset sale for marketable securities by considering the immediate convertibility of the securities to cash. *Id.*; *see, e.g.*, Geddes v. Anaconda Copper Mining Co., 254 U.S. 590, 597–98 (1921) (validating an asset sale for stock, where the shareholder could immediately "convert [the stock] into an adequate cash consideration for what his holdings were in the corporate property").

[33] Carney, *supra* note 22, at 88.

[34] *Id.* at 89 (chronicling the expansion of lawful sales, authorized by only a majority of shareholders, to situations where the business demonstrated no prospect for profits even though it may not yet be insolvent); *cf.* American Elementary Elec. Co. v. Normandy, 46 App. D.C. 329, 340–41 (1917) (citation omitted) (declaring that when faced with an "unprofitable and hopeless enterprise," a private corporation "'should cease to transact business as soon as, in the exercise of sound judgment, it is found that it cannot be prudently continued'"); Smith v. Stone, 128 P. 612, 617 (Wyo. 1912) (quoting NOYES ON INTERCORPORATE RELATIONS § 111) ("'When the further prosecution of the business of the corporation would be unprofitable, it is the duty, as well as the right, of the majority to dispose of its property and take action towards the liquidation of its affairs'").

[35] Carney, *supra* note 22, at 89.

[36] *Id.* at 89–90; *see* Bowditch v. Jackson Co., 82 A. 1014, 1017 (N.H. 1912), *appeal dismissed*, 239 U.S. 627 (1915) (suggesting that "[i]f the majority may sell to prevent greater losses, why may they not also sell for greater gains?" and finding "no substantial difference between the two cases . . . .").

[37] *See* Weiss, *supra* note 29, at 630 (observing that "[e]arly merger and consolidation statutes . . . generally required that the shareholders . . . receive shares in the surviving corporation").

[38] *See supra* note 29.

or stock followed by the dissolution of the seller were de facto mergers.[39] Considering the substance of those transactions, rather than just their form, courts put asset sales and mergers on parallel tracks, eventually permitting both to be carried out by a majority vote for stock or cash.[40]

Once asset sales and mergers could be effectuated for stock or cash consideration, there remained only one gap in the creation of appraisal rights. This missing link was to provide dissatisfied minority shareholders a cash option in stock transactions. Minority shareholders initially secured this cash option through litigation. While shareholders often sued for injunctive relief, courts usually instead awarded shareholders the fair value of their stock,[41] reasoning that their shares had been effectively converted.[42] Such litigation enabled minority shareholders to escape the choice of either forced membership in a new corporation or the acceptance of a pro rata share in cash of the transaction's proceeds;[43] instead, shareholders received the appraised fair value of their shares.

Litigation, however, proved unsatisfactory both for the corporation and for minority shareholders. The corporation feared the possibility of an injunction, and the shareholders disliked the expensive and risky process of judicially resolving their claims.[44] Both the corporation and the shareholders bettered their respective positions when they could settle on a cash payment and avoid judicial intervention, particularly given the likelihood that such litigation would require a cash settlement. Eventually, legislatures began to follow the courts' lead by enacting appraisal statutes.[45] While these statutes sought to ensure that dissenting shareholders received "value" for their shares,[46] courts interpret-

---

[39] *See, e.g.*, Carney, *supra* note 22, at 89 (observing that an asset sale to a new enterprise controlled by the majority shareholder was effectively a consolidation).

[40] *See id.* at 93 (interpreting judicial developments as leading toward a convergence in the treatment of asset sales and consolidation so that both could be effectuated by majority approval).

[41] *See id.* at 92–93 (observing that in the latter 19th century, courts began to question the propriety of granting injunctions to secure minority rights in consolidations and asset sales).

[42] Norman D. Lattin, *Remedies of Dissenting Stockholders Under Appraisal Statutes*, 45 HARV. L. REV. 233, 236 (1931).

[43] *See id.* (noting that prior to the enactment of appraisal statutes, courts examining asset sales, mergers, or consolidations "from a realistic standpoint" did not restrict a shareholder's consideration to a proportionate share of the proceeds).

[44] *See id.* at 236–37 (concluding that both the majority and the minority feared the risks of litigation).

[45] Carney, *supra* note 22, at 94 & n.102.

[46] *See* Lattin, *supra* note 42, at 243 (discussing contemporary standards for determin-

ing these statutes consistently awarded the stock's fair value in cash, rather than simply a cash pro rata share of the transaction's proceeds.[47] By 1927, at least twenty states had adopted appraisal

---

ing the cash value to which dissenting stockholders were entitled in exchange for their shares). Lattin singles out Ohio's treatment of the value to which dissenting shareholders were entitled as "illustrative" of a "usual provision." *Id.* at 243 & n.21. The Ohio statute awarded a minority stockholder "the *fair cash value* of his shares as of the day before the vote was taken authorizing [the] action, excluding from such fair cash value any appreciation or depreciation." OHIO GEN. CODE ANN. § 8623-72 (Page Perm. Supp. 1926–1935) (enacted 1931) (emphasis added).

Numerous other contemporary state statutes discussed a "value" to which dissenting minority stockholders were entitled as compensation for their shares. These statutory descriptions were neither uniform nor clear. At a minimum, these statutes entitled shareholders to "payment" for their shares. VT. LAWS § 5847 (1933) (enacted 1909). Other statutes required the corporation to pay qualified dissenting shareholders the "value" of their shares. CONN. GEN. STAT. § 3524 (1918); DEL. CODE REV. § 2093 (1935) (enacted 1929); IND. CODE ANN. § 4656.2 (Burns 1929); ME. REV. STAT. ch. 56, § 66 (1930); N.Y. STOCK CORP. LAW § 21 (Consol. 1930) (amended 1924); N.C. CUM. STAT. § 1224(c) (Michie 1925); MASS. GEN. LAWS ch. 156, § 46 (1921) (enacted 1903); S.C. CODE § 7759 (1932) (enacted 1925). Still other statutes provided for compensation to the dissenting shareholder for the "fair value" or the "fair cash value" of their shares. ALA. CIV. CODE § 1743 (1923); ARK. STAT. § 2225 (Pope 1931) (amending ARK. CODE ANN. § 1738n (Crawford & Moses Supp. 1927), which allowed for an appraisal to determine if a party's assertion regarding the stock's value equalled the true market value of the stock); FLA. COMP. GEN. LAWS § 6564 (1927) (enacted 1925); ILL. ANN. STAT. ch. 32, § 73 (Smith-Hurd 1927) (amended 1921); KY. STAT. ANN. § 558 (1922) (enacted 1893); MD. CODE ANN. art. 23, §§ 35–36 (Bagby 1924) (amended 1916); NEV. COMP. LAWS § 1640 (Hillyer 1929) (enacted 1925); N.H. LAWS ch. 225, § 54 (1926) (enacted 1919); N.J. COMPILED STAT. 1709–1910 Corporations § 108 (amended 1902); 1931–1932 R.I. PUB. LAWS ch. 1941, § 5 (amending R.I. GEN. LAWS tit. 24, § 56 (1923)); VA. CODE ANN. § 3822 (1924) (amended 1922). Still other statutes specified payment to the dissenting shareholders of the market value of their shares. PA. STAT. ANN. tit. 15, § 425 (1930) (amended 1924); TENN. CODE ANN. § 3752 (Michie 1932) (enacted 1929) (amending TENN. CODE § 2053a-10 (1917) (enacted 1907), which permitted dissenting stockholders a judicial determination of the value of their shares).

[47] *See* Jackson Co. v. Gardiner Inv. Co., 200 F. 113, 116 (1st Cir. 1912) (holding that minority stockholders are entitled to the "intrinsic value" of their shares, and should not be compelled to sacrifice "that to which they were lawfully entitled" for an arbitrarily fixed cash amount determined by the parties to the transaction); *see also* Cole v. Wells, 113 N.E. 189, 191 (Mass. 1916) (stating that a "dissatisfied" stockholder may either acquiesce to a transaction or "accept the fair value of his stock and retire from the corporation," which value is not to be determined based on the market price of the stock).

In other decisions, courts construed appraisal statutes in favor of dissenting stockholders, requiring the corporation to treat them equitably. *See* Manning v. Brandon Corp., 161 S.E. 405, 408 (S.C. 1931) (declaring that the South Carolina dissenting stockholder's rights statute "should be construed liberally in favor of the stockholder"). In some instances, courts went so far as to place the surviving majority stockholders/directors in the role of a trustee or fiduciary for the dissenting stockholder's investment. *See, e.g.*, Jones v. Missouri-Edison Electric Co., 144 F. 765, 770–71 (8th Cir. 1906) (holding in an appraisal case that the relationship between a minority stockholder and the corporation is one of "trust and confidence," and placing the directors of the corporation in a "fiduciary relation" with respect to such a stockholder, as "[t]he corporation holds its property in trust for its stockholders").

statutes.[48] Today, all jurisdictions include appraisal provisions in their corporate statutes.[49]

The evolution of appraisal rights is important for several reasons. By the time appraisal rights emerged, shareholders no longer enjoyed a vested contract right which could be breached, and majority rule was not perceived as wrongful. As a result, the remedy was not designed as a damage action for breach of contract. Instead of damages, shareholders were awarded the fair value of their stock.[50]

Second, the history helps explain why the corporate statutes differ so widely as to which transactions trigger the remedy.[51] The common feature of charter amendments, asset sales and mergers is that they all once required unanimous shareholder approval. Once that common link was severed, no unifying aspect remained to distinguish transactions that should offer appraisal rights from those that should not.[52] Today, all corporate statutes recognize appraisal rights for at least some mergers.[53] Almost all statutes also afford appraisal rights for short-form mergers[54] and sales of substantially all assets.[55] Many statutes

---

[48] *See supra* note 46 (listing early statutory appraisal rights provisions).

[49] *See infra* notes 53–58 and accompanying text.

[50] *See supra* notes 46–47 (discussing early appraisal rights provisions and judicial applications of these provisions). The appraisal remedy was thus analogous to an eminent domain proceeding. *See* Carney, *supra* note 22, at 70 & n.5.

[51] *See infra* note 52 (listing the number of jurisdictions granting appraisal rights for each type of transaction).

[52] *See* Joseph L. Weiner, *Payment of Dissenting Stockholders*, 27 COLUM. L. REV. 547, 548 (1927) (noting that state laws in force at the time granted appraisal rights in a myriad of transactions, including changes in corporate purposes, changes in share preferences, corporate consolidations, sales of assets for securities pursuant to a voluntary dissolution, and the issuance of a company's stock to its employees).

Currently, 45 states grant appraisal rights for mergers requiring shareholder authorization. MBCA ANN., *supra* note 2, § 13.02 statutory comparison, at 1371 (Supp. 1993). All jurisdictions, however, grant appraisal rights in at least some merger situations, and the 35 jurisdictions recognizing compulsory share exchanges grant appraisal rights in such transactions. *Id.* Appraisal rights are also recognized in 50 jurisdictions pursuant to a sale or exchange of assets. PRINCIPLES, *supra* note 9, pt. 7, ch. 4 introductory note, reporter's note 1. Approval of certain amendments to the articles of incorporation triggers appraisal rights in 33 jurisdictions. MBCA ANN., *supra* note 2, § 13.02 statutory comparison, at 1371 (Supp. 1993). Furthermore, five jurisdictions allow for appraisal rights after certain control shareholder acquisitions. *Id.* at 1372.1. In 24 jurisdictions, however, these appraisal rights are denied if the shares are either listed on a national exchange or held by a threshold number of shareholders. *Id.* at 1372.

[53] MBCA ANN., *supra* note 2, § 13.02 statutory comparison 1(a) (Supp. 1993).

[54] *See* Principles, *supra* note 9, pt. 7, ch. 4 introductory note, reporter's note 1 (listing 39 jurisdictions as affording appraisal rights for short-form mergers).

[55] *See* MBCA ANN., *supra* note 2, § 13.02 statutory comparison 1(c) (Supp. 1993) (noting 46 jurisdictions as granting appraisal rights for sales of substantially all assets).

also provide an appraisal remedy for certain charter amendments[56] and share exchanges,[57] and a few states provide the remedy for a variety of additional transactions.[58]

Most importantly, however, the history of the appraisal remedy is crucial to explaining why the appraisal statutes vary so greatly. While the legislatures could have overruled the rights that the courts had created, legislatures instead chose to codify these rights. Legislatures evidently found the appraisal remedy attractive; unfortunately, they explained neither the reasons for this attraction nor their intended purpose for this new remedy. The result has been a remedy built on quicksand, with shifting premises and purposes.

For example, some commentators look at the remedy's history and conclude that appraisal rights were designed to compensate the minority for the loss of their veto power: while shareholders could no longer veto a transaction, they could "veto" their continuing involvement in a "fundamentally different" corporation by requiring the corporation to cash them out.[59] Others interpret the remedy's history as creating a cash exit,[60] citing either the original requirement that mergers be effectuated only for stock consideration,[61] the common law prohibition against forcing shareholders to become shareholders of another corporation,[62] or the

---

[56] *See id.*, § 13.02 statutory comparison 1(d) (listing 33 jurisdictions as granting appraisal rights for charter amendments).

[57] *See id.*, § 13.02 statutory comparison 1(b) (listing 35 jurisdictions as providing appraisal rights in compulsory share exchanges).

[58] *See, e.g.*, OHIO REV. CODE ANN. § 1701.84(D) (Anderson 1992) (entitling dissenting shareholders of the acquiring corporation to appraisal in a majority share acquisition); OKLA. STAT. ANN. tit. 18, § 1091(B)(3) (West Supp. 1994) (permitting appraisal rights pursuant to a distribution of assets in kind); 15 PA. CONS. STAT. ANN. § 1931 (Pamph. 1994) (allowing appraisal in certain share acquisitions in which one person acquires control of the corporation).

[59] Kanda & Levmore, *supra* note 2, at 429, 434.

[60] *See* EISENBERG, *supra* note 11, at 78. Although Eisenberg phrases the goal as one of fairness, his conclusion that fairness requires a cash exit when the corporation is fundamentally changed demonstrates his use of the cash exit theory. *Id.*

[61] A merger for cash consideration was contrary to the "contract" view of stock ownership and historically was not allowed. *See* Weiss, *supra* note 29, at 626–29.

[62] Yanow v. Teal Indus., Inc., 422 A.2d 311, 316 n.5 (Conn. 1979); *see* Irving J. Levy, *Rights of Dissenting Shareholders to Appraisal and Payment*, 15 CORNELL L.Q. 420, 421 (1930) (observing that appraisal rights allowing dissenters "the right to receive the cash value of their stock" arose as a statutory compromise to satisfy both majority and minority shareholders, upon the realization that it was unfair to force a minority to continue in a fundamentally different enterprise). Compounding the confusion about the remedy's original purpose is the fact that among commentators who agree that the remedy's purpose is to protect shareholders from coerced participation in new or fundamentally different corporations, some view that purpose as a veto issue while others see it as a cash exit issue. *See* Fischel, *supra* note 18, at 875–78 (outlining

cash exit compromise struck by the majority and minority shareholders.[63] One eminent scholar has a third view of the remedy: Dean Bayless Manning argues that the effect of the appraisal remedy, if not its purpose, was to protect the majority from the tyranny of the minority.[64] This theory posits that the remedy decreased both the minority's incentive to seek injunctions and the courts' willingness to issue them.[65]

With Delphic ambiguity and without explaining whether the appraisal remedy was needed to compensate for the loss of veto power, to effectuate a cash exit, or to protect the majority, the legislatures of all fifty states have incorporated the remedy into their corporate statutes. While the original purpose of the appraisal remedy cannot be further clarified, Part II will examine whether any of the three suggested original purposes for the appraisal remedy remain viable today.

II. Analyses of the Asserted Purposes of the Remedy

A. *Compensation for Loss of Veto*

Virtually none of the modern corporate statutes include a unanimity requirement for any transaction.[66] Unanimity is simply an option the statutes allow.[67] Due to the difficulty of obtaining unanimous agreement, that option is common only when the stock is held by a small number of shareholders.

Instead, modern corporate statutes require a majority,[68] or at most, a supermajority,[69] shareholder vote. As long as the major-

---

different views of the function of appraisal rights); Folk, *supra* note 1, § 251.4, § 251:29 (observing that Delaware law early on deprived a shareholder of his veto power, replacing it with a right to cash). Folk views the cash as "compensation" for the shareholder's lost veto. *Id.*

[63] Lattin, *supra* note 42, at 236–37.

[64] Manning, *supra* note 3, at 227–29.

[65] *Id.*

[66] One notable exception is the California statute. Under Cal. Corp. Code §§ 1101, 1101.1 (West 1990), an acquiror or its parent corporation owning more than 50% (but less than 90%) of the voting stock of a target corporation that wishes to merge must obtain the unanimous consent of the target's shareholders if the merger consideration is anything other than the acquiror's (or its parent's) nonredeemable common stock. The statute provides some exceptions to its unanimity rule.

[67] *See, e.g.*, Del. Code Ann. tit. 8, § 251 (1991) (requiring a majority vote for mergers but permitting the corporate charter to specify a higher percentage).

[68] *See, e.g.*, Del. Code Ann. tit. 8, § 251(c) (1991); MBCA § 11.03(e) (1984).

[69] *See, e.g.*, Va. Code Ann. § 13.1-718E (Michie 1993) (requiring a two-thirds vote to approve a merger unless the board of directors requires a greater or lesser percentage, which must always be at least a majority).

ity fulfills its fiduciary duties and statutory obligations, it enjoys almost unlimited power.[70] This shift to a relatively unqualified system of majority rule suggests that current corporate statutes do not perceive their failure to require a unanimous vote as a wrong for which the appraisal remedy must offer minority shareholders compensation. To the contrary, many legislatures, courts, and commentators have discredited minority veto power as an infringement on corporate democracy.[71] Therefore, if compensation for the loss of the veto were the only justification for the appraisal remedy, the remedy should have been abolished long ago. While utilizing the appraisal remedy as a substitute for the veto power clearly seems anachronistic in the modern corporate legal environment, vestiges of this original view nevertheless persist.[72]

B. *Cash Exit at Fair Value*

As explained above,[73] the requirement that mergers could be effectuated only for stock consideration eventually gave way to cash consideration. By permitting transactions for cash—and later through appraisal rights requiring a cash option for shareholders—corporations afforded their shareholders a cash alternative to forced membership in a different entity. Moreover, since appraisal statutes required the corporation to pay the fair value of stock, rather than simply a shareholder's pro rata share of the proceeds of the transaction,[74] the appraisal remedy alternative provided a cash exit at fair value. While the appraisal literature has consistently recognized that the remedy was designed to create a cash exit, the requirement that the exit be for fair value has largely been ignored.[75]

---

[70] For example, the majority can cash out minority shareholders without a valid corporate purpose. *See* Weinberger v. UOP, Inc., 457 A.2d 701 (Del. 1983). The majority can amend the corporation's governance rules and shareholder rights through charter and bylaw amendments. *See, e.g.*, MBCA §§ 10.03, 10.20 (1984). The majority can mortgage all of the corporate assets. *See, e.g.*, MBCA §§ 12.01–12.02. Unless limited by the corporate charter's purpose clause, the majority can undertake new ventures. *See, e.g.*, MBCA § 2.02(b)(2)(i) (allowing corporate charter to include provisions regarding the purposes for which the corporation is organized).

[71] Yanow v. Teal Indus., Inc., 422 A.2d 311, 317 n.5 (Conn. 1979).

[72] *See infra* notes 250–252 and accompanying text (discussing those statutes which, consistent with the veto theory, limit the remedy to voting stock).

[73] *See supra* note 40 and accompanying text.

[74] *See supra* note 47 (discussing judicial interpretations of early appraisal statutes).

[75] *See infra* text accompanying note 125 (arguing that contemporary scholars believe

In modern corporate statutes, appraisal rights continue to serve the fair cash exit function in certain transactions. While the market duplicates the remedy's cash exit component for market corporations, there is no such mechanism for non-market corporations.[76] The transactions which trigger appraisal rights are, however, few in number and consist of only major changes to the corporation.[77] In addition, the flexibility of contemporary corporate statutes often affords the majority the option of structuring a transaction to avoid appraisal rights.[78] As a result, appraisal-triggering transactions are an infrequent occurrence.

Curiously, while the total number of appraisal cases is relatively small,[79] nearly all these cases were brought by shareholders in market corporations.[80] Evidently, these shareholders perceived some defect in the cash exit provided by the market; undoubtedly, they questioned whether the value set by the market was a fair value. In all states, the appraisal remedy's fair value component has overshadowed its cash exit component; all modern appraisal statutes grant the remedy largely without re-

---

their monitoring theory is novel because they fail to recognize that appraisal has always been intended to award the fair value of stock).

[76] *See infra* notes 157–175 and accompanying text (discussing the contractual nature of minority exit rights in non-market corporations).

[77] Prevalent triggers of appraisal rights include statutory mergers, short-form mergers, asset sales not in the regular course of business, amendments to corporate articles, and compulsory share exchanges. PRINCIPLES, *supra* note 9, pt. 7, ch. 4 introductory note, reporter's note 1; MBCA ANN., *supra* note 2, § 13.02 statutory comparison (Supp. 1993). Less common triggering transactions include "control share" acquisitions, majority share acquisitions, distributions of assets in kind, and transactions involving 10% shareholders. PRINCIPLES, *supra* note 9, pt. 7, ch. 4 introductory note, reporter's note 3; MBCA ANN., *supra* note 2, § 13.02 statutory comparison (Supp. 1992 & Supp. 1993).

[78] F. HODGE O'NEAL & ROBERT B. THOMPSON, O'NEAL'S OPPRESSION OF MINORITY SHAREHOLDERS § 5:29, at 173 (2d ed. 1985) [hereinafter O'NEAL'S OPPRESSION]. For example, similar results can be achieved by designing a transaction as a merger between two parties, as a merger of a wholly owned subsidiary and a third party, or as a cash purchase of all of another corporation's assets. Depending on the state and the structure of the transaction, some of these combinations will permit the appraisal remedy, while others will deny it.

[79] Professor Seligman observed that from 1972 to 1981, there were 19 reported state court decisions involving appraisal rights. *See* Seligman, *supra* note 3, at 829 n.3 (listing reported state appraisal rights cases during the relevant period); Thompson, Work in Progress (finding 103 decisions involving appraisal rights from 1984 to 1994). There are 84 transactions involved in the 103 decisions cited by Professor Thompson. *Id.*

[80] Of the 84 transactions cited by Professor Thompson, four did not reveal enough information to be classified as a market or non-market corporation. Thompson, *supra* note 79. Of the remaining 80 transactions, 14 involved non-market corporations, and 66 involved market corporations. *Id.*

gard to whether the transaction is for cash or non-cash consideration.[81]

While all corporate statutes award the appraisal remedy regardless of the type of consideration received, some corporate statutes create an exception when there is an available cash exit through the market. The premise of the "market out" exception, adopted by twenty-four states,[82] is that the market or a large number of shareholders creates a cash exit at fair value, thereby making the appraisal remedy superfluous. Some statutes make the exception applicable as long as the stock surrendered is widely traded or held.[83] The market-out exception in other statutes is applicable only when both the stock surrendered and the stock received as consideration are widely traded or held.[84]

---

[81] For example, the Model Act grants appraisal rights for a wide variety of transactions without regard to the type of consideration offered. MBCA § 13.02 (1984). The only exception is that appraisal rights for sales of all assets are denied when the sale is for cash followed by a liquidation. MBCA § 13.02(a)(3). This exception, however, is linked more to the liquidation feature than to the cash feature. *Cf. infra* note 84 (observing that Delaware's market-out exception, which denies appraisal rights, is dependent in part on the type of consideration received in the appraisal-triggering transaction).

[82] The 24 states that have adopted a market-out exception are Alabama, Alaska, Arizona, California, Delaware, Florida, Georgia, Indiana, Kansas, Louisiana, Maine, Maryland, Michigan, Nevada, New Jersey, Oklahoma, Oregon, Pennsylvania, Rhode Island, Tennessee, Texas, Utah, Virginia, and Wisconsin. Not surprisingly, the majority of trading corporations are incorporated in jurisdictions with a market-out exception. *See* Seligman, *supra* note 3, at 835 n.21 (noting that 73% of the corporations listed on the New York Stock Exchange are incorporated in states with a market-out exception). On the other hand, 26 states, the MBCA, and the PRINCIPLES do not have any market-out exception. While the MBCA and the PRINCIPLES articulate many objections to a market-out exception, their primary concern is that the exit provided by the market does not necessarily represent a stock's fair value. *See infra* notes 216–221 and accompanying text.

[83] *See, e.g.*, IND. CODE ANN. § 23-1-44-8(b) (Burns 1989); MD. CODE ANN., CORPS. & ASS'NS § 3-202(c) (Supp. 1993).

[84] For example, while the Delaware appraisal statute contains a market-out exception, the exception is inapplicable if the stock received as consideration is either not listed on a national exchange or is held by less than 2000 stockholders of record. DEL. CODE ANN. tit. 8, § 262(b) (1991 & Supp. 1992); *see* FOLK, *supra* note 1, § 262.2.3 (discussing exceptions to the Delaware Code's market-out exception). Delaware recently amended its appraisal statute to broaden the market-out exception to include stock represented by widely held or publicly traded depository receipts. 1994 Del. Laws ch. 262 (S.B. 324) (1994). Interestingly, the Delaware market-out exception is also inapplicable if the shareholders are to receive anything other than stock as consideration for their shares. DEL. CODE ANN. tit. 8, § 262(b)(2) (1991 & Supp. 1992); *see* Folk, *supra* note 1, § 262:15 (observing that any required consideration other than stock plus cash for fractional shares negates the market-out exception, restoring appraisal rights). In so providing for appraisal rights where the consideration is cash, the Delaware statute departs from the historical roots of the appraisal remedy, which was designed to create a cash exit in non-cash transactions. Professor Fischel speculates that this departure is an implicit recognition that minority shareholders face the greatest danger

In sum, while the veto function of the appraisal remedy is nonexistent in the modern corporate statute, the cash exit at fair value theory remains viable. For both market and non-market companies, the remedy continues to serve this function. The market-out exception to the remedy, adopted by almost half of the states,[85] is an endorsement of the fair cash exit theory because such statutory exceptions are premised on the belief that the market offers a fair cash exit.

C. *Protection of the Majority*

While the first two theories of the remedy's function are premised on protecting minority shareholders, a third theory posits that the remedy serves to protect the majority by facilitating corporate transactions.[86] Dean Manning argues that unhappy shareholders will be less likely to institute litigation, and courts will be less likely to grant injunctions, if minority shareholders have the right to be paid the fair value of their stock in cash.[87]

Modern corporate law discredits this premise. First, the exercise of appraisal rights does not necessarily facilitate the majority's will because the costs associated with the remedy can deter transactions. For example, the cash drain resulting from a sig-

---

of appropriation of the value of their shares in cash-out mergers. Fischel, *supra* note 18, at 885 & n.45.

An alternative explanation for the stricter treatment of cash mergers may be a lingering mistrust of them, as mergers were originally permitted only for stock consideration. *See supra* notes 32–33 and accompanying text; *cf.* Kanda & Levmore, *supra* note 2, at 447 (intimating that § 262(b)(2) of the Delaware corporate code reflects a goal of inframarginality in the treatment of dissenting target shareholders); *see infra* note 117 (setting forth definition of "inframarginality"). Still another possibility is that the Delaware statute reflects the "coattails" theory, which posits that shareholders want stock, rather than cash, consideration so as to ride their corporation's "coattails" to success in the new venture. *See* Kanda & Levmore, *supra* note 2, at 435–37. The coattails theory not only contradicts the cash exit premise of the appraisal remedy, but it is also unbolstered by any empirical evidence demonstrating harm suffered by shareholders who receive cash consideration. *Id.* at 436–37. For example, the coattails theory fails to explain why shareholders who receive cash are worse off than those who receive stock, fails to address fairness concerns where shareholders in stock-for-stock mergers are offered inadequate consideration for their stock, and fails to explain why the majority of corporate statutes give appraisal rights in stock mergers. *See, e.g.*, MBCA § 13.02(a) (1984).

[85] *See supra* note 82 (listing the 24 states that have adopted a market-out exception); *infra* notes 212–236 and accompanying text (discussing and suggesting reform of the market-out exception).

[86] *See* Manning, *supra* note 3, at 227 (arguing that while the remedy's purpose is debatable, the major effect of appraisal statutes has been to facilitate the majority's will).

[87] *Id.*

nificant appraisal demand may cause a proposed transaction to be aborted.[88] Transactions may also be abandoned if the appraisal demand prevents the transaction from qualifying for either a pooling of interests accounting treatment[89] or a tax-free treatment.[90] The abandonment of transactions because of the effects of appraisal demands is antimajoritarian in that it makes paramount the minority shareholders' interest in having their stock repurchased for cash, sometimes to the detriment of the majority's plans. Thus, contrary to Dean Manning's theory, the exercise of appraisal rights may derail, rather than effectuate, the majority's will.

Furthermore, Dean Manning's assertion that the grant of appraisal rights reduces litigation is based on questionable assumptions. For example, if shareholders prefer to remain shareholders,[91] they must attempt to stop the transaction by securing an injunction. As the standards for injunctive relief are difficult to meet, however, it is improbable that courts will grant injunctions simply because shareholders lack appraisal rights.[92] Rather than deter suits for injunctive relief, appraisal rights for shareholders who could not satisfy the demands for injunctive relief are a costly gift to the minority; unless the appraisal remedy is exchanged for an otherwise meritorious cause of action, the remedy does not substantially reduce litigation.

---

[88] *Id.* at 235.

[89] TED J. FIFLIS, ACCOUNTING ISSUES FOR LAWYERS 458 (4th ed. 1991).

[90] Section 368 and related provisions of the Internal Revenue Code permit a number of different tax-free reorganizations of corporate interests. All are premised, however, on the assumption that there is a continuation of the shareholder's investment in the old corporation in an altered (but unliquidated) form in the new corporation. This investment continuity requirement is enforced by compelling the transferor corporation (and/or its shareholders) to obtain an equity interest in the purchasing corporation. The qualifying equity consideration given to the transferor corporation and/or its shareholders by the purchasing corporation will vary depending on the particular type of reorganization involved. Cash is never a qualifying consideration in any reorganization, however, and the amount of cash permitted in the overall transaction is always limited. Pinellas Ice & Cold Storage Co. v. Commissioner, 287 U.S. 462 (1933); Cortland Specialty Co. v. Commissioner, 60 F.2d 937 (2d Cir. 1932), *cert. denied*, 288 U.S. 599 (1933). In some transactions, the presence of any cash consideration, including cash payments, pursuant to an appraisal remedy, will destroy the tax-free nature of the transaction for all parties. *See, e.g.*, I.R.C. § 368(a)(1)(B); Reeves v. Commissioner, 71 T.C. 727 (1979).

[91] The coattails theory states that shareholders want only stock consideration, as they want to ride their corporation's coattails to success in its new venture. *See supra* note 83.

[92] Among the requirements for an injunction are irreparable injury and an inadequate remedy at law. *See* STEVEN BAICKER-MCKEE ET AL., FEDERAL CIVIL RULES HANDBOOK 1995, at 591 (2d ed. 1995) (discussing the requirements of FED. R. CIV. P. 65). The lack of appraisal rights is insufficient to demonstrate that there is no adequate remedy at law, as plaintiffs could pursue damage actions.

Therefore, whether the appraisal remedy facilitates transactions by deterring litigation depends on whether a rational shareholder with a viable claim will be satisfied with appraisal or instead will seek alternative remedies. The existence of other remedies is a moot issue if the appraisal remedy is exclusive. Overwhelmingly, however, legislatures and courts have made the appraisal remedy non-exclusive. Only two jurisdictions[93] statutorily make the appraisal remedy exclusive, and the highest court in one of these states has nonetheless created an exception.[94]

The limited goal of the appraisal remedy explains the bias against exclusivity: most states have designed the remedy solely to assess whether directors misvalued their corporation's stock.[95]

---

[93] *See, e.g.*, CONN. GEN. STAT. ANN. § 33-373(f) (West 1987); IND. CODE. ANN. § 23-1-44-8(c) (Burns 1989).

[94] Gabhart v. Gabhart, 370 N.E.2d 345 (Ind. 1977) (allowing dissolution as an alternative remedy where the merger had no valid purpose). The Connecticut Supreme Court has left open the possibility of personal claims against corporate officers for their own misconduct. Yanow v. Teal Industries, Inc., 422 A.2d 311 (Conn. 1979).

In numerous jurisdictions, the legislatures have created one or more statutory exceptions, which the courts in some cases have expanded. In Minnesota, the exception is limited to fraud. MINN. STAT. ANN. § 302A.471.4 (West 1985). In 11 jurisdictions, an exception to exclusivity exists only in the event of either fraud or illegality. *See* COLO. REV. STAT. § 7-4-123(4) (1986); FLA. STAT. ANN. ch. 607.1302(5) (Harrison Supp. 1992); GA. CODE ANN. § 14-2-1302(b) (1994); IDAHO CODE § 30-1-80(d) (1994); MASS. ANN. LAWS ch. 156B, § 98 (Law. Co-op. 1979); MICH. STAT. ANN. § 21.200(762)(3) (Law. Co-op. 1993); MONT. CODE ANN. § 35-1-827(e)(2) (1993); NEB. REV. STAT. § 21-2079(4) (1991); N.H. REV. STAT. ANN. § 293-A:81(IV) (1987); N.M. STAT. ANN. § 53-15-3(D) (Michie 1993); OR. REV. STAT. § 60.554(2) (1993). Pennsylvania allows an exception in cases of either fraud or unfairness. 15 PA. CONS. STAT. ANN. § 1105 (Pamph. 1994). California, meanwhile, provides an exception to exclusivity only in cases where a party to the merger either directly or indirectly controls the corporation. CAL. CORP. CODE § 1312 (West 1990). Three jurisdictions, in addition to the statutory exceptions of fraud and illegality, provide through judicial precedent for an exception in the event corporate officers breach their fiduciary duties. *See* Walter J. Schloss Assocs. v. Arkwin Indus., 460 N.E.2d 1090 (N.Y. 1984) (interpreting N.Y. BUS. CORP. LAW § 623(k) (McKinney 1986)); Umstead v. Durham Hosiery Mills, Inc., 578 F. Supp. 342 (M.D.N.C. 1984) (interpreting N.C. GEN. STAT. § 55-13-02(b) (1990)); Farnsworth v. Massey, 365 S.W.2d 1 (Tex. 1963) (interpreting TEX. BUS. CORP. ACT ANN. § 5.12(G)). New Jersey adds to these three exceptions a statutory *ultra vires* exception. N.J. STAT. ANN. § 14A:11-5(2) (1969); Mullen v. Academy Life Ins. Co., 705 F.2d 971 (8th Cir.), *cert. denied*, 464 U.S. 827 (1983).

In 18 jurisdictions, namely Alabama, Arkansas, Delaware, Hawaii, Illinois, Iowa, Kentucky, Maine, Maryland, Missouri, Ohio, Rhode Island, South Carolina, South Dakota, Tennessee, Virginia, Washington, and Wisconsin, courts have created at least one judicial exception, even though exclusivity is not addressed by statute. *See* PRINCIPLES, *supra* note 9, § 7.24 reporter's note 1 (listing state treatments of appraisal).

Thirteen jurisdictions, namely Alaska, Arizona, the District of Columbia, Kansas, Louisiana, Oklahoma, Mississippi, Nevada, North Dakota, Utah, Vermont, West Virginia, and Wyoming, have neither a relevant statutory provision nor any court decisions creating exceptions to the exclusivity requirement.

[95] As the Delaware Supreme Court stated, "[a]n appraisal proceeding is a limited legislative remedy intended to provide shareholders dissenting . . . on the grounds of inadequacy of the offering price with a judicial determination of the intrinsic worth

When misvaluation is the only issue, transactions can proceed while the dispute over valuation is resolved.

When shareholders claim their directors breached their fiduciary duties or committed a fraudulent or illegal act, however, the claim relates to the stock's value but does not relate to misvaluation. There are three ways to resolve these other claims. First, legislatures could make the appraisal remedy exclusive, thereby extinguishing all claims other than misvaluation. This option leaves unpunished such undesirable conduct as fraud, illegality, or breach of fiduciary duty. Moreover, the flexibility of most modern corporate statutes permits management to achieve desired results via more than one route. If an exclusive appraisal remedy extinguished all other shareholder claims, directors would have greater incentives to achieve their goals through appraisal-triggering transactions. Viewed in this light, the appraisal remedy could become a vehicle for the unjust enrichment of corporate directors or controlling stockholders at the expense of shareholders, since the shareholders' right to the fair value of their stock may depend on the resolution of non-appraisal claims.

As a second option, legislatures could expand the appraisal process to encompass all claims relating to the stock's value.[96] Legislatures rarely select this route because it undermines any objective to make the appraisal remedy expedient and inexpensive; the resulting remedy is useless to all but large or wealthy shareholders.[97]

The third and most popular route limits the appraisal remedy to misvaluation claims and requires all other claims to be brought in another proceeding. Most courts and statutes have chosen this approach as the best way to expedite the appraisal process,[98]

(fair value) of their shareholdings." Cede & Co. v. Technicolor, Inc., 542 A.2d 1182, 1186 (Del. 1988).

[96] *See, e.g.*, Steinberg v. Amplica, Inc., 729 P.2d 683, 689-90 & n.16 (Cal. 1986) (interpreting CAL. CORP. CODE. § 1312(a), (b) & (c) (West 1994) to permit a shareholder's claim of misconduct to be aired in an appraisal proceeding); Sturgeon Petroleums Ltd. v. Merchants Petroleum Co., 195 Cal. Rptr. 29, 33 (Cal. Ct. App. 1983) (allowing shareholders to litigate their claims of misconduct in appraisal proceedings "where one of the merging corporations is controlled by the other or where the corporations are under common control").

[97] *See supra* note 3 (discussing objections to the cumbersome and expensive appraisal process); *see also* Cede & Co. v. Technicolor, Inc., C.A. Nos. 7129 & 8358, slip op. at 20 (Del. Ch. Jan. 13, 1987) ("the legislature has narrowed the issues involved [in an appraisal proceeding] in order to provide a fair and economical remedy for a specific problem").

[98] *See, e.g., id.*

punish wrongdoers,[99] and avoid inconsistent judgments.[100] As the Delaware Supreme Court explained in contrasting appraisal and non-appraisal claims:

> To summarize[,] in a[n] . . . appraisal action the only litigable issue is the determination of the value of the appraisal petitioners' shares on the date of the merger, the only party defendant is the surviving corporation and the only relief available is a judgment against the surviving corporation for the fair value of the dissenters' shares. In contrast, a fraud action asserting fair dealing and fair price claims affords an expansive remedy and is brought against the alleged wrongdoers to provide whatever relief the facts of a particular case may require.[101]

With the appraisal remedy in most jurisdictions limited by statute or case law to claims of misvaluation, shareholders must either discard their claims of fraud, misrepresentation, or illegality or sue for damage or injunctive relief.[102] Moreover, while

---

[99] The public policy concern is that appraisal is not a deterrent to individual misconduct because the corporation or its transactional partner is liable for the stock's fair value regardless of the cause of the original undervaluation. *See, e.g., Steinberg*, 729 P.2d at 691; Rabkin v. Philip A. Hunt Chem. Corp., 498 A.2d 1099 (Del. 1985). In some fact patterns, however, such as when a corporation goes private, the resulting corporation consists only of the wrongdoers; thus, when the corporation pays the appraisal award, the award does have a deterent effect. In both *Rabkin* and *Cede & Co.*, 542 A.2d 1182, for example, no innocent shareholders remained in the resulting corporation.

[100] PRINCIPLES, *supra* note 9, § 7.22 reporter's note 5, at 333.

[101] *Cede & Co.*, 542 A.2d at 1187–88; *see also* Alabama By-Products Corp. v. Neal, 588 A.2d 255, 256 (Del. 1991) (discussing the limited scope of an appraisal proceeding). *Rabkin* illustrates the limited utility of appraisal. Shareholders alleged that their directors manipulated the timing of the transaction so as to deprive shareholders of a contractual right to receive a fixed price if the transaction closed within a certain timeframe. As that fixed price was arguably more than the stock's fair value, appraisal would not compensate shareholders for the loss. The Delaware Supreme Court agreed, holding that plaintiffs' claims based on alleged breaches of fiduciary duties raised issues that an appraisal could not address. *Id.* at 1106; *see also* Weinberger v. UOP, Inc., 457 A.2d 701, 714 (Del. 1983) (stating that the appraisal remedy "may not be adequate in certain cases, particularly where fraud, misrepresentation, self-dealing, deliberate waste of corporate assets, or gross and palpable overreaching are involved").

[102] Shareholders can sue for a preliminary injunction, based on fraud or breach of fiduciary duties, to stop the transaction prior to the merger vote or prior to consummation of the merger. *See* Steven D. Gardner, *A Step Forward: Exclusivity of the Statutory Appraisal Remedy for Minority Shareholders Dissenting from Going-Private Merger Transactions*, 53 OHIO ST. L.J. 239, 282 (1992). It is difficult, however, for plaintiffs to meet the irreparable injury and inadequate remedy at law requirements for an injunction. Fischel, *supra* note 18, at 900; *see also* James Vorenberg, *Exclusiveness of the Dissenting Stockholder's Appraisal Right*, 77 HARV. L. REV. 1189, 1190 n.6 (1964) (listing cases denying an injunction because there was an adequate legal remedy). Rescissory damages grant plaintiffs the value of their shares given up in a merger, calculated either at the time of the merger or at the time the damages are awarded.

shareholders need not prove any wrongdoing to recover in appraisal, they are unlikely to successfully challenge the transaction price absent a conflict-of-interest transaction.[103] Thus, both appraisal and non-appraisal actions are likely to be brought only when shareholders have some grievance about management's conduct.

The issue underlying Dean Manning's thesis is whether the appraisal remedy deflects these non-appraisal suits. A related but more subtle question is whether plaintiffs, in order to utilize their non-appraisal remedies, are tempted to characterize every parent-subsidiary cash-out merger as self-dealing and every misrepresentation or omission as fraudulent.[104] For a variety of reasons, shareholders have incentives to pursue class actions instead of, or in addition to, their appraisal action.[105] The primary incentive is the cost, fee, and recovery structure attendant to each remedy. In class actions, financial concerns are minimized because the class representative initially pays the costs and the expert fees, and the attorneys' fees are apportioned from a suc-

---

Fischel, *supra* note 18, at 898 n.99. Professor Fischel criticizes an award of rescissory damages because they compensate plaintiffs for a post-merger increase in market value unrelated to the fraud. He argues that this award gives plaintiffs a costless option to speculate after the fraud is discovered. *Id.* at 901.

[103] *See infra* notes 222–228 and accompanying text (discussing that a distortion in the transaction price is most likely to occur when management has a conflict of interest).

[104] Fischel, *supra* note 18, at 901 (asserting that plaintiffs have incentive to characterize all misrepresentations or omissions as fraudulent).

[105] Plaintiffs may simultaneously bring appraisal and class actions for damage or injunctive relief and request a stay of the appraisal action until the issues that form the basis for the class action are resolved. Pence v. Nat'l Beef Packing Co., 1976 WL 1703 (Del. Ch. Jan. 9, 1976) (holding the grant of a stay of an appraisal action during pendency of non-appraisal causes of action within the court's discretion); Welsbach Corp. v. Sley, 1971 WL 3 (Del. Ch. May 7, 1971) (same); *see also*, William Prickett & Michael Hanrahan, *Weinberger v. UOP: Delaware's Effort to Preserve a Level Playing Field for Cash-Out Mergers*, 8 DEL. J. CORP. L. 59, 81 (1983) (commenting that uncertainty may lead some shareholders to file both appraisal and class action suits); Kanda & Levmore, *supra* note 2, at 470 nn.129–30 (discussing possible exclusivity rules where shareholders seek remedies in addition to appraisal). In *Cede & Co.*, 542 A.2d at 1188, the Delaware Supreme Court held that a shareholder who uncovered grounds for a fraud suit during appraisal discovery need not choose between remedies. Instead, the court consolidated both actions with orders to decide the fraud case first. The court held that if plaintiffs could prove fraud in the merger, they would be entitled to rescissory damages; if plaintiffs could not prove fraud, that claim would be dismissed and the plaintiffs could then proceed with their appraisal claim. *Id.* at 1191; *see also* Beard v. Ames, 468 N.Y.S.2d 253, 258 (N.Y. App. Div. 1983) (ruling that shareholders can simultaneously pursue appraisal and injunctive actions). While plaintiffs may bring both claims, defendants may argue that the appraisal demand constitutes an adequate legal remedy and therefore urge the court to deny the requested injunction. Due to appraisal procedures requiring that steps be completed within certain timeframes, shareholders who forego their appraisal remedy for a class action and then lose the class action may not be able to assert their appraisal rights at a later time. *See*

cessful recovery.[106] Moreover, in class actions, shareholders suffer no financial downside, because at a minimum, they retain the transaction price while seeking damages. Appraisal, in contrast, neither automatically nor uniformly relieves the individual from financial concerns; only some jurisdictions assess appraisal costs against the corporation.[107] The appraisal structure in other jurisdictions is more chilling, as courts have discretion to allocate costs between the parties.[108] Furthermore, shareholders in appraisal actions risk the possibility of receiving less than the transaction price. Finally, shareholders in jurisdictions without appraisal prepayment provisions may receive no return on their investment for prolonged periods of time.[109]

Just as shareholders have financial incentives to pursue non-appraisal actions, plaintiffs' attorneys are similarly motivated by the size of potential fees. While most jurisdictions provide that attorneys' fees in appraisal awards may be apportioned from the recovery,[110] as are fees in class actions, these equivalent structures often do not produce equivalent results. The potential amount of the attorneys' fees—and therefore their willingness to undertake a matter—is directly linked to the number of shares in the plaintiff class. In appraisal proceedings, the class tends to be small.[111] In contrast, the representative nature of a class action

---

Rabkin, 498 A.2d 1099 (holding shareholders may lose appraisal rights by instituting class actions without first perfecting their appraisal rights).

[106] Prickett & Hanrahan, *supra* note 105, at 77 n.116. For a discussion of attorneys' fees in class actions, *see* John C. Coffee, Jr., *Understanding the Plaintiff's Attorney: The Implications of Economic Theory for Private Enforcement of Law Through Class and Derivative Actions*, 86 COLUM. L. REV. 669, 678 & n.26 (1986). This fee structure also enables shareholders with small holdings to be represented in a class action, when otherwise, the costs in appraisal might not offset the benefits. *Id.* at 679.

[107] *See, e.g.*, MBCA § 13.31 (1984).

[108] *See, e.g.*, DEL. CODE ANN. tit. 8, § 262(j) (1991).

[109] *See, e.g.*, *In re* Appraisal of Shell Oil Co., No. 8080 (consolidated), 18 DEL. J. CORP. L. 1156, 1165 (Del. Ch. 1992) (awarding interest to former shareholders to cover seven-year appraisal process); Cede & Co. v. Technicolor, Inc., 542 A.2d 1182, 1184 (Del. 1988) (noting that discovery lasted two years). In contrast, under the Model Act the corporation at an early stage must give the shareholders the undisputed amount for their stock, thereby giving them access to at least a portion of their capital. MBCA § 13.25(a) (1984).

[110] *See, e.g.*, MBCA § 13.31(c) (1984); DEL. CODE ANN. tit. 8, § 262(j) (1991); *see also In re* Appraisal of Shell Oil Co., 18 DEL. J. CORP. L. at 1164, 1169–71 (holding that the standards governing discretionary awards of attorneys' fees in appraisal class actions are identical to those in other types of shareholder litigation, and awarding the attorneys 25% of the benefits awarded to shareholders, consisting of the difference between the merger consideration and what shareholders received as a result of the appraisal action, including increases from settlement offers, plus interest).

[111] Some shareholders will fail the individual statutory eligibility requirements while others will be either ignorant of, apathetic about, or simply unable to analyze the

does not require any action by individual shareholders, except for those shareholders desiring to "opt out" of the class. Ease of formation, coupled with a lack of financial concerns, tends to make the plaintiff group in class actions relatively large. The allocation of attorneys' fees as a percentage of the recovery of the class, when the process is skewed toward creating a large class, may be the pivotal reason for the preference for class actions.[112]

In addition to being better suited to shareholders' resources, the class action may also be better suited than is appraisal to shareholders' goals. The ability to seek an injunction or rescissory damages[113] significantly strengthens the minority's bargaining power. As a result, plaintiffs are drawn to class actions to air a broader range of grievances.

Thus, if, as Dean Manning states, the purpose of appraisal rights is to protect the majority's transaction from derailment by the minority, this purpose has largely been thwarted. That purpose is served only when shareholders forego litigation in order to pursue appraisal rights. In an effort to make the remedy easier to use, however, no claims other than misvaluation are heard in appraisal actions. This division of issues has ironically dictated the remedy's downfall; shareholders and their attorneys, with good reason, prefer class actions to appraisal actions. Thus, as shareholders often choose a non-appraisal remedy,[114] the appraisal remedy today does not provide the protection for majority shareholders that Dean Manning envisioned.

In sum, two of the three purposes originally advanced for the remedy do not fare well in the modern corporate statute, leaving the cash exit at fair value as the only viable function. Appraisal today provides neither continued compensation for the loss of veto power nor protection for the majority from derailment of

---

appraisal transaction procedure within the required timeframe. *See* Pellman v. Cinerama, Inc., 503 F. Supp. 107, 110 (S.D.N.Y. 1980) (noting that the economic realities of small appraisal classes, unlike in class actions, make extensive discovery and vigorous prosecution unlikely).

[112] In jurisdictions that allocate costs to the corporation, *see supra* note 15, the correlation between the size of the class and the expected fees explains the preference for class actions. *Cf.* Roberta S. Karmel, *Shareholder Concerns Not Fully Alleviated by Revisions in ALI's Provisions on Appraisal Proceedings*, BNA's CORPORATE COUNSEL WEEKLY, Sept. 11, 1991, at 8 (suggesting that despite the requirement in the PRINCIPLES that the corporation pay all appraisal costs, shareholders may find appraisal more difficult and expensive than a class or injunctive action).

[113] *See supra* note 102 (discussing rescissory damages).

[114] *See supra* notes 105–112 and accompanying text (explaining why rational shareholders might prefer non-appraisal remedies).

corporate transactions. The next section analyzes the monitoring theory, which some contemporary scholars advance as the remedy's current function.

III. Analysis of the Monitoring Theory

Several prominent scholars have suggested that all three original theories justifying appraisal rights are outmoded.[115] While advocating slightly different perspectives on the remedy's current function, these commentators share the view that appraisal serves as either an *ex ante* or an *ex post* monitor of management's conduct.[116] For example, Professors Kanda and Levmore suggest that the appraisal remedy serves a "discovery" function, "uncovering suspiciously non-arm's-length bargains or side payments to the target's managers, guiding future fiduciary suits, and, generally, deterring misbehavior."[117]

Professor Fischel's thesis is that appraisal rights provide minority shareholders with *ex ante* protection from appropriation of the value of their shares as a result of either a lack of bargaining power or managerial conflicts of interest.[118] He further

---

[115] *See, e.g.*, Fischel, *supra* note 18, at 876 (discussing the "inadequacy of common explanations for the appraisal remedy"); Kanda & Levmore, *supra* note 2, at 432 (discounting three commonly perceived goals of the appraisal remedy as "fail[ing] to illuminate current statutes").

[116] *See* Eisenberg, *supra* note 11, at 69–77; Fischel, *supra* note 18, at 876; Kanda & Levmore, *supra* note 2, at 434–37.

[117] Kanda & Levmore, *supra* note 2, at 444. Professors Kanda and Levmore discuss two other possible goals of the appraisal remedy: inframarginality and reckoning. *Id.* at 437–43. They explain "inframarginality" as follows:

> Shareholders, or legislators acting on their behalf, may realize *ex ante* that they will not all "appreciate" their shares identically, that the marginal, or market, price therefore understates their average valuation of these shares, and that appraisal may serve to protect these inframarginal valuations.

Id. at 437–38. They describe "reckoning" as follows:

> Appraisal at the time of the change thus may serve as a point of "reckoning"; prior performance is reckoned and future performance can be judged from the bench mark determined at appraisal . . . . In short, since appraisal of some shares requires appraisal of the enterprise's value as a whole, it may be sensible for monitoring purposes to allow for or even to encourage appraisal, or reckoning, at important junctures.

Id. at 441–42. Professor Eisenberg makes a similar suggestion, positing that in some situations, appraisal rights serve a valid goal of protecting shareholder expectations. Eisenberg, *supra* note 11, at 83–84.

[118] Fischel, *supra* note 18, at 881–87. Neither Fischel nor other advocates of the monitoring theory address why a monitor is needed when all appraisal-triggering transactions (with the exception of short-form mergers) provide a monitor by requiring shareholder approval. Presumably, advocates of the monitoring theory believe shareholder approval is an insufficient safeguard, either because the directors' disclosure of

posits that this *ex ante* protection encourages investment, resulting in higher share prices which benefit all shareholders.[119] As an example, Professor Fischel presents a "prisoner's dilemma" in which an acquiror announces a two-tiered offer: a tender offer at $60 for fifty-one percent of the target shares selling at $50, and a subsequent freeze-out merger for the remaining forty-nine percent at $30, for a blended rate of approximately $45.[120] Due to a lack of coordination, no shareholder can negotiate for all shareholders. Fischel thus concludes that shareholders would tender as much of their stock as possible at $60 to avoid the possibility of receiving $30 per share.[121] Appraisal rights in the second step merger, however, might allow minority shareholders to recover something approximating the $50 pre-transaction market value, rather than the $30 price set by the acquiror. From this example, Professor Fischel derives his thesis that appraisal

---

information is deficient or because voting serves a *per se* inadequate monitoring function. Directors must make a full and fair disclosure of all material facts before shareholder approval is binding. Weinberger v. U.O.P., Inc., 457 A.2d 701 (Del. 1983) (finding that approval by a majority of the minority shares did not preclude a fairness inquiry where the vote was predicated on inadequate disclosure). The presumed defect, therefore, must lie in the vote. If management has sufficient voting power to dictate the outcome, the shareholder vote is mere window dressing. Where non-management votes are needed, however, studies negate the claim that shareholders simply rubber-stamp whatever their directors recommend, especially when the vote concerns a significant issue, such as an appraisal-triggering transaction. Frank H. Easterbrook & Daniel R. Fischel, *Voting in Corporate Law*, 26 J. LAW & ECON. 395, 416–17 (1983). Easterbrook and Fischel conclude as follows:

> If shareholders' voting serves as a monitoring device on self-interested behavior by management, shareholders should vote against these [charter] amendments.
>
> The evidence is consistent with this hypothesis. Many institutional investors depart from their customary adherence to the Wall Street Rule (vote with management or sell your shares) and vote against shark repellent amendments.

Id.; *see also* Louis Lowenstein, *Pruning Deadwood in Hostile Takeovers: A Proposal for Legislation*, 83 COLUM. L. REV. 249, 332 (1983) ("A tender offer commands attention. Given the substantial stakes involved, the decision will be important to the shareholders personally . . . . A greater concern may be that shareholders reject automatically all responses by management.") Yet, when faced with a management buyout or other transaction in which management has a strong conflict of interest, concerns about the policing power of a shareholder vote linger. Lucian A. Bebchuk, *Limiting Contractual Freedom in Corporate Law: The Desirable Constraints on Charter Amendments*, 102 HARV. L. REV. 1820, 1839 (1989) (suggesting that shareholders rationally would vote for all proposed amendments); Robert C. Clark, *Vote Buying and Corporate Law*, 29 CASE. W. RES. L. REV. 776, 779 (1979) (discussing shareholders' "rational apathy"). Therefore, while a monitor clearly is needed when shareholder approval is a foregone conclusion because of voting control, a monitor in addition to the shareholder vote may be useful in conflict transactions, even when approval by non-management votes is required. The appraisal remedy supplies this second layer of protection.

[119] Fischel, *supra* note 18, at 879.

[120] *Id.* at 878–79.

[121] *Id.* at 879.

rights benefit all shareholders: "The appraisal remedy, by reducing the probability that the shares of the minority will be acquired at a price unilaterally set by the majority, increases the price the minority will pay for shares to the benefit of both the majority and the minority."[122]

Despite its superficial appeal, Fischel's theory is disquieting. First, he offers no support for his conclusion that avoidance of the prisoner's dilemma increases, rather than reallocates, the wealth of all shares.[123] Second, his premise that the appraisal remedy deters transactions that aim to appropriate wealth from minority shareholders lacks empirical support. How can we know how many self-dealing transactions did not occur or how many transpired at higher prices due to the specter of appraisal rights?[124] Third, his assumption about shareholder behavior is problematic. Based on his example, appraisal rights will deter a two-tier offer at $60 and $30 (for a blended rate of $45) where the stock is traded at $50 before the announcement of the transaction. Offerors will not necessarily be so deterred, however, because not all shareholders will obtain the appraisal price. Many investors, particularly those holding small numbers of shares, will forego their appraisal rights in favor of a more expedient route, while the appraisal procedures will foil others. The point, therefore, is a simple one: offerors may have incentives to proceed with two-tiered offers and gamble that the blend of $60 at the front end, $30 to those not obtaining appraisal, and $50 to those who correctly exercise their appraisal rights will yield a net benefit.

Regardless of the merit in Fischel's theory, it raises several concerns. First, it should be recognized that Professor Fischel's argument that the appraisal remedy enhances the value for all shareholders by providing *ex ante* protection against conflict transactions is really just a simple repackaging of an old theory. Because Fischel and other scholars fail to recognize that one original purpose for the remedy was to provide a cash exit at

---

[122] *Id.* at 880.

[123] From a practical perspective, it is unclear why shareholders would value protecting their downside more than the potential loss of their upside gain, and therefore pay a premium to own stock in a corporation with appraisal rights. Furthermore, assuming an acquiror will pay a certain sum and no more, Fischel's theory relates only to how an acquisition price will be allocated among shareholders, rather than explaining why an acquiror will pay more than its fixed sum for a corporation with appraisal rights.

[124] Even if Fischel is correct that the specter of appraisal as a remedy deters two-tiered and other conflict transactions geared at appropriating the minority's wealth, he provides no evidence that the appraisal remedy is the best and most efficient mechanism for policing such transactions.

fair value,[125] they believe that their identification of the remedy's reservation price, and its concomitant monitoring effect on conflict transactions, is new. This monitor, however, is simply one consequence of providing a cash exit at fair value; a fair value exit reclaims for minority shareholders the wealth that the directors or the controlling shareholders attempted to appropriate.[126]

Second, because Fischel believes the remedy's utility is *ex ante* the transaction, rather than *ex post*,[127] he discounts the paucity of appraisal demands. By providing a justification for why the remedy is seldom used, Fischel ignores the possibility that the small number of appraisal cases could be the result of defects in the remedy.[128] If these defects are significant but ignored, the remedy's weakened capacity to monitor management self-dealing motivates acquirors to gamble on two-tiered and other value-appropriating transactions at the expense of minority shareholders.

Finally, Fischel's monitoring theory is flawed because he does not limit the remedy to the context from which he draws his support. Like so many other advocates of the appraisal remedy,[129] Fischel extrapolates his justification for the appraisal remedy from one subset of appraisal-triggering transactions, namely, conflict-of-interest mergers. From this example, Fischel reasons that the appraisal remedy establishes "a reservation price for all or part of the firm in situations where coordination or conflict-of-interest problems might otherwise lead to" appropriation of the value of the minority's shares.[130] While his logic suggests limiting the remedy to conflict transactions, he never so concludes.[131] Deriving a monitoring goal for the remedy from one subset of trans-

---

[125] *See supra* note 60 and accompanying text.

[126] Fischel's theory is novel in his assertion of a benefit for all shareholders, for the traditional theory of appraisal rights identified minority shareholders as the remedy's sole beneficiaries. Although not articulated by Fischel, this difference is important. If Fischel is correct, hostility toward the appraisal remedy for deterring transactions desired by the majority may be overstated. According to Fischel, even if an appraisal demand were to thwart a particular transaction, the majority would have nonetheless already benefited from higher share prices generated by the remedy's mere existence. *See supra* notes 88–90 and accompanying text (discussing how the cash drain in appraisal actions can prevent consummation of corporate transactions).

[127] Fischel, *supra* note 18, at 884.

[128] *See supra* note 17 (discussing the paucity of appraisal cases).

[129] PRINCIPLES, *supra* note 9, § 7.21 cmt. d; Buxbaum, *supra* note 11, at 1720; Alfred F. Conard, *Amendments of Model Business Corporation Act Affecting Dissenters' Rights (Sections 73, 74, 80 and 81)*, 33 BUS. LAW. 2587, 2595–96 (1978).

[130] Fischel, *supra* note 18, at 885.

[131] In a footnote, Professor Fischel comments that it is anomalous for appraisal rights to be available in arm's-length mergers. *Id.* at 884 n.36.

actions, and then failing to restrict the remedy to that same subset, is inconsistent.

This inconsistency has its costs. Allowing appraisal rights in non-conflict transactions contradicts the corporate law's well-established business judgment rule. That rule presumes that when corporate directors make a business decision, they act independently, on an informed basis, in good faith, without abusing their discretion.[132] Thus, unless shareholders can dislodge the rule's presumption by showing that management violated its duties of care or loyalty, shareholders are precluded from second-guessing the wisdom of their directors' business decisions. Appraisal, however, contradicts the business judgment rule: the remedy second-guesses the price that disinterested management acting with due care negotiates for its shareholders. As directors are rarely liable for a breach of their duty of care unaccompanied by a conflict of interest or other breach of their duty of loyalty,[133] there is nothing to monitor if management has no conflict of interest in the transaction. Thus, any return to the shareholders above the transaction price is a windfall.

Fortunately, such windfalls are atypical. While all statutes give appraisal rights without regard to whether management has a conflict of interest, shareholders and their attorneys appear to have given effect to this distinction. The vast majority of litigated appraisal cases involve allegations of conflict of interest.[134] The reasons are obvious. Disinterested management is likely to produce a better transaction price for its shareholders than are directors with a conflict. Furthermore, while management in non-conflict situations is, by definition, not self-interested when it agrees to a transaction price, shareholders exercising their ap-

---

[132] DENNIS J. BLOCK ET AL., THE BUSINESS JUDGMENT RULE: FIDUCIARY DUTIES OF CORPORATE DIRECTORS 20 (4th ed. 1993); *cf.* R. Franklin Balotti & James J. Hanks, Jr., *Rejudging the Business Judgment Rule*, 48 BUS. LAW. 1337 (1993) (exploring the different formulations of the business judgment rule).

[133] Cinerama, Inc. v. Technicolor, Inc., Del. Ch., C.A. Co. 835 at 34 (Oct. 6, 1994); BLOCK, ET AL., *supra* note 132, at 72–75 (citing only 10 modern cases in which directors were liable for breach of their duty of care without a concurrent breach of their duty of loyalty).

Advocates of appraisal rights also assume that judges are more capable of ascertaining the stock's fair value than are directors. That assumption is highly questionable, given the complexity of valuation. *See infra* notes 256–267 and accompanying text (discussing valuation methodologies).

[134] Elmer J. Schaefer, *The Fallacy of Weighting Asset Value and Earnings Value in the Appraisal of Corporate Stock*, 55 S. CAL. L. REV. 1031, 1032 & n.6 (1982) (noting that of 13 appraisal cases, 11 involved conflict-of-interest transactions); *see also* Thompson, *supra* note 79.

praisal rights are self-interested in seeking the highest appraisal award. In this context, a rational court should not accept the views of the interested shareholders' experts above those of disinterested management. There is thus little, if any, need for appraisal rights in non-conflict situations, as shareholders are unlikely to win on the merits. The existence of the remedy, however, regardless of whether management has a conflict, has nuisance value, since the remedy's availability creates the potential for a cash drain.[135] Thus, while Fischel is correct that the remedy can monitor transactions, he fails to reach the right conclusion: appraisal rights should monitor only conflict-of-interest transactions.[136]

In sum, providing a cash exit at fair value remains a useful function for the modern appraisal remedy, both to provide a cash exit and to police conflict-of-interest transactions. This monitoring role is demonstrated by the large percentage of appraisal cases brought by shareholders in market companies[137] who disregarded the market's cash exit in favor of a judicial review of the transaction price. Moreover, despite the advantages of non-appraisal actions when a conflict of interest exists,[138] these statistics also demonstrate that some shareholders nevertheless utilize their appraisal remedy.

The examination above in Parts II and III concerning the historical and modern goals of appraisal demonstrates several factors that have restricted the remedy to its current narrow sphere of operation: (1) the appraisal-triggering transactions are infrequent events; (2) other shareholder remedies provide advantages over appraisal proceedings; and (3) the market, when available, provides an alternative cash exit, arguably at fair value. The net result is a remedy that today has limited utility. The remedy provides a fair cash exit in non-market corporations, but the events triggering that exit rarely occur. The remedy's fair cash exit may or may not be available in market corporations, depending on whether the applicable statute contains a market-out exception. When the remedy is available in market corporations, it

---

[135] *See supra* notes 88–90 and accompanying text (discussing possible negative effects of appraisal rights on effectuating the majority's will).

[136] *But see supra* notes 105–113 and accompanying text (discussing the advantages in conflict transactions of non-appraisal remedies).

[137] *See supra* note 80 (noting that of 103 decisions involving appraisal rights, 66 transactions involved shareholders in market companies).

[138] *See supra* notes 105–113 and accompanying text (discussing the advantages of non-appraisal actions).

is unlikely to be exercised in non-conflict transactions; in conflict transactions, however, the remedy may be either used or foresaken for other remedies. Thus, the remedy's transaction base makes it of limited utility in non-market corporations, and both the market and the market-out exception make the remedy either unnecessary or unavailable in many market corporations. The remaining question is whether and how to extricate the remedy from its narrow sphere to make optimal use of it in the next generation of corporate statutes.

IV. The Next Generation of Appraisal Statutes

Some of the existing appraisal statutes have been amended to increase the remedy's efficacy. Most notably, the Model Act has simplified the procedural requirements for demanding appraisal,[139] required the corporation to prepay its estimate of fair value,[140] and, absent unusual circumstances, allocated all appraisal costs to the corporation.[141] This Article incorporates all of these amendments as the foundation for its recommendations and seeks to further improve the remedy.

The Model Act's improvements mitigate only some of shareholders' preferences for class actions. As noted above,[142] the significant factor causing this preference may be the "opt out" feature of most class actions for damages. An "opt out" feature tends to produce larger classes, thereby giving the plaintiffs' attorneys a larger recovery pool on which their fees are based. Despite this important factor, however, this Article does not recommend changing the appraisal remedy from its current "opt in" system to an "opt out" class. One reason for not doing so is that groups demanding appraisal, particularly in close corporations, often would not meet the numerosity requirement for class actions.[143] More importantly, conscribing shareholders to the ap-

---

[139] MBCA §§ 13.20–13.23 (1991).

[140] MBCA § 13.25(a) (1991).

[141] MBCA § 13.31(a) (1991).

[142] *See supra* note 112 and accompanying text.

[143] Rule 23(a)(1) of the Federal Rules of Civil Procedure requires that the class in a class action be sufficiently large so that the alternative of joinder is "impracticable." *See, e.g.*, Robidoux v. Celani, 987 F.2d 931, 935 (2d Cir. 1993) (emphasizing that for certification, joinder need only be impracticable, not necessarily impossible). There is no threshold number of class members guaranteed to satisfy the numerosity requirement of Rule 23(a). *See, e.g.*, Kelley v. Norfolk and Western Ry., 584 F.2d 34 (4th Cir. 1978). A class comprised of many hundreds of members will likely meet this test.

praisal remedy unless they take affirmative action and opt for the transaction price makes an unwarranted assumption about shareholder preferences. A shareholder's failure to vote for a transaction may result either from apathy or from an objection to the transaction. Under either scenario, it is unsafe to assume that such a shareholder would prefer appraisal rights to the transaction price. The transaction itself, rather than the price, may be the objectionable aspect. Alternatively, the shareholder may prefer immediate payment of the transaction price to the complete or partial delayed payment in appraisal. Further, the shareholder may simply be averse to the potential financial risks attendant to the appraisal process.[144] The *Principles*, in rejecting the reformation of appraisal into an "opt out" system, cited a related reason:

> Given the well-known phenomenon of shareholder apathy, any assumption that all shareholders who failed to vote for [the] proposed transaction wished to dissent and pursue their appraisal remedy, unless they explicitly opted out, would be of doubtful validity. Such a rule would treat shareholders who failed to vote in favor of a proposed charter amendment, which sometimes may be of no more than ministerial significance, as having elected to dissent. Again, this seems over-inclusive . . . . It would be an unsound presumption in such a case to treat apathetic shareholders as if they wished to dissent.[145]

Thus, while converting the appraisal remedy into an "opt out" system would strengthen the remedy against its competitors, policy reasons counsel against such a change. Nevertheless, the Model Act's improvements, coupled with the other recommendations in this Article, can together mold appraisal rights into a viable and attractive remedy.

Utilizing the next generation of appraisal statutes as a fair cash exit and as a monitor of certain conflict-of-interest transactions requires resolution of three issues. First, at what junctures should

---

Classes of 10 litigants or less, however, will almost certainly not meet this test. BAICKER-MCKEE, *supra* note 92, at 269. *But cf.* Grant v. Sullivan, 131 F.R.D. 436 (M.D. Pa. 1990) (noting that as few as 14 people have been certified as a class). Smaller classes are usually subject to joinder of parties under Rule 20. When the number of proposed class members falls between approximately 25 and 100, the probability of meeting the numerosity requirement varies among judicial districts. *See, e.g.*, Hernandez v. Alexander, 152 F.R.D. 192, 194 (D. Nev. 1993) (holding class of 52 too small where proponents of class did not demonstrate any unusual difficulties with joinder). Most states have similar numerosity rules. *See, e.g.*, DEL. CT. CH. R. 23.

[144] *See supra* notes 107–109 and accompanying text (discussing financial risks of exercising appraisal rights).

[145] PRINCIPLES, *supra* note 9, pt. 7, ch. 4 introductory note, reporter's note 6, at 299.

the remedy provide this exit and monitoring function? Second, which shares should be afforded this exit right? Third, how should fair value be derived, given that the efficacy of the appraisal remedy as a fair exit and as a monitor is largely determined by the derivation of fair value?

While recognizing the difficulty in effecting widescale reformation of any established remedy, this Article proffers the following recommendations:

- In non-market corporations, appraisal rights should be available for mergers, share exchanges, sales of substantially all assets, certain charter amendments, and as stated in the corporate charter, bylaws or board resolutions. In market corporations, appraisal rights should be available for mergers, share exchanges, sales of substantially all assets, and certain charter amendments if any of these transactions is a conflict-of-interest transaction, and as stated in the corporate charter, bylaws or board resolutions.
- Appraisal rights should be available for all shares participating in earnings as well as all shares adversely affected by charter amendments.
- Fair value should be determined by consideration of all relevant valuation techniques. The award in appraisal-triggering transactions should apply key person but not minority or non-marketability discounts, should exclude synergy gains, and should award compound interest at the average rate at which the corporation borrows funds, or if no such rate is available, at a rate determined by the court.

A. *The Appraisal Triggers*

Unlike nineteenth-century corporations,[146] modern corporate statutes grant corporations perpetual life.[147] Corporate immortality stands in sharp contrast to other business associations. Nei-

---

[146]It was common in the early 19th century for corporate charters to limit the corporation's life to periods ranging from 5 to 30 years. LAWRENCE M. FRIEDMAN, A HISTORY OF AMERICAN LAW 190 (2d ed. 1985). Over time, however, business needs demanded the extension of corporate life. *See id.* at 191 (observing that over the period from 1800 to 1850, the business corporation developed into a form subject to "few real restrictions" on its duration).

[147]All jurisdictions specify that unless otherwise stated, a corporation's existence is perpetual. MBCA ANN., *supra* note 2, § 3.02 statutory comparison 2 (Supp. 1993); *see, e.g.*, MBCA § 3.02 (1984); DEL. CODE ANN. tit. 8, § 102(b)(5) (1991). Perpetual

ther limited liability companies[148] nor partnerships[149] enjoy perpetual life. Perpetual life is generally regarded as one of the principal advantages of incorporation[150] because it affords investors and creditors greater certainty that a dissolution will not readily occur.[151] Such certainty encourages long-term capital commitments to the enterprise[152] and permits long-term planning.

Perpetual life also has significant costs, borne primarily by shareholders in non-market corporations. Immortality has the concomitant effect of making an investment permanently illiquid. This illiquidity, coupled with delegated corporate governance, increases the potential for investors in non-market corporations to be abused or otherwise become dissatisfied.[153] For example, long periods of time with no ready exit may turn amiability into enmity; shareholders who are employees may become incapacitated or disfavored; priorities may change; or heirs and spouses desiring employment may complicate a once harmonious relationship.[154] The large number of shareholder dissension cases in non-public corporations is thus not surprising:[155]

---

life stems from the perception of the corporation as an entity separate from its shareholders.

[148] Limited liability companies have a fixed duration. *See, e.g.*, DEL. CODE ANN. tit. 6, § 18-801 (1993) (limiting the duration of a limited liability company to 30 years, unless the agreement states otherwise).

[149] Partnerships dissolve upon the death, bankruptcy or withdrawal of any partner. UNIFORM PARTNERSHIP ACT §§31–32 (1914) [hereinafter UPA]. The same result occurs if a general partner in a limited partnership dies, becomes bankrupt, or withdraws from the partnership. A partner may withdraw at any time, thereby effecting a dissolution, and retrieving the investment. REVISED UNIFORM LIMITED PARTNERSHIP ACT § 801 (1985). Nevertheless, any partner may be liable for damages if the petition for dissolution is contrary to the partnership agreement. Charles R. O'Kelley, Jr., *Filling Gaps in the Close Corporation Contract: A Transaction Cost Analysis*, 87 NW. U. L. REV. 216, 234 (1992). Concurrent ownership functions like a partnership in that each owner has the right at any time to compel a partition of the property. CORNELIUS J. MOYNIHAN, INTRODUCTION TO THE LAW OF REAL PROPERTY 213 (2d ed. 1988).

[150] HARRY G. HENN & JOHN R. ALEXANDER, LAWS OF CORPORATIONS § 75 (3d ed. 1983).

[151] JESSE H. CHOPER ET AL., CASES AND MATERIALS ON CORPORATIONS 2 (3d ed. 1989).

[152] *Id.*

[153] *See generally id.* at 5 (discussing certain costs connected with the long-term nature of corporations).

[154] *See* Steven C. Bahls, *Resolving Shareholder Dissension: Selection of the Appropriate Equitable Remedy*, 15 J. CORP. L. 285, 287 (1990) (noting that familial conflicts are a major reason for close corporation failures); Hetherington & Dooley, *supra* note 21, at 3 (discussing the effects of "time and human nature" on consensus in close corporations). In contrast, in public corporations the market both offers an exit and disciplines management. *See* Bahls, *supra*, at 289 (discussing the role of the market in ensuring that management acts in the best interests of shareholders).

[155] *See* Harry J. Haynsworth, *The Effectiveness of Involuntary Dissolution Suits as a*

> The position of the minority in the close corporation is as unique as it is precarious: no other form of business organization subjects an owner to the dual hazards of a complete loss of liquidity and an indefinite exclusion from sharing in the profitability of the firm. The majority's exploitative power is not, of course, unlimited, and demonstrably oppressive or unlawful conduct may give the minority a claim for judicial relief. But the legal limits on the majority's conduct are generous, and the right of the excluded faction to obtain relief is dependent on its ability to prove that the limits have been overstepped. Obtaining relief can be a lengthy, difficult, and uncertain process which is invariably costly, whatever the outcome. The hazards of litigation thus serve to broaden the area of marginally lawful behavior within which the majority can promote its own interests at the expense of the minority.[156]

Shareholders in non-market corporations rely primarily on private contracts either among themselves or with the corporation to provide an exit.[157] Contractual exits, however, are problematic. The long-term nature of the corporation makes it impossible to anticipate all events that might warrant an exit. In addition, shareholders in close corporations tend to begin a venture with strong positive feelings toward each other, thereby making them less prone to create complex contractual exits.[158]

Remedy for Close Corporation Dissension, 35 CLEV. ST. L. REV. 25, 87 (1987) (citing O'NEAL'S OPPRESSION, *supra* note 78, at iii) (describing as "phenomenal" the growth in recent years in the number of cases decided on grounds of shareholder oppression).

[156] Hetherington & Dooley, *supra* note 21, at 6; *see also* Robert B. Thompson, *The Shareholder's Cause of Action for Oppression*, 48 BUS. LAW. 699, 699 (1993) ("The statutory norms of centralized control and majority rule, when combined with the lack of a public market for shares in a close corporation, leave a minority shareholder vulnerable in a way that is distinct from the risk faced by investors in public corporations").

[157] Such contracts may be contracts to have the corporation and other shareholders agree to purchase stock from the shareholders upon the occurrence of certain events. Such contracts may also call for dissolution by less than majority vote. *See, e.g.*, MBCA § 7.32 (1984) (allowing shareholders to contract to dissolve by less than majority vote). Instead of directly addressing the problems associated with illiquidity, corporate statutes focus primarily on corporate governance, presumably to create a fair environment for long-term investors. Fair governance is a poor substitute, however, for those who are dissatisfied but not abused. Shareholders may simply disagree with the course their managers have chosen or object to being ousted from management. Moreover, litigating governance issues is both problematic and unsatisfactory. The time and financial constraints, the requirement of some cognizable wrong, and fact that the wrongdoers stay in power make litigation unresponsive to the needs of many minority shareholders.

[158] *See* F. HODGE O'NEAL & ROBERT B. THOMPSON, O'NEAL'S CLOSE CORPORATIONS § 1.18, at 103 (3d ed. 1992) (noting that participants in small enterprises often cannot foresee possible future problems in their relationship); Hetherington & Dooley, *supra*

Corporate statutes do not significantly alleviate the problems illiquidity generates. Some statutes authorize an array of remedies, including partitioning the assets, appointment of a receiver or provisional director, payment of dividends, or invalidation of corporate actions,[159] each of which is problematic.[160] While all corporate statutes provide for voluntary dissolution,[161] the requisite majority approval[162] precludes a unilateral minority exit. Most corporate statutes also permit involuntary dissolution if there is deadlock at the shareholder or managerial level[163] or managerial misconduct.[164] Most involuntary dissolution statutes define managerial misconduct narrowly, as activity that is illegal, fraudulent or wasteful.[165] Minority shareholders are rarely successful in meeting these stringent standards.[166] Thirty-one statutes also authorize dissolution for "oppressive" conduct.[167] Courts have defined "oppression" in a variety of ways, ranging from "harsh and wrongful conduct,"[168] to a breach of the fiduciary duties of good faith and

---

note 21, at 2–3 ("Typically [close corporations] are founded by individuals who have a virtually complete identity of interests and strong feelings of trust and confidence for one another").

[159] *See* Bahls, *supra* note 154, at 293, 312–13 (discussing the various approaches of state statutes addressing dissension); *see also* Thompson, *supra* note 156, at 723–26 (discussing additional remedies, both statutorily and judicially created, allowed in cases of shareholder dissension in various jurisdictions). Sometimes, courts employ these remedies without legislative authority. Bahls, *supra* note 154, at 313 & n.177 (describing the use of judicial license in courts' findings that legislatively granted tools for resolving dissension do not limit courts' common law equitable powers).

[160] Court-appointed receivers or directors create nervous customers, creditors and suppliers and impose additional personnel costs. Bahls, *supra* note 154, at 309; *see* O'NEAL'S OPPRESSION, *supra* note 78, §§ 7.22, 7.23 (discussing, respectively, issues and developments in state corporate laws relating to appointment of custodians and provisional directors). Other remedies, such as mandating the payment of dividends, require the courts to make business decisions, a task from which courts normally refrain; *cf.* BLOCK, *supra* note 132, at 8 (observing that the constraints of the business judgment rule keep a generally ill-equipped judiciary from becoming involved in corporate decisionmaking).

[161] *See, e.g.*, MBCA §§ 14.01–14.02 (1991); DEL. CODE ANN. tit. 8, § 275 (1991).

[162] *See, e.g.*, MBCA §§14.01–14.02 (1991). Some statutes give the shareholders the option of providing in the charter for voluntary dissolution by less than majority vote. *See, e.g.*, MBCA ANN., *supra* note 2, § 14.02 statutory comparison 2 (Supp. 1992) (listing eight jurisdictions that permit corporate provisions for a less than 50% approval of a voluntary dissolution).

[163] Bahls, *supra* note 154, at 295–96. *See, e.g.*, MBCA § 14.30(2)(i), (iii) (1984).

[164] *See, e.g.*, MBCA § 14.30(2)(ii) (1984).

[165] *See* Bahls, *supra* note 154, at 295–96 (noting that most states permit judicial dissolution where the directors act fraudulently, oppressively, or illegally); Thompson, *supra* note 156, at 708.

[166] *Cf.* Thompson, *supra* note 156, at 708 & n.61 (observing that courts rarely find sufficient grounds for dissolution on the basis of only one of these factors).

[167] *Id.* at 709 n.70. Courts in six additional states without statutory authority have dissolved corporations for oppressive conduct. *Id.*

[168] *See id.* at 711–12 (discussing judicial interpretations of the meaning of "oppression" in corporate law statutes).

fair dealing,[169] to conduct designed to defeat the minority's reasonable expectations that were central to the minority's decision to invest.[170] Even the most liberal standards for court-ordered dissolution generate problems for minority shareholders who simply disagree with the majority[171] or who lack the resources or will to litigate.[172] Thus, the most expansive standard remains unresponsive to the plight of the typical minority shareholder in a non-market corporation. Finally, half of the statutes or courts authorize buyouts either as an alternative to involuntary dissolution or in other shareholder litigation.[173] This remedy is of limited aid because the minority must still litigate[174] unless the majority voluntarily chooses to buy the minority's shares.[175]

Recognizing the limitations of the statutory exits, some scholars have argued that increased help for minority shareholders may be needed. Describing the liberalization of the grounds for involuntary dissolution,[176] Professor Thompson notes that court-ordered dissolution can properly balance the investment interests of the minority shareholder with the business interests of the majority.[177] While elastic standards offer minority shareholders a

---

[169] *Id.* at 712; *see* Baker v. Commercial Body Builders, Inc., 507 P.2d 387, 394 (Or. 1973). A single breach of fiduciary duty or a vague concern about future problems will be insufficient to cause dissolution. Thompson, *supra* note 156, at 712; *see also Baker*, 507 P.2d at 394 (concurring with the assertion that a single act, a continual course of generally "oppressive" conduct, or vague apprehensions are insufficient to warrant judicial dissolution). Even so, the linkage of oppression to fiduciary duties is beneficial to the minority, particularly in those jurisdictions that take an expansive view of such duties in close corporate settings. *See, e.g.*, Donahue v. Rodd Electrotype Co., 328 N.E.2d 505 (Mass. 1975) (quoting Cardullo v. Landau, 105 N.E.2d 843 (Mass. 1952)) (holding that stockholders in close corporations owe each other a duty similar to the duty of "'utmost good faith and loyalty'" among partners in an enterprise).

[170] *See, e.g., In re* Kemp & Beatley, Inc., 473 N.E.2d 1173 (N.Y. 1984); Meiselman v. Meiselman, 307 S.E.2d 551, 558 (N.C. 1983); Balvik v. Sylvester, 411 N.W.2d 383, 387 (N.D. 1987).

[171] *See* Thompson, *supra* note 156, at 715 (discussing the limits beyond which a court probably will not find oppression).

[172] *Cf.* O'Kelley, *supra* note 149, at 217 (pointing out that judicial interpretations of the close corporation contract leave majority shareholders many options for ignoring the minority's interests). Satisfying shareholders' reasonable expectations is also problematic where shareholders receive their shares by gift or inheritance.

[173] Thompson, *supra* note 156, at 718. Some courts, however, refuse to order buyouts absent statutory authority. *Id.* at 722 n.154.

[174] See *supra* notes 157, 166, 171 and accompanying text (describing problems for the minority in meeting the required standards in litigation).

[175] In this scenario, oppressive conduct is no longer an issue. *See* Thompson, *supra* note 156, at 719 (observing that once a party has elected to buy out an oppressed minority shareholder, the court may be precluded from interfering). If the corporation or the majority is not so willing, a court may order a buyout against the majority's will. *See id.* at 720–22 (discussing judicially directed buyouts where the purchasers had not expressed such a desire).

[176] *See id.* at 706.

[177] *Id.* at 706–07.

broader base for relief, such relief is only available upon undertaking the substantial financial and emotional burdens of litigation.

On the other hand, Professors Hetherington and Dooley propose an immutable statutory provision requiring closely held corporations or their majority shareholders either to purchase the minority's shares upon request or to effect a dissolution.[178] Professors Hetherington and Dooley found that in the vast majority of involuntary dissolution cases, the corporation or the majority buys out the minority.[179] Their proposal, which eliminates judicial discretion and the element of wrongdoing, equips minority shareholders with a viable remedy by giving their investments a partnership-like liquidity.

Yet the premise of the Hetherington-Dooley proposal is questionable because its broad conclusion is based only on the subset of cases in which minority shareholders pursue litigation. Professors Hetherington and Dooley do not provide any evidence that parties resolve their differences through buyouts when the minority does not institute suit. Committing the necessary resources and will to litigation is a significant expression of dissatisfaction. The Hetherington-Dooley proposal dramatically expands available minority remedies by granting an at-will exit to shareholders who have not demonstrated significant dissatisfaction.

The Hetherington-Dooley proposal also suffers from several practical flaws. An at-will exit right provides the same danger of minority opportunism that eventually caused the demise of shareholder veto power.[180] Furthermore, it is unreasonable to expect all corporations to function like mutual funds, continuously maintaining either sufficient cash or access to funds to repurchase each minority shareholder's interest upon demand.[181]

---

[178] Hetherington & Dooley, *supra* note 21, at 6, 50–59; *cf.* Kanda & Levmore, *supra* note 2 (discussing the possibility of granting appraisal at will).

[179] Hetherington & Dooley, *supra* note 21, at 31.

[180] *See* O'Kelley, *supra* note 149, at 243–44 & n.110 (critiquing the Hetherington-Dooley proposal); FRANK H. EASTERBROOK & DANIEL R. FISCHEL, THE ECONOMIC STRUCTURE OF CORPORATE LAW 148 (1991) (criticizing an at-will appraisal right as encouraging managers to reject risky projects); *see also* MICHAEL P. DOOLEY, FOUNDATIONS OF CORPORATION LAW 1069–70 (1995) (advocating that adding partnership-like default rules onto the corporate form is wasteful).

[181] For example, the majority may have to sell assets, reduce its business, secure a loan, or admit new shareholders to meet the financial demands of an at-will buyout. Thompson, *supra* note 156, at 706. The Hetherington-Dooley proposal, which would

Finally, the Hetherington-Dooley proposal does not reflect the problems created when one aspect of partnership law is grafted onto the corporate form. Partnership law defines a package of interrelated rights for partners and creditors. While the Uniform Partnership Act grants each partner at-will dissolution rights, the statute also requires unanimity for major decisions,[182] thereby decreasing the likelihood that a partner will trigger the dissolution provision upon personal dissatisfaction.[183] Similarly, as all general partners are personally liable for partnership debts[184] if the partnership's assets at dissolution are insufficient to satisfy creditors, all general partners have personal incentives to resolve differences rather than to dissolve the partnership. Slicing off only the partner's liquidity right distorts the balance of rights among all the parties. Armed with the dissolution trigger, but not subject to the unanimity rule or personal liability rule, close corporate shareholders would be more likely than partners to resolve or threaten to resolve their disputes by dissolution. Furthermore, such at-will dissolution would come at the expense of corporate creditors, who normally have access only to corporate assets.[185] Thus, at-will dissolution in the corporate context would clearly benefit minority shareholders at the expense of majority shareholders and creditors. The Hetherington-Dooley proposal offers no justification for minority shareholders to enjoy the benefits of the corporate form as well as partnership-like liquidity,[186] particularly when such liquidity comes at the expense of creditors and other shareholders.

The Model Act's newly adopted section 14.34[187] is a response to the findings in the Hetherington-Dooley article that buyouts

---

allow installation payments for the buyout, is an incomplete solution to these problems. *See* Hetherington & Dooley, *supra* note 21, at 51.

[182] UPA § 18(g)–(h) (requiring unanimous consent to either expand the membership of a partnership or to undertake actions in contravention of the partnership agreement).

[183] *Cf.* O'Kelley, *supra* note 149, at 231–32 (observing that the partnership dissolution mechanism does not create the same risks of minority opportunism in a partnership that such a mechanism would create in a corporation, because partnership law permits all partners full information concerning partnership affairs and prevents withdrawal of assets except by negotiation or court order).

[184] UPA § 15.

[185] *See* Lawrence E. Mitchell, *The Fairness Rights of Corporate Bondholders*, 65 N.Y.U. L. REV. 1165, 1193 (1990) (discussing the rights of creditors and shareholders in corporate dissolution).

[186] *See generally* O'Kelley, *supra* note 149, at 218 (arguing that wholesale grafting of partnership liquidity onto the corporate form is not necessarily the result for which the parties would have bargained).

[187] Section 14.34 of the Model Act was adopted in 1991.

are the most common dispute resolution mechanism. MBCA section 14.34(a) gives the corporation or its shareholders the option of purchasing the petitioning shareholder's interest instead of proceeding with the petitioner's claim for involuntary dissolution. With the standards for involuntary dissolution becoming increasingly liberalized,[188] the MBCA's buyout provision promises to become increasingly used.

The Model Act's buyout provision is a boon to majority shareholders: it reduces their downside in involuntary dissolution cases by offering them an alternative to dissolution. For minority shareholders, however, MBCA section 14.34 does little. Unlike the majority, the minority has no option but to dedicate the resources to bring a proceeding for involuntary dissolution and hope the majority opts for the buyout. If the majority opts for the buyout, the minority is relieved from proving wrongful behavior and ends up in a comparable, although not equivalent,[189] position as would result from a successful petition for involuntary dissolution. If, however, the majority opts not to utilize the buyout provision, the minority must litigate its claim for involuntary dissolution. Moreover, the majority, but not the minority, retains the option to proceed with dissolution if dissatisfied with the buyout's derived valuation.[190]

Thus, the statutorily created exits for minority shareholders in non-market corporations are limited. Their remaining option is the appraisal remedy. Currently, the remedy is only a contingent exit, little used because the appraisal triggers are transaction-based, limited in number, and otherwise unresponsive to the dissatisfaction that illiquidity generates.

Several avenues exist for modern corporate statutes to maximize use of the appraisal remedy's ability to effectuate a cash exit at fair value. At a minimum, corporate statutes should adopt

---

[188] *See supra* notes 163–175 and accompanying text (discussing various standards for involuntary dissolution).

[189] Comparing the minority's position under § 14.34 with § 14.30, the minority benefits under § 14.34(e) because the shareholder may recover reasonable fees and expenses of counsel and experts if the court finds the shareholder had probable grounds for relief under § 14.30(2)(ii) or (iv). Conversely, the MBCA's involuntary dissolution provisions have no such reimbursement provisions. On the other hand, successful petitioners under the involuntary dissolution provisions receive their pro rata share of the liquidated corporation as soon as the liquidation is completed; in contrast, § 14.34(e) permits the court to "enter an order directing the purchase upon such terms and conditions as the court deems appropriate, which may include payment of the purchase price in installments . . . ." MBCA § 14.34(e) (1984).

[190] MBCA § 14.34(g).

an enabling provision that permits expansion of appraisal rights as provided in the corporate articles, bylaws or board resolutions.[191] Given the vast area in which corporate statutes permit private ordering,[192] little justification exists for precluding private ordering of appraisal rights.[193] Expanding appraisal rights through the articles of incorporation, corporate bylaws or board resolutions, however, requires at least majority approval. Thus, private ordering does not provide the minority with unilateral protection.

Therefore, the pivotal issue is whether corporate statutes should expand the availability of appraisal rights. As noted above,[194] the list of transactions that trigger the appraisal remedy varies among the states. The line distinguishing those transactions that offer appraisal rights from those that do not is often subjective and arbitrary.[195] Even bypassing Dean Manning's claim that granting appraisal rights in mergers is in itself irrational,[196] it is impossible to create any meaningful list of appraisal triggers that depends on whether a corporate transaction generates a fundamental change. Unfortunately, having identified the remedy's proper function as providing a cash exit at fair value does nothing to

---

[191] *See, e.g.*, MBCA § 13.02(a)(5); *cf.* PRINCIPLES, *supra* note 9, § 7.21(e) (permitting a grant of additional appraisal rights in the corporate articles).

[192] *See, e.g.*, MBCA § 7.32 (1984) (allowing shareholders wide latitude to contract on governance issues); MBCA § 2.02(b)(4) (allowing shareholders option largely to eliminate or limit the liability of their directors).

[193] The only policy reason to forbid private ordering of appraisal rights is when the exclusivity provision is interpreted to eclipse all other remedies, thereby sheltering fraud and other undesirable conduct. *See supra* notes 96–97 and accompanying text.

[194] *See supra* note 52.

[195] There is a stark inconsistency between transactions that afford appraisal rights and those that do not because a corporation can fundamentally change in a variety of ways, some of which yield appraisal rights while others do not. For example, a steel corporation (with a flexible purpose clause) can purchase for cash a larger corporation whose sole business is making guacamole. Notwithstanding this fundamental change in the corporation's business and the variation among state statutes in the appraisal area, no state would grant appraisal rights to the shareholders of the steel corporation for this cash purchase. Were the steel corporation to engage in a cash merger with the guacamole corporation, however, all states would grant appraisal rights to these same shareholders. MBCA ANN., *supra* note 2, § 13.02 statutory comparison, at 1371 (Supp. 1993) (observing that all 51 jurisdictions grant appraisal rights in some merger situations). The weakness of the "fundamental change" theory is its inability to explain such disparities. Indeed, the "fundamental change" theory is doomed, both because the definition of "fundamental change" is so subjective and because the flexibility of modern corporate statutes provides several routes to achieve the same result.

[196] *See* Manning, *supra* note 3, at 248 (concluding that "[t]he appraisal remedy as applied to mergers is a pure anachronism—a residual adaptation to an extinct theological problem").

further the analysis of when appraisal should provide either an exit or a monitor.[197]

Any list of appraisal-triggering transactions is likely not worth the considerable amount of ink that has already been spilled on this topic.[198] Whatever list may be crafted, it will not meet the needs of shareholders in non-market corporations, whose dissatisfaction is often not transaction-based. For shareholders in market corporations, who have a continuous exit available through the market, the list of appraisal triggers is important primarily as a monitor. Near uniformity exists among the states that mergers and asset sales should trigger the remedy.[199] All states that authorize share exchanges agree that such transactions should trigger the remedy.[200] There is some disagreement about charter amendments, which thirty-three jurisdictions list among their appraisal triggers.[201] As meaningful distinctions cannot be drawn, this Article recommends adopting the most common appraisal triggers: mergers, sales of substantially all assets, share exchanges and certain charter amendments.[202]

---

[197] Since the appraisal remedy is not based on any contract breach or other wrong, the list of triggering transactions should not be derived from the fundamental change premise. *See supra* note 50 and accompanying text. Moreover, with the vested contract rights theory discredited, no logic exists explaining why shareholders whose corporations undergo a fundamental change—however defined—should be provided any remedy.

[198] The futility of attempting to craft an exhaustive list of appraisal-triggering transactions is evidenced by the efforts of those who have tried. *See, e.g.*, EISENBERG, *supra* note 11, at 69–84; Kanda & Levmore, *supra* note 2, at 432; Manning, *supra* note 3.

[199] *See supra* note 52 (noting that all states give appraisal rights in some mergers, and that 50 jurisdictions, with the notable exception of Delaware, provide for appraisal rights in sales of substantially all assets).

[200] *See supra* note 52.

[201] Unlike the merger context, appraisal rights for charter amendments reallocate wealth from those shareholders remaining in the corporation to those shareholders demanding appraisal. This resource transfer is less apparent in mergers, where the acquiror most often pays the appraisal bill. *See supra* note 2 (discussing payment obligations in appraisal proceedings).

[202] This Article recommends adopting the Model Act's list of any amendment of the articles of incorporation that materially and adversely affects rights in respect of a dissenter's shares because it:

> (i) alters or abolishes a preferential right of the shares;
> (ii) creates, alters, or abolishes a right in respect of redemption, including a provision respecting a sinking fund for the redemption or repurchase, of the shares;
> (iii) alters or abolishes a preemptive right of the holder of the shares to acquire shares or other securities;
> (iv) excludes or limits the right of the shares to vote on any matter, or to cumulate votes, other than a limitation by dilution through issuance of shares

Currently, shareholders may increase their liquidity by contractual buyouts, and, in some jurisdictions, through expanded appraisal and minority dissolution rights.[203] The transaction base of the appraisal triggers does not adequately respond to the illiquidity plight of minority shareholders in non-market corporations. Although minority shareholders may be dissatisfied with their illiquidity, the question remains whether the appraisal remedy is the appropriate solution to their dissatisfaction.

Professors Hetherington and Dooley mount a strong argument, based on both equity and efficiency, for providing liquidity for minority shareholders.[204] While their proposed at-will dissolution right disproportionately favors minority shareholders, a periodic buyout—such as every ten years—would better balance the needs of those shareholders most in need of exit rights with the needs of the corporation and the majority for stability, continuity, and long-range planning.[205] In addition, unlike an at-will buyout, a periodic buyout would force the minority to maintain its investment for a reasonable, although not indefinite, period of time,[206] would decrease the minority's chance for opportunistic behavior, and would alleviate the corporation's need to maintain continuous access to funds. Moreover, both creditors and majority shareholders would have notice of the potential loss in funds,[207] thereby deterring the majority's opportunistic behavior. Furthermore, a periodic buyout option would offer advantages over a limited liability company and a partnership because those entities offer liquidity only through a dissolution.[208] Finally, unlike under the Model Act,[209] minority shareholders exercising a periodic buyout option would neither have to institute litigation nor have to prepare to sustain the grounds for involuntary dissolution.

---

or other securities with similar voting rights; or

(v) reduces the number of shares owned by the shareholder to a fraction of a share if the fractional share so created is to be acquired for cash . . . .

MBCA § 13.02(a)(4) (1984).

[203] *See, e.g.*, MBCA § 13.02(a)(5) (allowing expansion of appraisal rights by corporate charter, bylaws or board resolution); MBCA § 7.32 (permitting shareholders to contract to achieve dissolution by less than majority vote).

[204] Hetherington & Dooley, *supra* note 21, at 43–44 (noting that liquidity is an essential component for capital market competition).

[205] *Id.* at 50–51.

[206] The Hetherington-Dooley proposal would permit shareholders to exercise their buyout right only after holding shares for two years. *Id.* at 51.

[207] The potential depletion of corporate funds in a periodic buyout occurs any time the corporation contracts to repurchase a shareholder's stock.

[208] *See supra* notes 148–149.

[209] *See supra* notes 187–188 and accompanying text.

Despite the merits of a periodic buyout, expanding appraisal rights beyond a transaction base is problematic. Such public ordering of exit rights is a dramatic solution to the real but limited problem of shareholder illiquidity. Given the contractual flexibility by which corporate shareholders can provide their own exit,[210] and given that alternative business associations could have been selected to provide greater liquidity,[211] it is unwise to assume all shareholders would be better off with a mandatory periodic buyout. Shareholders, however, should carefully consider the merits of contractually expanding their appraisal remedy to include a periodic buyout so as to alleviate their illiquidity problem.

B. *Market-Out Exception for Non-Conflict Transactions*

This Article recommends that appraisal statutes contain a market-out exception that is limited to non-conflict transactions. The limitation of the market-out exception to non-conflict transactions runs counter to all state corporate statutes, the *Principles*, and the Model Act. All current statutory law takes an absolute position either for or against a market-out exception. Such absolute positions ignore the substantial benefits of a market-out exception, and intermingle those benefits with the potential problems of conflict transactions.

The proposed limited market-out exception would grant appraisal rights for conflict transactions but deny them in non-conflict transactions. While there are different ways to define conflict-of-interest transactions, this Article recommends targeting certain relationships which create conflicts of interest, such as the following: (1) beneficial ownership of a substantial amount of stock; (2) power of appointment of directors, whether through contract, stock ownership, or otherwise; (3) transactions with an officer of the corporation; or (4) transactions in which any director receives any financial benefit not available to the shareholders, with the exception of the director's continued employment with the corporation.[212]

---

[210] *See supra* note 157.

[211] *See supra* notes 148–149 and accompanying text.

[212] Another approach would label certain transactions, such as parent-subsidiary transactions or management buyouts, as conflict transactions. This approach is questionable because appraisal rights may not be available in all transactions that warrant

This Article's proposal for a market-out exception limited to non-conflict transactions thus reflects the remedy's two primary functions: providing a fair cash exit and monitoring conflict transactions. Outside these two contexts, the appraisal remedy creates costs without offsetting benefits.[213] It is uncontroverted that the remedy's cash exit function is generally replicated by the market, with lower valuation and transaction costs than the appraisal remedy involves. The controversy between advocates and opponents of a market-out exception instead centers primarily on whether the market's cash exit is at fair value.

As noted above,[214] twenty-four jurisdictions have adopted a market-out exception. The other jurisdictions, the Model Act, and the *Principles* have declined to add this exception to their statutes.[215] Interestingly, in rejecting the market-out exception, the Model Act and the *Principles* neither acknowledge the cash exit as the remedy's original purpose nor ascribe any current purpose to the remedy.

The Model Act and the *Principles* articulate several reasons why the market may not deliver fair value:

> The 1970s have demonstrated again the possibility of a demoralized market in which fair prices are not available, and in which many companies publicly offer to buy their own shares because the market grossly undervalues them. Under these circumstances, access to market value is not a reasonable alternative for a dissenting shareholder.[216]

Similarly, concerned that market value and fair value are not necessarily synonymous, the *Principles* argue that "the market's valuation may reflect the prospect of future adverse impact or

---

scrutiny. A third approach would exempt from appraisal rights any appraisal-triggering transaction that is approved by a majority of disinterested shares, thus leaving to the shareholders the decision of whether they have received fair value. This approach creates questions, however, concerning which shares are disinterested and whether reliance on shareholder voting is a sufficient monitor in a given transaction. *See supra* note 118 (discussing the monitoring effect of a shareholder vote).

[213] *See supra* notes 132–133 and accompanying text (discussing the remedy's transgression of the business judgment rule in non-conflict transactions); *supra* notes 88–90 and accompanying text (discussing the negative effect appraisal rights can have on transactions); *infra* notes 258–259 and accompanying text (discussing the speculative valuation process).

[214] *See supra* note 82 (observing that the vast majority of corporations whose stock is traded are incorporated in jurisdictions with a market-out exception).

[215] PRINCIPLES, *supra* note 9, § 7.21 cmt. d; Conard, *supra* note 129, at 2595–96 (discussing why the MBCA deleted its market-out exception in 1978).

[216] Conard, *supra* note 129, at 2595–96; *see* Buxbaum, *supra* note 11, at 1248 (arguing that market does not always reflect fair value). Professor Buxbaum also notes that many appraisal-triggering transactions involve conflict-of-interest transactions. *Id.*

other disadvantage to the minority shareholders . . . and hence does not provide a protection against it."[217] The Model Act makes a more direct attack on market price, arguing that the market reflects the transaction price rather than fair value.[218] The *Principles* further attack market value, arguing that the market's valuation can reflect only publicly available information. Therefore, management's knowledge of non-public information makes it alone privy to when stock is undervalued. In fact, this undervaluation can simply be generated by management by withholding favorable information. The *Principles* conclude as follows: "This potential informational asymmetry is likely to have special significance with respect to transactions in which management may itself be interested, either directly or indirectly, such as parent/subsidiary mergers, management buyouts, and 'going private' transactions."[219]

Thus, the MBCA's and the *Principles*' objections to the market-out exception reflect a variety of concerns about whether market price represents fair value. It is significant that neither statute's commentary argues that the market never delivers fair value; such a position would contradict the widely accepted efficient market hypothesis,[220] which states that stock prices fluc-

---

[217] PRINCIPLES, *supra* note 9, § 7.21 cmt. d. The PRINCIPLES use Fischel's hypothetical, albeit with different numbers. *See supra* notes 120–124 and accompanying text (restating and critiquing Fischel's hypothetical and conclusions). It should be noted that some states have a controlled business combination statute that would permit the transaction described by the PRINCIPLES only if the second-step transaction is approved by disinterested shares or if the price paid in the second step at least equals the price in the first step. Thus, in some jurisdictions Fischel's hypothetical transaction could not legally occur. *See, e.g.*, ME. REV. STAT. tit. 13A, § 910 (West Supp. 1992); MD. CODE ANN., CORPS. & ASS'NS § 3-601 *et seq.* (1993).

[218] Conard, *supra* note 129, at 2595–96. Moreover, the market cannot exclude the effects of the transaction on market price. Some statutory appraisal provisions define fair value as the value of stock immediately prior to the effectuation of the appraisal-triggering transaction. *See, e.g.*, MBCA § 13.01(3) (1984) (defining fair value); *see also infra* notes 276–282 and accompanying text (discussing synergy gains in appraisal valuation).

[219] PRINCIPLES, *supra* note 9, § 7.21 cmt. d, at 310.

[220] There are four popular theories that analyze whether market value and fair value are synonymous: the Efficient Market Hypothesis (EMH), the Capital Asset Pricing Model (CAPM), the Security Analysis Theory, and the Noise Theory. Of these, only the EMH has gained widespread acceptance. *See* Ronald J. Gilson & Reinier H. Kraakman, *The Mechanisms of Market Efficiency*, 70 VA. L. REV. 549, 549 (1984) ("Of all recent developments in financial economics, the efficient market hypothesis . . . has achieved the widest acceptance by the legal culture"). In fact, two of the other four theories are simply variations of the EMH. The CAPM assumes that the market is efficient but also assumes that all investors have identical views of the risks and returns of stock. *See* Jeffrey N. Gordon & Lewis A. Kornhauser, *Efficient Markets, Costly Information, and Securities Research*, 60 N.Y.U. L. REV. 761, 776 (1985); Lynn A. Stout, *Are Takeover Premiums Really Premiums? Market Price, Fair Value, and*

tuate with insignificant variations around the stock's fair value, because the market reflects prices upon which informed buyers and sellers agree.[221] Instead, both the MBCA and the *Principles* reason solely from the effect that conflict transactions have on the market value of stock.

Specifically, the *Principles* extrapolate from a conflict transaction that the market price may reflect the possibility of future adverse impact on minority shareholders.[222] Similarly, the MBCA's concern about a "demoralized market" is appropriate only if there is a conflict transaction. Absent a conflict, the market is not "demoralized" because present or future adverse impacts on shareholders are appropriate elements of fair value if the directors have made decisions consistent with their fiduciary duties. The MBCA's concern that the market price will reflect the transaction price is similarly misguided; the transaction price represents the quintessential fair value if the directors negotiated it in accordance with their fiduciary duties:[223]

Corporate Law, 99 YALE. L.J. 1235, 1239 (1990). The Noise Theory concludes that a market absent human traders is efficient but that human instincts, such as investors reacting to other investors' trading decisions, skew that efficiency. *See* Donald C. Langevoort, *Theories, Assumptions, and Securities Regulation: Market Efficiency Revisited*, 140 U. PA. L. REV. 851, 859 (1992). Only the Security Analysis Theory disputes the assumption that the market is efficient, positing that the market is not efficient because it fails to account for inaccurate or late information, tax consequences and transaction costs. The Security Analysis Theory has been widely discredited, however, by empirical evidence supporting the EMH. *See* Gordon & Kornhauser, *supra*, at 834–46; Gilson & Kraakman, *supra*, at 551–52.

In its integrated disclosure and shelf registration rules, the Securities and Exchange Commission has accepted the EMH for stock that is widely traded. *See* 47 Fed. Reg. 11,380, 11,382 (1982) ("Form S-3, in reliance on the efficient market theory, allows maximum use of incorporation by reference of Exchange Act reports"); 48 Fed. Reg. 52,889, 52,892 (1983) ("Forms S-3 and F-3 recognize the applicability of the efficient market theory to those companies which provide a steady stream of high quality corporation information to the marketplace"). Similarly, for purposes of the "fraud on the market" theory in SEC Rule 10b-5, the Supreme Court has assumed that a well developed market is efficient. Basic v. Levinson, 485 U.S. 224, 241–45 (1987) (assuming the efficiency of a well-developed market for purposes of accepting the fraud on the market theory). *But see* Viacom Int'l Inc. v. Icahn, 946 F.2d 998, 1000–01 (2d Cir. 1991) (finding that despite the EMH, market price is not the sole measure of fair value).

[221] Fama, *Efficient Capital Markets: A Review of Theory and Empirical Work*, 25 J. FIN. 383 (1970), *cited in* Lynn A. Stout, *The Unimportance of Being Efficient: An Economic Analysis of Stock Market Pricing and Securities Regulation*, 87 MICH. L. REV. 613, 615 n.1 (1988).

[222] The only example the PRINCIPLES cite in support of this concern is a conflict transaction. In the hypothetical's second-step merger, both corporations are under common control. *See supra* note 217.

[223] Furthermore, to the extent that the deal price is poor, this objection fails to explain why shareholders should be entitled to an alternative to the market in appraisal-triggering transactions but not in other transactions.

> In the absence of a conflict of interest, less policy justification exists for upsetting the bargain reached between the selling and purchasing firms. The law of contracts does not generally give a seller the right to ask a court to reform a contract and impose a higher price simply because it appears in hindsight to have sold at too low a price. The appraisal remedy should not invite the court to substitute freely its own judgment as to fair value for that of the board, unless special factors are present.[224]

The *Principles*' concern about the market not reflecting nonpublic information is sensible only for conflict transactions. Management has only a few reasons to withhold information. First, information may not be ripe for disclosure; such information may, however, be insufficiently definite to be quantified in an appraisal valuation. Second, management may in good faith believe some corporate purpose will be served by keeping the information secret, such as a purchase of land believed to be valuable. Given that corporations whose stock is subject to a market-out exception are also subject to the disclosure requirements[225] and liability provisions[226] of the federal securities laws, these corporations are not likely—absent a conflict of interest—to enter into appraisal-triggering transactions when they cannot publicly disclose material information. The only real danger to a fair market valuation of stock exists when management has a self interest in distorting the market price. Indeed, the *Principles* concede that management's ability to unfairly affect the market price is particularly acute in conflict transactions.[227] In short, by reasoning solely from conflict-of-interest examples but extending their concerns to non-conflict transactions, the MBCA and the *Principles* have unfairly tarnished the market-out exception. Their broad-based concerns about market price lack support in the literature[228] and in actual practice.

The remaining two objections of the Model Act[229] and the *Principles*[230] to a market-out exception are not related to fair

---

[224] PRINCIPLES, *supra* note 9, § 7.22 cmt. c.

[225] Securities Exchange Act of 1934 §§ 12, 13, 15 U.S.C. §§ 78l, 78m (1988); Securities Act of 1933 § 5, 15 U.S.C. § 77(e) (1988).

[226] Securities Act of 1933 §§ 11, 12, 17, 15 U.S.C. §§ 77k, 77l, 77q (1988); Securities Exchange Act of 1934 §§ 10(b), 13(e), 14(e), 15 U.S.C. §§ 78j, 78m, 78n (1988).

[227] *See* PRINCIPLES, *supra* note 9, § 7.22 cmt. d, at 319–20 (suggesting that courts should be cautious in deferring to the board's valuation when they have a conflict of interest in the transaction).

[228] *See supra* note 220 (discussing the efficient market hypothesis).

[229] Conard, *supra* note 129, at 2595–96.

[230] PRINCIPLES, *supra* note 9, § 7.21 cmt. d.

value. First, they lament the loss of exclusivity caused by a market-out exception; without appraisal rights, the remedy cannot be exclusive. The exclusivity concern is that collateral litigation can spark uncertainty about whether the transaction will close.[231] This policy tension between the benefits of a cash exit through the market and the costs of a decline in exclusivity, however, is a false issue. A shareholder without appraisal rights may seek injunctive relief to resolve issues about whether the transaction will close. Moreover, given that most appraisal cases involve conflict transactions,[232] in which exclusivity is largely nonexistent,[233] the trade-off between a market-out exception and loss of exclusivity is more theoretical than real.[234] The proposal of a market-out exception limited to non-conflict transactions, however, largely satisfies any remnant concern about the market-out exception's impact on exclusivity; appraisal rights would remain intact for conflict transactions, which provide the underlying rationale for the MBCA's and the *Principles*' support for retaining the appraisal remedy.

Finally, the MBCA opposes a market-out exception because of legitimate concerns that the denial of appraisal rights creates exit problems for the seller who legally or functionally cannot use the market.[235] If the market cannot provide a fair cash exit to those who are legally precluded from using the market, a market-out exception should provide for this group to retain its appraisal rights.[236] Another exception, although less compelling, might encompass large holdings that practically cannot be sold in the market.

---

[231] *Id.*

[232] *See* Thompson, *supra* note 79 (finding 103 decisions involving appraisal rights from 1984 to 1994, wherein 64 of 80 transactions identified were conflict transactions); *see also* Schaefer, *supra* note 134, at 1032 & n.6 (listing 13 appraisal cases, of which 11 involved conflict-of-interest mergers).

[233] *See supra* notes 98–103 and accompanying text.

[234] At least two authors have observed that courts in Delaware have intervened at the same frequency in cases where appraisal was permitted as in cases where appraisal was denied. ERNEST L. FOLK III, THE DELAWARE GENERAL CORPORATION LAW 395–96 (1972); Cyril Moscow, *Aspects of Shareholder's Rights*, 18 WAYNE L. REV. 1003, 1028 (1972).

[235] Conard, *supra* note 129, at 2595–96. These shareholders include those who have restricted shares, shareholders who cannot sell within a short timeframe without violating the federal securities laws, and shareholders whose holdings are too large for the market to absorb.

[236] *See, e.g.*, CAL. CORP. CODE § 1300(b)(1) (West 1990 & Supp. 1994) (providing an exception to the market out exception for shares whose transfer is restricted either by the corporation or by law).

C. *Which Shares Should Have Appraisal Rights?*

This Article proposes granting appraisal rights to all shares participating in residual earnings regardless of their voting power and to all shares adversely affected by charter amendments. All states that grant appraisal rights for certain charter amendments do so for all shares, regardless of their voting power. The proposal of offering the remedy to residual claimants, however, although supported by the *Principles*,[237] is not contained in any corporate statute.

As with most other aspects of the remedy, the state appraisal statutes vary considerably over whether to restrict the remedy to voting stock. The Model Act limits the remedy to voting stock[238] with two exceptions: all shareholders of a subsidiary merging with its parent in a short-form merger;[239] and all shareholders adversely affected by certain charter amendments.[240] Twenty-five jurisdictions follow the Model Act's pattern of primarily limiting the remedy to voting stock.[241] Only three jurisdictions explicitly

---

[237] Section 7.21 of the PRINCIPLES limits appraisal rights to an "eligible holder," which is defined in § 1.17(a) as the holder of "one or more shares, whether common or preferred, that (1) carry voting rights with respect to the election of directors, (2) are entitled to share in all or any portion of current or liquidating dividends after the payment of dividends on any shares entitled to a preference, or (3) are adversely affected by an amendment of the certificate of incorporation as described in § 7.21(d) . . . ." PRINCIPLES, *supra* note 9, § 1.17(a). The PRINCIPLES thus differ from this Article's proposal by offering appraisal rights to non-participating preferred shares as well as to stock with voting rights but no equity ownership.

[238] MBCA § 13.02(a) (1984). MBCA § 13.02(a)(1)(i), (2), and (3) entitle a shareholder to appraisal in mergers, share exchanges, and sales or exchanges of all or substantially all assets if the shareholder is entitled to vote.

[239] MBCA § 13.02(a)(1)(ii).

[240] MBCA § 13.02(a)(4).

[241] ALA. CODE §§ 10-2A-162 to -163 (1992) (repealed by Acts 1994, No. 94-245, § 13, effective January 1, 1995) (excepting shareholders of a subsidiary merging with its parent in a short-form merger, all shareholders adversely affected by certain charter amendments); ARK. CODE ANN. § 4-27-1302 (Michie 1991) (excepting shareholders of a subsidiary merging with its parent in a short-form merger, all shareholders adversely affected by certain charter amendments); CAL. CORP. CODE § 1300 (West 1990 & Supp. 1994) (excepting shareholders of a disappearing corporation in a short-form merger); FLA. STAT. ANN. §§ 607.1302, § 607.0902 (West 1993) (excepting approval of control-share acquisitions, shares adversely affected by an amendment to the articles of incorporation); GA. CODE ANN. § 14-2-1302(a)(1), (2), (3) (1989) (excepting shareholders of a subsidiary merging with its parent in a short-form merger, all shareholders adversely affected by certain charter amendments); IND. CODE ANN. § 23-1-44-8 (Burns 1989) (excepting only approval of control-share acquisition); IOWA CODE ANN. § 490.1302 (West 1991) (excepting shareholders of a subsidiary merging with its parent in a short-form merger, all shareholders adversely affected by certain charter amendments); KY. REV. STAT. ANN. § 271B.13-020 (Baldwin 1989) (excepting shareholders of a subsidiary merging with its parent in a short-form merger, all shareholders adversely affected by certain charter amendments); LA. REV. STAT. ANN. § 12:131

provide the appraisal remedy for all equity holders, regardless of voting power.[242]

The language of the remaining appraisal statutes is remarkably ambiguous on this point. These statutes typically state that the remedy is available to "any shareholder" "who shall not have voted in favor" of the proposed change.[243] It is debatable whether

---

(West 1994) (excepting shareholders of a subsidiary merging with its parent in a short-form merger); MICH. COMP. LAWS § 21.200(762) (1993) (excepting parent-subsidiary short-form mergers, amendments to articles of incorporation giving rise to appraisal rights, and approval of control-share acquisitions); MISS. CODE ANN. § 79-4-13.02 (Supp. 1994) (excepting shareholders of a subsidiary merging with its parent in a short-form merger, all shareholders adversely affected by certain charter amendments); MO. ANN. STAT. § 351.875 (Vernon 1991) (excepting shareholders of a subsidiary merging with its parent in a short-form merger, all shareholders adversely affected by certain charter amendments); MONT. CODE ANN. § 35-1-827 (1993) (excepting shareholders of a subsidiary merging with its parent in a short-form merger, all shareholders adversely affected by certain charter amendments); NEV. REV. STAT. § 78.481 (1994) (excepting shareholders of a subsidiary merging with its parent in a short-form merger, all shareholders adversely affected by certain charter amendments); N.Y. BUS. CORP. LAW §§ 806(b)(6), 910 (McKinney 1986) (excepting shareholders of a subsidiary merging with its parent in a short-form merger, all shareholders adversely affected by certain charter amendments); OHIO REV. CODE ANN. §§ 1701.74, .76, .84, .85 (Anderson 1992) (excepting shareholders of a subsidiary merging with its parent in a short-form merger, all shareholders adversely affected by certain charter amendments); OR. REV. STAT. § 60.554 (1993) (excepting shareholders of a subsidiary merging with its parent in a short-form merger, all shareholders adversely affected by certain charter amendments); S.C. CODE ANN. § 33-13-102 (Law. Co-op. 1991) (excepting subsidiary merger with a parent, parent merger with a subsidiary, material and adverse amendment to the articles of incorporation, approval of a control-share acquisition, unless otherwise explicitly provided for); TENN. CODE ANN. § 48-23-102 (1988) (excepting shareholders of a subsidiary merging with its parent in a short-form merger, all shareholders adversely affected by certain charter amendments); TEX. BUS. CORP. ACT ANN. art. 5.11 (West Supp. 1994) (excepting disposition of substantially all corporate property and assets, exchange of shares of the corporation of the class or series held by the shareholder); VT. STAT. ANN. tit. 11A, § 13.02 (1993) (excepting shareholders of a subsidiary merging with its parent in a short-form merger, all shareholders adversely affected by certain charter amendments); VA. CODE ANN. § 13.1-730 (Michie 1993) (excepting shareholders of a subsidiary merging with its parent in a short-form merger); WASH. REV. CODE ANN. § 23B.13.020 (West 1994) (excepting parent-subsidiary short-form merger, amendment to articles that materially reduces number of shares owned by shareholder); WIS. STAT. ANN. § 180.1302 (West 1992) (excepting voting stock only in the case of a plan of share exchange); WYO. STAT. § 17-16-1302 (1992) (excepting shareholders of a subsidiary merging with its parent in a short-form merger, all shareholders adversely affected by certain charter amendments).

[242] KAN. STAT. ANN. § 17-6712 (1988) (including shareholders not entitled to vote); MASS. ANN. LAWS ch. 156, § 46E (Law. Co-op. 1992) (explicitly allowing remedy for non-voting stock); UTAH CODE ANN. § 16-10a-1302 (Supp. 1994) (allowing remedy for stock "whether or not entitled to vote").

[243] ALASKA STAT. § 10.06.574 (1989); ARIZ. REV. STAT. ANN. §§ 10-080 to -081 (West 1990); COLO. REV. STAT. §§ 7-4-123 to -124 (1986); CONN. GEN. STAT. ANN. § 33-373 (West 1987); DEL. CODE ANN. tit. 8, § 262 (1991 & Supp. 1992); D.C. CODE ANN. § 29-373 (Supp. 1994); HAW. REV. STAT. § 415-80 (1993); IDAHO CODE § 30-1-80 (Supp. 1994); ILL. ANN. STAT. ch. 805, para. 11.65 (Smith-Hurd 1993); ME. REV. STAT. ANN. tit. 13-A, § 908 (West 1981); MD. CODE ANN., CORPS. & ASS'NS § 3-202 (1993); MINN. STAT. ANN. § 302A.471 (West 1985 & Supp. 1994); NEB. REV. STAT.

that language restricts the remedy to voting shares or is available as well to non-voting shares, since those shares, by definition, do not "vote for" the transaction.[244] The historical basis of this language, stemming back to the English Companies Act of 1862,[245] provides little clarification; non-voting stock was probably non-existent at that time.[246] Case law on the issue is scant.[247]

Given the vast literature on appraisal rights, there is a surprising dearth of comment about which shares should qualify for the remedy.[248] As a result, the policy arguments on each side have scarcely been raised. The primary reason to limit the remedy to voting stock focuses on the costs of expanding the remedy to a larger class. First is the cost of satisfying the appraisal demands of a larger class. Second is the increased likelihood that a large cash demand will thwart a transaction or prevent favorable accounting or tax treatment.[249] Thus, the remedy indirectly enfranchises non-voting stock, giving it the power to deter transactions by increasing the costs beyond the point of efficiency for the parties.

Commentators who advocate the extension of the remedy to all shares discount this indirect enfranchisement effect. Without

---

§ 21-2079 (1987); N.H. REV. STAT. ANN. § 293-A:81, :82 (1987); N.J. STAT. ANN. § 14A:11-1 (West Supp. 1994); N.M. STAT. ANN. § 53-15-3 (Michie 1993); N.C. GEN. STAT. § 55-13-02 (1990); N.D. CENT. CODE § 10-19.1-87, -88 (1985); OKLA. STAT. ANN. tit. 18, §§ 1091, 1155 (West 1986 & Supp. 1994); 15 PA. CONS. STAT. ANN. § 1930 (Pamph. 1994); R.I. GEN. LAWS § 7-1.1-73 (1992); S.D. CODIFIED LAWS ANN. § 47-6-23 (1991); W. VA. CODE § 31-1-122 (1988 & Supp. 1994).

[244] One commentator cites statutory notice provisions to clarify this ambiguity. He argues that when non-voting shareholders are not given notice of the shareholder meeting called to authorize the transaction, non-voting shares do not have appraisal rights. William H. Pittman, Comment, *Corporations—Are Non-Voting Shares Entitled to Appraisal Rights?*, 28 MO. L. REV. 246, 247 (1963). This argument is further supported by the notice procedure, when notice is given only to voting shares for mergers but to all shareholders for charter amendments. *Id.* at 249.

[245] *Id.* at 248.

[246] *Id.*

[247] *See, e.g., In re* Harwitz, 80 N.Y.S.2d 570, 572–73 (Sup. Ct. 1948) (interpreting statute to limit appraisal to voting stock); Newman v. Arabol Mfg. Co., 245 N.Y.S.2d 442, 444 (Sup. Ct. 1963) (same); *In re* Bowman, 414 N.Y.S.2d 951, 955 (Sup. Ct. 1978) (same); *cf.* Duvall v. Moore, 276 F. Supp. 674, 688 (N.D. Iowa 1967) (finding moot whether non-voting shares have appraisal rights, as non-voting shares should have been permitted to vote on the proposed transaction).

[248] Only a handful of articles address this issue. *See* Levy, *supra* note 62, at 427–28 (advocating offering appraisal to non-voting and voting shares); Weiner, *supra* note 52, at 552 (finding courts likely to offer appraisal to non-voting and voting shares); Pittman, *supra* note 244, at 246 (interpreting language of appraisal statutes to limit the remedy to voting shares); Melvin A. Eisenberg, *The Legal Roles of Shareholders and Management in Modern Corporation Decisionmaking*, 57 CAL. L. REV. 1, 91 (1969) (asserting that appraisal need not be limited to voting stock).

[249] *See supra* notes 88–90 and accompanying text.

the remedy, it is argued, those lacking voting rights are powerless to protect their interests or to prevent consummation of the appraisal-triggering transaction. Non-voting shares are thus depicted as being in greater need of the remedy than are voting shares.[250]

The statutes, the few cases and the relevant scholarship explicate only the absence of any foundation on which to make a recommendation. This void results from the failure to tie the eligible class of stock to the function of the remedy. Assigning appraisal rights according to voting power stems from the common law right to veto fundamental changes. Modern corporate statutes discredit such a right.[251] Moreover, all other direct ties between the appraisal remedy and voting have been eliminated; only one corporate statute requires shareholders to dissent in order to demand appraisal rights.[252]

Therefore, the line of demarcation assigning appraisal rights should be drawn to include those groups of shareholders that need a fair cash exit, irrespective of their voting rights. Resolution of this issue depends on which classes of shares typically negotiate their rights extensively (as do classes of debt) and on which classes negotiate for fixed-income returns (as opposed to those classes dependent for some or all of their return on the directors' fulfillment of their fiduciary duties). Those that typically negotiate their rights extensively can provide for a fair cash exit as part of their package of rights.[253] For example, non-participating preferred shares, like debt instruments, are fixed-income investments that are extensively negotiated. While participating preferred and common shares are negotiated as well, their return is dependent on management's fulfillment of their fiduciary duties. Therefore, the remedy is needed to monitor whether management fulfills those duties in allocating the residual claim:

---

[250] PRINCIPLES, *supra* note 9, § 7.21 cmt. e (discussing the definition of "eligible holder").

[251] *See supra* notes 66–72 and accompanying text.

[252] *But see* LA. REV. STAT. ANN. § 12:131(C) (West Supp. 1994) (requiring that a shareholder dissent to claim appraisal right). Ordinarily, shareholders must simply not vote to approve the transaction. *See supra* note 243 and accompanying text. Even the Model Act, which labels its appraisal chapter "Dissenters' Rights," does not require that the shareholder dissent. MBCA § 13.21(a)(2) (1984).

[253] *Cf.* George S. Corey et al., *Are Bondholders Owed a Fiduciary Duty?*, 18 FLA. ST. U. L. REV. 971, 975 (1991) (explaining that the indenture contract largely determines a debtholder's rights); Morey W. McDaniel, *Bondholders and Corporate Governance*, 41 BUS. LAW. 413, 413 & n.1 (1986) (stating that a bondholder's rights are largely a matter of contract); Mitchell, *supra* note 185, at 1174 (noting that treatment of bondholders as contract claimants is assumed in modern case law and commentary).

> Conferring an appraisal right on those sharing in the firm's residual earnings is justified because of the absence of contractual protections guaranteeing such shareholders any fixed rate of return. Given their greater exposure, it is a fair response for the law to accord minority shareholders a right to exit at a fair value if the terms of their participation in the firm are fundamentally altered . . . This need exists even if these shareholders lack voting rights, because those lacking voting rights may be even more exposed as a result of their disenfranchisement. In contrast, shareholders who do not share in the firm's residual earning may be expected to rely principally on their contractual rights and need an appraisal remedy only when those contractual rights (including any voting rights they possess) are to be modified by a charter amendment . . . .[254]

For transactions other than charter amendments, the eligible class proposed by this Article is thus larger than that in statutes limiting the remedy to voting shares but smaller than that in statutes providing the remedy to all equity holders. This Article's recommended eligible class in the charter amendment context, however, is consistent with current corporate statutes, which grant the remedy to all shares for charter amendments.[255]

D. *Fair Value*

While most appraisal statutes entitle shareholders to their stock's fair value,[256] these statutes generally provide little, if any, guidance about how fair value should be derived.[257] This Article makes recommendations regarding the key issues in valuation:

---

[254] PRINCIPLES, *supra* note 9, § 7.21 cmt. e; *cf. In re* FLS Holdings, Inc. Shareholders Litigation, 1993 WL 104562 (Del. Ch. Apr. 2, 1993) (explaining preferred shareholders were dependent on fulfillment of directors' fiduciary duties to allocate merger price fairly between preferred and common shareholders).

[255] *See, e.g.*, ARK. CODE ANN. § 4-27-1302 (Michie 1991); GA. CODE ANN. § 14-2-1302 (1989); IOWA CODE ANN. § 490.1302 (West 1991).

[256] *See, e.g.*, MBCA § 13.02(a) (1984).

[257] *See, e.g.*, MBCA § 13.01(3):

> "Fair value," with respect to a dissenter's shares, means the value of the shares immediately before the effectuation of the corporate action to which the dissenter objects, excluding any appreciation or depreciation in anticipation of the corporate action unless exclusion would be inequitable.

The Model Act's official comment explains as follows:

> The definition of "fair value" in section 13.01(3) leaves to the parties (and ultimately to the courts) the details by which "fair value" is to be determined within the broad outlines of the definition. This definition thus leaves untouched the accumulated case law about market value, value based on prior sales, capitalized earnings value, and asset value.

MBCA § 13.01 cmt. 3.

methodology, discounts, synergy gains, and interest awards. Without such recommendations, the remedy's function of providing a fair cash exit can degenerate into a hollow promise.

Determination of fair value is highly dependent on the methodology[258] and variables chosen, the weight ascribed to each variable, and the assumptions made. The consensus of opinion recognizes that there is no universally correct way to determine fair value. Each appraisal circumstance generates its own *sui generis* valuation issues.[259]

Since no single definition can accommodate the nuances attendant to each fact pattern, professional consensus should determine the appropriate valuation techniques.[260] This recommen-

---

[258] For example, prior to 1983, Delaware determined fair value through the "block method": the court determined the corporation's market value, earnings value and asset value, ascribed to each value a different percentage that collectively totalled 100%, and multiplied each value by its assigned percentage. This approach was highly criticized. *See, e.g.*, Joseph M. Coleman, *The Appraisal Remedy in Corporate Freeze-Outs: Questions of Valuation and Exclusivity*, 38 SW. L.J. 775 (1984); Elmer J. Schaefer, *The Fallacy of Weighing Asset Value and Earnings Value in the Appraisal of Corporate Stock*, 55 S. CAL. L. REV. 1031 (1982). Despite this criticism, some courts continue to utilize the block method. *See, e.g.*, Leader v. Hycor, Inc., 479 N.E.2d 173 (Mass. 1985); Blasingame v. American Materials, Inc. 654 S.W.2d 659, 668 & n.1 (Tenn. 1983). Since 1983 Delaware has primarily used a discounted cash flow analysis which computes the present value of a corporation's projected future cash flow. *See* Neal v. Alabama By-Products Corp., C.A. No. 8282, slip op. at 16 (Del. Ch. Aug. 1, 1990), *aff'd*, 588 A.2d 255 (Del. 1991); Cede & Co. v. Technicolor, Inc., C.A. No. 7129, slip op. at 17 (Del. Ch. Oct. 19, 1990), *aff'd in part and rev'd in part*, 634 A.2d 345 (Del. 1993). Where considered "appropriate," Delaware has also endorsed other methodologies. *See* Harris v. Rapid-American Corp., C.A. No. 6462, slip op. at 7 (Del. Ch. Oct. 2, 1990) (using a comparative valuation analysis in which one determines fair value by reference to other similarly situated companies), *aff'd in part and rev'd in part*, 603 A.2d 796 (Del. 1992) (endorsing a segmented valuation technique in which the Supreme Court separately valuated each of three subsidiaries).

[259] *See, e.g.*, PRINCIPLES, *supra* note 9, § 7.22 cmt. d, at 318, including the following comments:

> Techniques used to value a natural resources corporation, which is in the business of exploiting a finite resource, would normally be inappropriate in valuing a publicly held retail store chain. Similarly, a start-up company, dependent on the viability of a product or technology not yet commercially successful, could not be valued sensibly based primarily on its past earnings or dividends or on its balance sheet.

[260] Given the array of valuation issues, appraisal has become a battle of experts, as illustrated by one case where each side had presented the chancery court with three different valuation methods:

> [W]e take the occasion to comment upon a recurring theme in recent appraisal cases—the clash of contrary, and often antagonistic, expert opinions on value. The presentation of widely divergent views reflecting partisan positions in appraisal proceedings adds to the burden of the Court of Chancery's task of fixing value.

In re Appraisal of Shell Oil Co., 607 A.2d 1213, 1222 (Del. 1992). The Delaware Supreme Court thereafter remarked that the trial court should consider, when appropriate, appointing its own expert witness to provide an objective presentation of evidence.

dation is consistent with the position taken by Delaware case law, the Model Act, and the *Principles* regarding non-conflict transactions. For example, the Delaware Supreme Court has held that determination of fair value "requires consideration of all relevant factors involving the value of a company,"[261] including "proof of value by any techniques or methods which are generally considered acceptable in the financial community."[262] The Model Act's Official Comment on the definition of fair value specifically endorses Delaware's open-ended approach.[263] For non-conflict transactions, the *Principles* similarly require fair value to be "determined using the customary valuation concepts and techniques generally employed in the relevant securities and financial markets for similar businesses in the context of the transaction giving rise to appraisal."[264]

For conflict transactions, however, the *Principles* direct the court to "give substantial weight to the highest realistic price that a willing, able, and fully informed buyer would pay for the corporation as an entirety."[265] Deference to professional consensus regardless of whether the transaction is a conflict transaction, however, is preferable. A conflict transaction merits the appraisal remedy but not a different methodology. Since the parties with the conflict are not calculating the appraisal award, their self interest is irrelevant.[266]

Id. at 1223. Indeed, 44 jurisdictions expressly authorize the court to appoint an appraiser. *See* MBCA ANN., *supra* note 2, § 13.30 statutory comparison 2 (Supp. 1993) (listing those jurisdictions that authorize court-appointed appraisers). A notable exception is the New York provision, which expressly prohibits the court from referring to an appraiser or a referee. N.Y. BUS. CORP. LAW § 623(h)(4) (McKinney 1986). Prior to 1982, the New York provision specifically permitted the court to appoint an appraiser to receive evidence and to recommend a decision on fair value. *Id.* historical note, at 304. Some statutes have provisions requiring the parties to pay the added cost of an independent appraisal. *See* CAL. CORP. CODE § 1305(e) (West 1990); ILL. ANN. STAT. ch. 805, para. 5/11.70(i) (Smith-Hurd 1993); *cf. In re* Shell Oil Co., 607 A.2d at 1223 (Del. 1992) (suggesting that any court-appointed expert would be paid by the parties). A court-appointed appraiser might be better trained than a judge to evaluate the numerous factors that often compete in deriving fair value. A handful of court-appointed appraisers might also develop a consistent approach, thereby providing better guidance to all parties on the likely outcome of the appraisal action. Litigation may decrease if the range of valuation results narrows. The court-appointed appraiser's hearing and ultimate report, however, adds yet another layer of procedure from which parties could appeal, undoubtedly adding cost and time to the appraisal process.

[261] Weinberger v. UOP, Inc., 457 A.2d 701, 713 (Del. 1983).

[262] *Id.*

[263] MBCA § 13.01 cmt. 3 (1984).

[264] PRINCIPLES, *supra* note 9, § 7.22(a).

[265] *Id.*, § 7.22(c).

[266] While the PRINCIPLES state that management's self interest to pay the lowest price possible in a management buyout warrants an auction standard so that the highest price

If parties contract for a periodic buyout, their valuation directions should prevail. In the absence of such direction, the fair value for such a buyout should be computed as for a going concern. Since this buyout would not be in lieu of dissolution, going concern value would be more appropriate than liquidation value. This position conflicts with the Hetherington-Dooley proposal, which defines fair value as "the liquidation value of the demanding shareholder's interest in the corporation, but taking into account the going concern value of the corporation, if any."[267]

This Article also recommends amendments to appraisal statutes embodying valuation issues that are dependent on policy rather than on facts. Such issues include the following: (1) whether the appraisal award should be a pro rata share of the corporation or should require that the individual shares be valued and discounted; (2) whether the appraisal award should include synergy gains expected from the transaction; and (3) at what rate interest should be computed and whether interest should be simple or compound.

1. Discounts

The first valuation issue is whether to ascribe to each share a pro rata value of the corporation as a whole or to value the shares held by a particular shareholder. The latter valuation technique would require discounting the shares for minority status and, where applicable, discounting for lack of marketability and for key persons. This Article recommends, in accordance with a majority of courts that have addressed the issue, that minority and non-marketability discounts should not apply to appraisal-triggering transactions.[268] Key person discounts should, however, apply in appraisal-triggering transactions. All applicable discounts should apply to any periodic buyout.

---

is obtained, the PRINCIPLES fail to explain why the highest price, rather than fair value, is appropriate. *Id.*, § 7.22 cmt. d.

[267] Hetherington & Dooley, *supra* note 21, at 56.

[268] *See* Jesse A. Finkelstein, *Appraisal Rights and Fairness of Price in Mergers and Consolidations*, Corp. Prac. Series No. 38-2d, at A-24 to -25 & n.60 (BNA June 1994) (discussing judicial rejections of minority and marketability discounts); *see also* LEWIS D. SOLOMON ET AL., CORPORATIONS, LAW AND POLICY 1320 (3d ed. 1994); *cf.* PRINCIPLES, *supra* note 9, § 7.22(a) (stating that no minority discount should apply and that absent extraordinary circumstances, there should not be a discount for lack of marketability).

There are several policy arguments against the use of minority and non-marketability discounts in appraisal-triggering transactions. Such discounts require courts to undertake complex and speculative inquiries.[269] Further, such discounts in an appraisal setting can "whipsaw" minority shareholders, who are first forced by majority shareholders into a transaction and then hammered by the discounts. Consequently, these discounts provide the majority with inappropriate incentives to activate appraisal-triggering transactions. The Delaware Supreme Court has addressed this possibility:

> [T]o fail to accord to a minority shareholder the full proportionate value of his shares imposes a penalty for lack of control, and unfairly enriches the majority shareholders who may reap a windfall from the appraisal process by cashing out a dissenting shareholder, a clearly undesirable result.[270]

The counterargument is that these discounts reflect the value that shares would command in the "real world." Given that corporate governance is effectuated by a majority vote, minority shareholders cannot dictate corporate policy:[271]

> "A minority shareholder could not have expected to receive a proportionate share of the going concern value of the assets if he had remained a stockholder . . . as a going concern, unless the assets as a whole, or the company as a whole, were to be sold. As a minority stockholder, he would have had no voice in a corporate policy, and no power to influence decisions as to whether to sell, and, if so, when and how. The control of these decisions is an element of value . . . ."[272]

Such logic is indisputable, but it reaches the wrong conclusion in the context of appraisal-triggering transactions. As all appraisal-triggering transactions are corporate transactions, none

---

[269] *See* PRINCIPLES, *supra* note 9, § 7.22 cmt. e, at 324.

[270] Cavalier Oil Corp. v. Harnett, 564 A.2d 1137, 1145 (Del. 1989); *see also In re* Valuation of Common Stock of McLoon Oil Co., 565 A.2d 997, 1005 (Me. 1989) (concurring with *Cavalier*); Hodas v. Spectrum Technology, Inc., C.A. No. 11265, 1992 Del. Ch. LEXIS 252, at *14 (Del. Ch. Dec. 7, 1992) (citing Rapid American Corp. v. Harris, 603 A.2d 796, 805 (Del. 1992)). *But see* Rapid-American Corp. v. Harris, 603 A.2d 796 (Del. 1992) (adding a control premium to the value of the parent corporation for its 100% ownership of its subsidiaries).

[271] Hernando Bank v. Huff, 609 F. Supp. 1124 (N.D. Miss. 1985), *aff'd*, 796 F.2d 803 (5th Cir. 1986); Perlman v. Permonite Mfg. Co., 568 F. Supp. 222 (N.D. Ind. 1983), *aff'd*, 734 F.2d 1283 (7th Cir. 1984); Atlantic States Constr., Inc. v. Beavers, 314 S.E.2d 245 (Ga. Ct. App. 1984); Stanton v. Republic Bank of South Chicago, 581 N.E.2d 678 (Ill. 1991); Moore v. New Ammest, Inc., 630 P.2d 167 (Kan. Ct. App. 1981); Blasingame v. American Materials, Inc., 654 S.W.2d 659 (Tenn. 1983).

[272] *Moore*, 630 P.2d at 177 (quoting the appraiser).

involve only the sale of individual shares. It is therefore inappropriate to value shares individually. Unlike minority and nonmarketability discounts, however, key person discounts affect the sale of the corporation as a whole. As such, key person discounts are appropriate in appraisal-triggering transactions.[273]

Appropriate valuation for a periodic buyout is not wholly analogous to the valuation for appraisal-triggering transactions. A periodic buyout would not involve any corporate action. As a buyout is more akin to a repurchase of individual shares, the shares should be valued individually. Therefore minority, nonmarketability and key person[274] discounts are appropriate in a periodic buyout.[275]

2. Synergy Gains

Another significant policy issue is whether an appraisal valuation should include the synergy gains expected from the transaction.[276] This Article recommends, in accordance with the majority of statutes, that synergy gains be excluded from the appraisal award.[277] Minority shareholders should not be able to "have it both ways": to exit the corporation through appraisal and at the

---

[273] *See Hodas*, 1992 Del. Ch. LEXIS 252 (allowing a 20% discount to reflect the significant role of one of the company's founders whose continuation in the corporation was vital to the corporation's existence). In *Hodas*, the court made clear that the key person's skills were probably not replaceable within a reasonable period of time. *Id.* at *13. The key person discount was nevertheless applied because the key person had a favored status as a minority contractor and several years of contacts with top personnel of the corporation's client. The court concluded that rather than replacing the key person, loss of this person would likely cause the corporation's demise. *Id.*

[274] Key person discounts impact both the value of the corporation and the value of individual shares.

[275] *Cf.* MBCA § 14.34 cmt. (1984) (stating that a minority discount is appropriate in calculating the fair value of petitioner's shares in a voluntary buyout).

[276] "The concept of synergy is, quite simply, that the whole may be greater than the sum of its parts." Simon M. Lorne, *A Reappraisal of Fair Shares in Controlled Mergers*, 126 U. PA. L. REV. 955, 974 (1978) (discussing different sources of synergy gains). The statutes generally phrase the issue as any appreciation or depreciation in value that can be attributed to the transaction giving rise to the appraisal remedy. *See, e.g.*, MBCA § 13.01(3) (1984).

[277] Finkelstein, *supra* note 268, at A-21; *see* MBCA § 13.01(3) (1984) (defining "fair value" as "the value of the shares immediately before the effectuation of [the transaction] . . . *excluding any appreciation or depreciation in anticipation of the corporate action*") (emphasis added); *see, e.g.*, DEL. CODE ANN. tit. 8, § 262(h) (1991); FLA. STAT. ANN. ch. 607.1301(2) (Harrison Supp. 1992); ILL. ANN. STAT. ch. 805, para. 5/11.70(j)(1) (Smith-Hurd 1993); 15 PA. CONS. STAT. ANN. § 1572 (Pamph. 1994); *cf.* Cede & Co. v. Technicolor, Inc., 542 A.2d 1182, 1187 (Del. 1988) (explaining that Court of Chancery must consider effect of future earnings on fair value of shares on day of merger).

same time to share in future expected gains.[278] Furthermore, valuing synergy gains is particularly speculative, requiring estimation of the future gains of an entity that was non-existent until the consummation of the appraisal-triggering transaction.

Moreover, synergy gains should be excluded regardless of whether the appraisal-triggering transaction is a conflict transaction. This position conflicts with that of the *Principles*:[279]

> In the case of freeze-out transactions, where the shareholder's interest in the corporation is being involuntarily terminated, this justification for denying the shareholder any share of the prospective post-merger or synergy gains simply does not apply. Where minority shareholders are being ousted by the majority shareholder or by management in a freeze-out merger, the equities favor the minority shareholders, and estoppel should not apply against them, because they have made no decision to exit. A proportionate allocation of synergy gains should not be used where it would result in an undeserved windfall.[280]

The *Principles*' position is that if shareholders are involuntarily terminated in a conflict-of-interest transaction, they deserve the synergy gains. Such a conclusion is non sequitur. Appraisal is designed to render the fair value of shares, not to compensate for the involuntary loss of shareholder status.[281] Despite the involuntary termination of the shareholder's status, an ousted shareholder bears neither the costs nor the risks of the future enterprise and thus should not share in its rewards.[282]

---

[278] PRINCIPLES, *supra* note 9, § 7.22 cmt. e, at 327.

[279] *See* PRINCIPLES, *supra* note 9, § 7.22(c):

> If the transaction giving rise to appraisal falls within § 5.10, § 5.15, or § 7.25, the court generally should give substantial weight to the highest realistic price that a willing, able, and fully informed buyer would pay for the corporation as an entirety. In determining what such a buyer would pay, the court may include a proportionate share of any gain reasonably to be expected to result from the combination, unless special circumstances would make such an allocation unreasonable.

[280] PRINCIPLES, *supra* note 9, § 7.22 cmt. e, at 327.

[281] Minority shareholders have argued against involuntary termination of shareholder status and have lost in both federal and state courts. *See, e.g.*, Santa Fe Indus., Inc. v. Green, 430 U.S. 462 (1977); Singer v. Magnavox Co., 380 A.2d 969 (Del. 1977).

[282] The PRINCIPLES' position is based on Victor Brudney & Marvin Chirelstein, *Fair Shares in Corporate Mergers and Takeovers*, 88 HARV. L. REV. 297 (1974). The premises and conclusions of that article have been critiqued by Simon Lorne. *See generally* Lorne, *supra* note 276. There are no synergy issues in a periodic buyout.

3. Interest

Issues regarding payment of interest diminish if, as recommended, jurisdictions adopt a prepayment provision for both appraisal transactions and for periodic buyouts.[283] If prepayment is made, interest is due only on the difference between the amount the corporation prepays and the amount it ultimately pays to those demanding appraisal.

Generally, state statutes grant the court discretion to order payment of interest in addition to the appraisal award.[284] Some states specify the interest rate;[285] others require the same rate of interest that is paid on judgments;[286] and still others require the same rate the corporation pays on its loans.[287] Many statutes, however, vest discretion in the court to determine the amount of interest and to decide whether interest is to be calculated as simple or compound.[288]

The appraisal award should include compound interest[289] at the average rate currently paid by the corporation on its principal bank loans.[290] Until the appraisal award is made, the corporation has, in effect, borrowed money from shareholders entitled to the award. The interest rate on the award should thus reflect the corporation's cost of borrowing money. Any statutorily fixed interest rate quickly becomes arbitrary as the corporate statute cannot be amended to keep pace with prevailing rates.

In addition, the interest should be compound rather than simple. This recommendation is consistent with use of compound

---

[283] *See supra* note 140 and accompanying text.

[284] *See, e.g.*, DEL. CODE ANN. tit. 8, § 262(h) (1991).

[285] *See, e.g.*, D.C. CODE ANN. § 29-373 (Supp. 1994) (setting five percent interest rate).

[286] *See, e.g.*, CAL. CORP. CODE § 1305(c) (West 1990); N.D. CENT. CODE § 10-19.1-88(1)(c) (Supp. 1993); UTAH CODE ANN. § 16-10a-1301(5) (1988); *cf.* TENN. CODE ANN. § 48-23-101(5) (Supp. 1994) (requiring the use of the average auction rate on U.S. Treasury Bills to determine the applicable interest rate).

[287] *See, e.g.*, HAW. REV. STAT. § 415-81(a) (Supp. 1992) (requiring the interest rate to equal the corporation's cost of borrowing money); MBCA § 13.01(4) (1984) (directing that interest be paid at average rate currently paid by the corporation on its principal bank loans or, if none, at a rate that is fair and equitable under the circumstances); *cf.* DEL. CODE ANN. tit. 8, § 262(h) (1991) (allowing a court to consider all relevant factors, including the rate of interest the corporation would have had to pay to borrow money during the appraisal proceedings).

[288] *See, e.g.*, DEL. CODE ANN. tit. 8, § 262(i) (1991).

[289] David S. Reid, *Dissenters' Rights: An Analysis Exposing the Judicial Myth of Awarding Only Simple Interest*, 36 ARIZ. L. REV. 515, 515 (1994).

[290] *See* MBCA § 13.01(4) (1984); DEL. CODE ANN. tit.8 § 262(h) (1991). If the corporation had no such loans, the court should award an equitable amount considering prevailing rates at which the corporation borrows funds.

interest by instruments that do not pay interest on a current basis[291] to compensate the lender for the loss of access to that money.[292] An award of simple interest would provide an incentive to the majority to unfairly appropriate wealth from the minority by making a low appraisal offer and prolonging the appraisal proceedings:

> In the absence of compound interest, the corporation could force the dissenter to sell his shares at less than fair value. If the corporation initially makes a low settlement offer, lengthy appraisal proceedings are inevitable. However, the allowance of only simple interest on the appraisal award could, in some cases, result in a situation where it would be more profitable for the shareholder to accept the settlement offer and invest the money in a savings account, drawing compounded interest quarterly, rather than go through the lengthy appraisal process. This result is clearly inconsistent with the purpose of the appraisal statutes which is to guarantee the dissenting shareholder fair value of his shares and to encourage the corporation to make a fair settlement offer . . . .[293]

Cases awarding simple interest have offered no policy justification for that position, instead noting the absence of statutory direction.[294]

V. Conclusion

The appraisal remedy has deep roots in corporate law. At its nineteenth-century origins, it was viewed as an important protection for both minority and majority shareholders. In the twentieth century, it has largely been viewed as unimportant because, for a variety of reasons, it has been little used. If the remedy is to be retained in the twenty-first century, some reason other than its long-standing history should compel its place in corporate

---

[291] For example, the principal portion of treasury strips award compound interest because they do not pay interest on a current basis.

[292] Thus, stock should be valued just before the transaction closes and interest should compound from that date.

[293] *In re* Valuation of Common Stock of McLoon Oil Co., 565 A.2d 997, 1007–08 (Me. 1989); *see also* Sarrouf v. New England Patriots Football Club, Inc., 492 N.E.2d 1122, 1128–29 (Mass. 1986) (finding that an award of compound interest is not an abuse of discretion); *cf.* Reid, *supra* note 289, at 538 (arguing that use of compound interest provides the corporation with incentives to resolve appraisal claims quickly).

[294] *See, e.g.*, Rapid American Corp. v. Harris, 603 A.2d 796, 808 (Del. 1992) (holding that it is not an abuse of discretion to award simple interest, given the lack of a statutory mandate); *In re* Shell Oil Co., 607 A.2d 1213 (Del. 1992) (same).

law. This Article suggests that private parties in non-market corporations consider contractually expanding their fair cash exit rights. This Article also recommends a limited market-out exception in order to utilize the remedy in an effective and efficient way. The remedy so configured is thus confined to its dual roles of providing a fair exit and monitoring certain conflict-of-interest transactions. If the recommendations contained in this Article are adopted, the appraisal remedy can regain a valued place in corporate law in the twenty-first century.

ARTICLE

IN FROM THE COLD: ENERGY EFFICIENCY AND THE REFORM OF HUD'S UTILITY ALLOWANCE SYSTEM

STEVEN FERREY*

The Department of Housing and Urban Development (HUD) provides assistance to the poor in meeting their energy needs through a number of federal programs. Under HUD's regulatory scheme, HUD provides utility allowances to local Public Housing Authorities (PHAs), which, in turn, provide such allowances to tenants. Since 1985, HUD has allowed PHAs broad discretion in setting utility allowances provided to tenants.

In this Article, Professor Ferrey revisits the subject of utility allowances, a subject he analyzed for this publication in 1986. He concludes that the current "deregulated" system fails to provide tenants with equitable allowances and recommends the use of more sophisticated statistical techniques to set allowances and investments in energy efficiency to minimize variations in energy consumption that are beyond tenants' control.

Congress appropriates more than one billion dollars annually for energy used in public housing and an equal amount for subsidized housing. According to an investigation by the U.S. General Accounting Office (GAO), the utility allowance system does not satisfy its statutory mandate. Administered by often overwhelmed local housing authorities, and overseen by a federal agency that has not tracked its own programs, the present utility allowance system denies tenants statutory protections, while also misallocating federal housing funds.

Public and Section 8 housing are the two largest sources of low-income rental housing in the nation, subsidizing several million units in more than 3000 communities nationwide. The utility allowance system for these housing units affects millions of low-income, elderly, and handicapped tenants. The failure to

---

*Professor of Law, Suffolk University School of Law; Former Senior Counsel, National Consumer Law Center, Inc.; B.A., Pomona College, 1972; M.A. (Urban and Regional Planning), University of California, Berkeley, 1977; J.D., University of California, Berkeley, 1975; postdoctoral Fulbright Fellow, University of London, 1975–1976. Professor Ferrey is the author of THE LAW OF INDEPENDENT POWER (6th ed. 1994), published by Clark Boardman Callaghan. Professor Ferrey is also the author of *Cold Power: Energy and Public Housing*, 23 HARV. J. ON LEGIS. 33 (1986), on which this Article follows. Comments about this Article can be addressed via Internet to HJOL@HULAW1.HARVARD.EDU.

administer more exactingly the federal housing law results in energy waste, tenant hardship, and administrative gridlock.

This Article reviews and documents housing energy problems at the federal and local levels and offers proposals for a revised federal utility allowance system that, while recognizing technological limitations facing public housing authorities (PHAs), would employ more sophisticated statistical tools to administer better the federal utility allowance system. It also suggests that a key to reform is implementation of available energy efficiency measures to narrow the range and variation of energy consumption in subsidized housing units. This Article also introduces legislative and administrative suggestions to improve the system at all levels and to bring it into compliance with current statutes.

Part I outlines the public housing and Section 8 housing legislation and regulatory systems for allocating utility allowances. Part II summarizes GAO's critique of these systems in practice. Part III examines the technological limitations in altering utility allowances, specifically focusing on different configurations of utility metering. These technology differences are isolated to examine options for standardizing technology. Part IV resolves this techological problem by proposing a means of devising meter-neutral utility allowances.

Parts V and VI identify two ways that the utility allowance system can be improved. First, Part V analyzes statistical tools to redesign and recalculate utility allowances and concludes that the current utility allowance calculation lacks statistical robustness (i.e., sophistication in the statistical tools used to isolate tenant-controlled versus building-determined variations in tenants' utility usage) and misallocates utility allowances in most housing authorities nationwide. This part introduces three models for calculating utility allowances and examines the benefits and limitations of each model. In addition, this part presents several additional statistical tools to reflect the central tendency, range, and skew of federal housing energy use, ultimately recommending a more statistically robust utility allowance scheme incorporating these tools.

Second, Part VI examines the interests of the various stakeholders in the utility allowance allocation and concludes that investments in energy efficiency offer an opportunity to all stakeholders to both save federal funds and improve tenants' comfort. An ambitious energy efficiency program, coupled with a more statistically sensitive and robust utility allowance formula, would

add flexibility and responsiveness to the current system and better satisfy the legislative directive of Congress.

I. The Federal Housing Assistance Programs and Their Rent Burden Mandates

A. *The Rent Burden Ceiling*

The public housing[1] and Section 8[2] certificate programs together provide federal housing and utility assistance to over three million households nationwide.[3] Federal housing law requires that those assisted households contribute no more than thirty percent of their adjusted incomes toward rent.[4] This policy is designed to maintain affordability and curtail undue financial hardship for the low-income population housed by these programs.[5]

The Department of Housing and Urban Development (HUD) is responsible for ensuring that federal housing funds appropriated by Congress are used pursuant to government regulations.[6]

---

[1] Public housing is composed of two major programs, low-rent public housing and Indian public housing, and operates under the U.S. Housing Act of 1937, ch. 896, 50 Stat. 888 (1937) (codified as amended at 42 U.S.C. § 1437 (1988)). Public housing represents approximately 20% of all low-income housing in the nation. Office of Technology Assessment, Energy Efficiency of Buildings in Cities 154 (1982) [hereinafter 1982 OTA Report].

[2] Section 8 housing assistance provides federal assistance to poor households in the private housing market through five programs: existing housing, moderate rehabilitation, substantial rehabilitation, new construction, and voucher programs. 42 U.S.C. § 1437(f) (1988).

[3] General Accounting Office, Assisted Housing: Utility Allowances Often Fall Short of Actual Utility Expenses, Vol. 1, at 2 (Mar. 1991) [hereinafter GAO, Vol. 1]. Public housing and Section 8 certificate programs are the two largest blocks of federally assisted rental housing in the nation. *Id.* at 10. Public housing assistance reaches roughly 3200 communities, where almost 10,000 separate housing projects provide subsidized dwellings for about 3.4 million individuals. 1982 OTA Report, *supra* note 1, at 154. The Section 8 certificate program is of nearly the same magnitude and size as the public housing program. One function of the Section 8 program is to provide "project-based certificates" for designated units in the moderate and substantial rehabilitation and new construction programs. In the Section 8 existing housing programs, the household leases a private housing unit of its choice from a Section 8-approved landlord participant. GAO Vol. 1, *supra*, at 70.

[4] The Brooke Amendment to the U.S. Housing Act of 1937 requires that the percentage of income paid by public housing tenants for "shelter" (rent and utilities combined) not exceed 30% of adjusted family income. 42 U.S.C. § 1437a(a)(1)(A) (1988). Alternatively, tenants may pay the lesser of 10% of gross income or their welfare rents provided by other assistance programs. 42 U.S.C. § 1437a(a)(1)(B)-(C) (1988).

[5] *See* GAO, Vol. 1, *supra* note 3, at 10.

[6] HUD meets its statutory, regulatory, and administrative requirements through its

To carry out its housing programs, HUD contracts with over 4100 public housing authorities (PHAs) nationwide that administer either one or both of the public housing or Section 8 housing programs at the local level.[7]

Although PHAs are guided by uniform HUD administrative, regulatory, and statutory requirements, PHAs differ markedly in the number, type, and overall composition of the housing units they administer.[8] In practice, PHAs face difficulties managing and implementing these uniform HUD requirements, depending upon PHA size and housing composition.[9]

Guidelines restrict eligibility for public housing and Section 8 assistance to those households earning less than the fiftieth, and sometimes the eightieth, percentile of the median income for their geographic area.[10] In contrast with public housing, which typically includes large, concentrated housing projects owned and operated by PHAs, Section 8 programs rely on voluntary participation by landlords who own widely diversified housing through PHA contracts that offer public assistance to pay rents in excess of fixed tenant rent levels.[11] On average, Section 8 households have higher annual incomes than public housing households.[12] The mechanics of federal housing assistance provided under each program, as well as the average incomes and energy use characteristics of participating residential households, differ considerably.[13]

---

staff at its headquarters, regional and area offices, and its Office of the Inspector General. *Id.* at 16.

[7] *Id.* at 10. Indian public housing is operated by Indian housing agencies. *Id.* at 70.

[8] *Id.* at 34–35. While there is no "typical" PHA, the majority of PHAs administer between 100 and 5000 units of public housing and a similar number of Section 8 units. Large urban housing authorities may administer tens of thousands of units. *Id.*

[9] *Id.* at 35–36. For example, a large PHA with several hundred employees administering 10,000 units spread across numerous housing projects and scattered site buildings must meet the same standards as a small PHA with one or two employees and less than 25 units to administer.

[10] 24 C.F.R. § 813.102 (1994); 24 C.F.R. §§ 913.102–.103 (1994).

[11] GAO, VOL. 1, *supra* note 3, at 10, 70.

[12] Based on the author's extrapolation from GAO study findings at approximately 4500 public housing and 5000 Section 8 households at six judgmentally selected PHAs. *See* GAO, VOL. 1, *supra* note 3, at 24–25. For a more complete discussion of GAO's survey techniques and findings, *see* part II.A, *infra* notes 63–86.

[13] For example, due to the fair market rent concept discussed in part IV.B, *infra* notes 186–212, Section 8 housing is more likely to include air conditioning and modern appliances typically not found in public housing projects. Generally, HUD considers air conditioners and other modern appliances, such as food freezers, as luxuries in public housing, but not necessarily in Section 8 housing. GAO VOL. 1, *supra* note 3, at 75–76.

Despite these important differences, the guidelines provided to assist PHAs to comply with HUD regulations regarding public housing and Section 8 housing utility assistance are similar.[14] While establishing a thirty percent rent burden mandate, "rent" is not defined in the statute.[15] HUD interprets "rent" to include the costs of shelter and reasonable utility consumption.[16] Because most Section 8 households and a large percentage of public housing households are now individually metered for some utility services, they pay utility bills directly to utility suppliers. Therefore, PHAs must set-off these utility costs against the mandated thirty percent of adjusted monthly income "rent" payment to the PHA.[17] This set-off is the utility allowance.

B. *The Utility Allowance System*

The amount of rent credit given to tenants for utility expenses is known as the utility allowance. Since 1985, HUD regulations have given PHAs almost unlimited discretion in setting local

---

[14] *Compare* 24 C.F.R. §§ 800–899 (1994) (Section 8 Housing Assistance Program); 886 (Special Allocations) *with* 24 C.F.R. §§ 900–999 (1994) (Public and Indian Housing); 965.470–.480 (PHA-Owned or Leased Projects—Maintenance and Operation: Tenant Allowance for Utilities). Unless otherwise indicated, subsequent cites to these sections or their subsections are to the 1994 edition of the Code of Federal Regulations.

[15] 24 C.F.R. § 813.102; 24 C.F.R. § 913.102.

[16] *See generally* 24 C.F.R. § 965.470–82 (outlining the mechanism for passing utility allowances from PHAs to individual tenants); *see also* U.S. PUB. HOUS. ADMIN., LOCAL PUBLIC HOUSING AUTHORITY MANAGEMENT HANDBOOK, pt. 2, § 9, 14 (1963) [hereinafter LOCAL PHA HANDBOOK].

[17] Historically, most Section 8 tenants have had individual meters. *See* URBAN SYSTEMS RESEARCH AND ENGINEERING, INC., RESEARCH AND EVALUATION REGARDING THE SECTION 8 HOUSING ASSISTANCE PROGRAM, REPORT ON SECTION 8 EXISTING FAIR MARKET RENTS—FINAL REPORT 5-7 (June 1977) [hereinafter URBAN SYSTEMS]. Although HUD originally encouraged master metering in public housing for heat and electric services, *see* Steven Ferrey, *Cold Power: Energy and Public Housing*, 23 HARV. J. ON LEGIS. 33, 51–56 (1986) [hereinafter Ferrey, *Cold Power*], after the 1970s "energy crisis," HUD promoted individual metering in new public housing construction, and Congress enacted statutes encouraging, and in certain cases requiring, the installation of or conversion to individual heat and electricity metering. *See* Public Utilities Regulatory Policies Act of 1978, Pub. L. No. 95-617, 92 Stat. 3117 (codified in scattered sections of 15, 16, 30, 42 U.S.C. (1988)). An unfortunate consequence of the government's pressure to retrofit individual meters to existing master-metered heating systems was that building owners and PHAs shut down master boilers and converted to individually metered electric resistance heating, which typically costs tenants twice as much as direct application fossil fuel heating such as gas or oil. *See generally* Ferrey, *Cold Power*, *supra*, at 51–56.

GAO now estimates that individually metered electricity predominates in both public housing and Section 8 housing programs. Natural gas individual metering is also a major metering configuration in both programs. GAO, VOL. 1, *supra* note 3, at 18.

utility allowances.[18] PHAs must maintain rent determination files for each assisted household, which include the methods used to arrive at household utility allowances.[19] Once PHAs establish allowances, HUD requires that PHAs review those allowances annually and revise them if required by regulatory standards.[20] HUD further requires PHAs to reexamine household income annually to recertify household eligibility for PHA occupancy and adjust household rental contributions in response to changing household income or circumstances.[21]

Substantial changes in rates charged by utility suppliers trigger an automatic PHA allowance review.[22] HUD also requires that its own field officers review the PHAs' methods and effectiveness in deriving and maintaining reasonable utility allowances to ensure compliance with the regulations.[23]

HUD does not require PHAs to account for individual building, unit, or occupancy characteristics when setting allowances for either program,[24] but rather requires only that PHAs set allowances for classes of housing units to reflect a reasonable use of utilities by "an energy-conservative household of modest circumstances."[25] Although HUD suggests that PHAs take into account the type of utility service provided (i.e., electric, gas, propane, oil) and its use (heating, non-heating, etc.), the housing structure, and the type and number of bedrooms in a unit when setting utility allowances, it does not require PHAs to do so.[26]

Formulating methods to gather utility consumption data in subsidized housing and establishing reasonable household energy consumption standards is a difficult task. PHAs usually gather consumption data and information from local PHA expe-

---

[18] 24 C.F.R. §§ 965.470–.480; *see also* Ferrey, *Cold Power*, *supra* note 17, at 77–81 (tracing the evolution of the utility allowance standards before 1985).

[19] 24 C.F.R. § 965.473(c).

[20] 24 C.F.R. § 965.478(a).

[21] 24 C.F.R. § 813.109(a); 24 C.F.R. 965.478(a).

[22] *See* 24 C.F.R. § 886.126 (requiring an allowance review for Section 8 housing whenever utility rates change "substantially"); 24 C.F.R. § 965.478(b) (requiring an interim review of utility allowances in public housing if rates change by a minimum of 10%).

[23] 24 C.F.R. § 965.473(d). However, GAO found that HUD's own staff did not adequately identify all deficiencies present in the PHA's methods for determining allowances at the 10 agencies it reviewed. GAO, VOL.1, *supra* note 3, at 57.

[24] 24 C.F.R. § 965.474.

[25] 24 C.F.R.§ 965.476(a). Comments on the proposed regulation argued that this provision was vague and unenforceable. *See, e.g.*, 49 Fed. Reg. 31,401–07 (1984) (reporting the comments of the Massachusetts Union of Public Housing Tenants).

[26] 24 C.F.R. § 965.474.

rience, if available, as the bases to establish utility allowances.[27] However, such data is not always readily available, making this method impractical for many PHAs.

For Section 8, this experience-based method is even more impractical due to the ever changing and diverse nature of participating private housing and the resulting limitation of precise current data.[28] To side-step these problems, PHAs typically base Section 8 allowances on utility consumption averages from the community at large. Where sample households are not available for a particular size unit, PHAs may extrapolate allowance estimates.[29] Therefore, Section 8 allowances often do not reflect actual utility use by Section 8 tenants. HUD deems a PHA Section 8 allowance scheme acceptable if it results in a majority of a PHA's Section 8 households receiving legally adequate allowances over any given twelve-month period.[30] However, HUD does not state an express standard of "adequacy" for the utility allowance.[31]

Other factors contribute to the imprecision of the utility allowance system in practice. First, PHAs of different sizes face different challenges in establishing and maintaining reasonable allowances based on identical HUD requirements.[32] Second, even when sufficient data indicate that an allowance is "reasonable," additional individual factors—such as unit placement within the building; efficiency of energy-consuming appliances; number, age and any special needs of occupants; and variations in tenant lifestyles—can create disparity between a uniform allowance and "reasonable" household consumption for a given tenant.[33] Third, while a household may experience an average *yearly* rent

---

[27] GAO VOL. 1, *supra* note 3, at 30.

[28] Other problems faced by PHAs in gathering data on Section 8 utility usage include confidentiality issues where utilities are individually metered and a lack of thorough recordkeeping by the PHAs.

[29] GAO, VOL. 1, *supra* note 3, at 34–35.

[30] *See* GAO, VOL. 1, *supra* note 3, at 41. However, GAO criticizes HUD for not enforcing this requirement. *Id.*

[31] HUD allows PHAs to grant additional relief to public housing tenants whose household expenses exceed their allowance using a standard of "reasonable grounds." 24 C.F.R. § 965.479. Individuals such as the elderly, ill, or handicapped often benefit from this discretion by the PHAs. There are no similar provisions for additional relief for Section 8 tenants. GAO, VOL. 1, *supra* note 3, at 35.

[32] *See* GAO, VOL. 1, *supra* note 3, at 35–36. Differences in the housing stock within individual PHAs, especially large PHAs with tens of thousands of units in hundreds of projects, pose difficulties in developing a workable standard to compute allowances. PHAs managing Section 8 units are especially burdened because of the greater variation in the size and characteristics of Section 8 housing.

[33] *See id.* at 36.

and utility burden of thirty percent of income, utility allowance schemes do not always account for seasonal changes in heating requirements.[34] Because most assisted households have limited disposable income after payment of their rent contribution, it is improbable that the households will put money aside to compensate for this utility expense fluctuation.[35]

Rental contributions from assisted tenant households and annual operating subsidies from HUD fund PHAs' operations.[36] Whenever a utility allowance increases, the PHA's, or in the case of Section 8 housing, the private landlord's, rental income declines proportionately.[37] Although HUD's operating subsidy is designed to cover the difference between PHA rental receipts and its estimated yearly operations budget, in some recent years, Congress has not fully funded estimated operating subsidies.[38] To increase their net revenues and cover their operating costs, PHAs may intentionally suppress utility allowances in order to inflate net rental payments due from tenants.[39]

Similarly, PHAs may minimize Section 8 utility allowances in order to encourage and maintain private landlord participation.[40] Landlords' rents are restricted to fair market rents (FMRs) set regionally by HUD.[41] Because utility expenses are included as part of the individually metered household's statutory thirty percent rent contribution for both Section 8 and public housing, the higher the utility allowance portion, the lower the remaining rental payment to the private landlord. When FMRs are perceived as too low, landlords are naturally discouraged from par-

---

[34] *Id.* at 47–48. Seasonal variations may lead households to be undersubsidized, that is, to consume energy in excess of the allowance amount, in some months and to be oversubsidized in other months. *Id.* at 28.

[35] *Id.* Although HUD regulations provide that PHAs may vary allowances throughout the year to compensate for seasonal fluctuations in utility expenses, it appears that few PHAs do. *Id.* However, it is worth noting that many utility companies allow households to guard against seasonal fluctuations through level payment, or "budget," plans. According to one survey, 17% of all households and 25% of households with a heavy energy burden use budget plans. U.S. DEP'T OF ENERGY, RESIDENTIAL ENERGY CONSUMPTION SURVEY: HOUSING CHARACTERISTICS 10 (1984).

[36] 24 C.F.R. § 990.101(a); 42 U.S.C. § 1437g(a) (1988). The Annual Contributions Contract between HUD and individual PHAs allows HUD to consider projections of income as well as past operating costs to determine the annual operating subsidy. 42 U.S.C. § 1437g(a)(3)(A) (1988).

[37] *See* GAO, VOL. 1, *supra* note 3, at 43.

[38] *Id.* (citing as an example that in FY 1989 only 97% of the estimated operating subsidy was funded).

[39] *Id.*

[40] *Id.* at 44.

[41] GAO, VOL. 1, *supra* note 3, at 44.

ticipation in the Section 8 program.[42] Some PHAs may compensate for the perceived low FMR by setting utility allowances lower than a reasonable amount, thus leaving a greater proportion of the FMR for the private landlord.[43]

C. *Metering Practices and Allowance Equity*

Utility metering configurations play a major role in the allowance system. Three utility metering alternatives exist: master, individual and check metering.[44] Master metering measures utility consumption through a single meter for the entire building; individual units are not metered.[45] All master-metered utilities are paid by the building owner, and tenants' rents are fixed at thirty percent of adjusted income with no utility allowance provided.[46]

Privately maintained check meters monitor individual units' consumption.[47] The building owner still pays the utility supplier for all utility costs[48] but then bills those costs to individual units that consume more than an established maximum utility allowance.[49] PHAs may provide rebates, or credits toward future rent or utility surcharges, if the household consumes less than its allowance.[50]

Individually metered households always receive a utility allowance.[51] These households are billed directly by utility suppliers, and any excess or deficiency in costs above or below the allowance amount causes deviation from each household's established thirty percent rent burden.[52] Therefore, maintaining realistic utility allowances in individually metered households is essential to a PHA's compliance with the rent burden mandate.[53]

---

[42] *See id.* (giving numerical analysis of this phenomenon).

[43] *Id.* GAO recommended correcting this incentive by having HUD adjust FMRs when they are too low rather than attempt to adjust utility allowances. HUD has not responded to this recommendation. *Id.* at 45.

[44] *See* Ferrey, *Cold Power*, *supra* note 17, at 46–47 nn.71–74.

[45] *Id.* at 46 n.71.

[46] *Id.*

[47] *Id.* at 47 n.72.

[48] *Id.*

[49] GAO, VOL. 1, *supra* note 3, at 45–46.

[50] 24 C.F.R. § 913.108.

[51] GAO, VOL. 1, *supra* note 3, at 12, 46.

[52] *Id.* at 78.

[53] *Id.* at 11. This conclusion is particularly true in cases in which assisted households cannot enroll in a budget plan with their utility supplier.

II. THE GAO STUDY: EMPIRICAL DOCUMENTATION OF SYSTEM ACCOUNTABILITY

HUD regulations give PHAs broad discretion to develop reasonable utility allowances.[54] Unfortunately, PHAs encounter significant difficulties in defining and estimating reasonable utility quantities for heterogeneous housing stock and diverse assisted household populations.[55]

In 1987, Congress directed GAO to conduct a study of this utility allowance system and report its findings to the Congress.[56] Specifically, Congress directed the GAO to report on how many households rely on utility allowances and how these allowances are provided, rent burdens actually incurred by these households, and options to improve the allowance system; GAO also was directed to determine how metering affects the utility allowance system and how the system could foster energy conservation.[57] After three years of effort, GAO reported its findings to the Congress in a 1991 study addressing allowances administered by PHAs for public housing and Section 8 housing certificate programs.

First, GAO collected data concerning how many and which PHAs provide utility allowances, the number of tenant households that receive allowances, and overall allowance-expense accuracy.[58] To obtain this information, as well as to determine PHA practices for deriving and reviewing allowances, GAO solicited questionnaire responses from approximately 1500 stratified and randomly selected PHAs nationwide and received an eighty-three percent response rate to its questionnaire.[59]

Second, to examine more closely how PHAs derived, monitored, and revised their allowances and their overall effectiveness in achieving desired household rent burdens, GAO gathered detailed financial records of tenant household incomes, utility costs, and allowances at six judgmentally selected PHAs.[60] GAO ran-

---

[54] 24 C.F.R. § 965.473(e).

[55] GAO, VOL. 1, *supra* note 3, at 34. This is because of differences among housing units in thermal characteristics, occupant behavior or conservation, and household characteristics.

[56] Housing and Community Development Act of 1987, Pub. L. No. 100-242, 101 Stat. 1815 (1987) (codified in scattered sections of 12, 42 U.S.C. (1988)) [hereinafter 1987 Act].

[57] *Id.* § 102.

[58] GAO, VOL. 1, *supra* note 3, at 14.

[59] *Id.* This represents a response rate of 90% of PHAs administering public housing and 46% of PHAs administering Section 8 housing.

[60] GAO reviewed household data at the following PHAs: City of Chandler, Housing

domly sampled twenty percent of public housing and Section 8 households in these six PHAs and collected twelve months of tenant income and utility allowance data from their files where such data was available.[61] In addition, GAO gathered utility cost data for these households from local utility companies and PHAs. Finally, GAO interviewed PHA and HUD officials responsible for devising and implementing allowance standards and policy for these PHAs and obtained pertinent office records and handbooks.[62]

A. *Results of the PHA Survey: Utility Allowances and Basic Human Needs in Assisted Households*

The GAO survey revealed a fascinating profile of the role of energy in subsidized housing. Utility costs represent as much as eighty-five percent of HUD operating subsidies and can constitute up to fifty percent of an individual PHA's operating budget.[63] However, despite the importance of utility allowances, prior to the GAO study, HUD did not know how many or which PHAs provide allowances.[64] GAO estimates that three-fifths of public housing households and four-fifths of Section 8 households receive utility allowances for at least one utility service.[65] Over eighty percent of PHAs administering public housing, and over ninety-five percent of PHAs administering Section 8 housing, employ a utility allowance program,[66] and electricity and natural gas are the most prevalent metered utilities they provide.[67] Metered utility services meeting the heating and/or cooling needs of assisted households rank as a primary concern of public hous-

---

Division (Arizona); City of Phoenix Neighborhood Improvement and Housing Dept. (Arizona); Cuyahoga Metropolitan Housing Authority (Ohio); Dakota County Housing and Redevelopment Authority (Minnesota); East Detroit Housing Commission (Michigan); West Memphis Housing Authority (Arkansas).

The selection of the six PHAs was neither random nor representative because several more troubled PHAs originally selected for the study did not participate. *Id.* at 15.

[61] *Id.*

[62] *Id.* In addition, GAO corrected errors it found in PHA data. *Id.*

[63] WAYNE SHERWOOD ET AL., COUNCIL OF LARGE PUBLIC HOUSING AUTHORITIES SURVEY OF ENERGY CONSUMPTION IN PUBLIC HOUSING AUTHORITIES, 1978–1981 at 1 (June 25, 1982); Evan Mills et al., *Deterrents to Energy Conservation Investment in Public Housing*, 11 ENERGY SYSTEMS AND POLICY 170 (1987).

[64] GAO, VOL. 1, *supra* note 3, at 14.

[65] *Id.* at 18.

[66] *Id.* at 3.

[67] *Id.* at 19.

ing and Section 8 tenants, who collectively constitute over one-third of all poor households in the nation.[68]

PHAs employ variant approaches in developing utility allowances,[69] and the elected approach may depend on the availability or accessibility of data.[70] When developing allowances for *both* their Section 8 and public housing programs, most PHAs utilize community-wide consumption data made available by local utility companies.[71] Some PHAs also use assisted households' actual consumption data as a basis for developing and evaluating allowances in both housing programs.

The somewhat more prevalent approach employed by PHAs when developing allowances solely for public housing uses households' actual consumption data derived from sampled public housing households.[72] Typically, the surveyed PHAs draw samples and extrapolate allowances for a class of public housing units from selected sample households' consumption.[73] Approximately two-fifths of PHAs also turn to this actual consumption method when developing their Section 8 allowances.[74]

However, the most prevalent approach employed for developing Section 8 allowances entails calculating average community-wide consumption from data supplied by utility companies and then choosing percentiles that represent reasonable consumption for particular housing types and unit sizes.[75] Unfortunately, this method employs extraneous data that may not reflect actual Section 8 housing types, locations, or tenancies of federally subsidized housing. Approximately one-half of PHAs employ this approach for developing public housing allowances as well.[76]

Overall, PHA approaches show no consensus in defining the reasonable consumption standard for either housing program; implementation of allowance standards is likewise inconsistent

---

[68] *Id.* at 22.

[69] *See, e.g., id.* at 40 (finding no clear consensus on what constituted "reasonable consumption").

[70] *Id.* at 34–35. *See supra* notes 29–32 and accompanying text.

[71] *Id.* at 40. Moreover, PHAs often discard high and low consumption values often from the database. *Id.*

[72] *Id.* at 72. It is worth remembering that the surveyed PHAs were among the most cooperative. Not all PHAs may have similar levels of administrative skill or resources. *Id.*

[73] *Id.* The consumption data is multiplied by applicable utility rates to determine the allowance value. *Id.*

[74] *Id.* at 73.

[75] *Id.* at 72.

[76] *Id.* at 73.

from PHA to PHA.[77] In addition, GAO discovered that the federal allowance framework is further destabilized due to widespread PHA non-compliance with HUD's annual review requirements.[78] GAO estimates from its survey that less than half of PHAs actually review their allowances annually as required by regulation and that between only thirty and forty percent conduct annual reviews in any given year.[79] Aggravating this situation, when PHAs actually conduct reviews, they frequently ignore other HUD requirements relating to methods of evaluation.[80] PHAs typically review and adjust allowances only when utility rates change, rather than gathering actual consumption data to match against established allowances and then evaluating allowance adequacy, as HUD regulations require.[81]

HUD also requires that its own field officers periodically review PHA allowances to monitor their compliance with PHA standards and regulations and provide guidance to PHAs regarding their allowance derivations and review policies.[82] HUD requires documentation of actual allowance calculations to evaluate both the documentation methods and mathematical accuracy and to determine the effectiveness of these methods in implementing HUD's "reasonable" consumption standard and in their overall compliance with federal housing law.[83]

However, GAO's study discovered that HUD does not itself document how PHAs either collectively or individually arrive at their allowances.[84] Thus, HUD does not know how PHAs define and implement the reasonable consumption standard it promulgates.[85] GAO concluded that HUD has not adequately defined

---

[77] *Id.* at 74 (noting that PHAs make many adjustments, which can be arbitrary or idiosyncratic, to the data they employ).

[78] *Id.* at 54.

[79] *Id.* From 1985 to 1989, 32% to 55% of PHAs reported reviewing public housing allowances in a given year. Between 38% and 65% of PHAs reported reviewing Section 8 allowances in a given year. *Id.*

[80] *Id.* at 55 (reporting that PHA reviews are often noncomprehensive and undocumented).

[81] *Id.* at 55–56.

[82] 24 C.F.R. §§ 901 (Public Housing Management Assessment Program); 965.473. GAO found that HUD typically schedules these reviews every four to eight years, or more frequently where warranted. GAO, VOL. 1, *supra* note 3, at 57.

[83] 24 C.F.R. § 965.473(d).

[84] GAO, VOL. 1, *supra* note 3, at 59 (reporting that some PHAs could not replicate how their allowances were derived, could not justify current levels, had not read tenant meters, or had left in place faulty meters).

[85] *Id.* at 41, 74. HUD staff did not focus on detailed reviews of important areas. *Id.* at 60.

the reasonable consumption standard in theory and does not oversee it in practice.[86]

B. *The Household Survey: A Closer Look at Household Utility Allowances Exceeding the Rent Ceiling*

GAO studied 9500 households in six participating PHAs[87] to determine their energy use characteristics, unit sizes and structures, utility allowances, actual utility expenses incurred, resulting rent burdens, and related information.[88] GAO found that two-thirds of the public housing households studied did not, in fact, pay thirty percent of their incomes toward rent and utilities, and a whopping ninety-three percent of Section 8 households studied had rent burdens that deviated from the thirty percent standard.[89]

Of these deviating households, most pay more than the thirty percent rent contribution mandate.[90] Approximately fifteen percent of all households in each housing program had notably excessive rent burdens, i.e., burdens exceeding thirty-three percent of adjusted income for public housing and exceeding forty percent of adjusted income for Section 8.[91] The average rent

---

[86] *Id.* at 12, 60.

[87] The ability to draw strong inferences from GAO's household study is tempered by a variety of survey problems that affected GAO's judgmental sampling of PHAs during the period between July 1988 and April 1990. *See generally id.* at 14–17. GAO originally selected 10 PHAs to review in detail how they derived, monitored, and revised their respective utility allowances only to discover that many PHAs neither kept centralized accounts of allowances they provide nor collected data on households' actual utility costs. *Id.* at 14. As a result, the GAO collected this primary data that should have been routinely maintained by PHAs.

Moreover, administrative problems at several PHAs forced GAO to drop four of the 10 PHAs selected for its detailed review. *Id.* at 15. These four had been designated in the GAO selection process as representative of PHAs that likely faced impediments administering utility allowance systems. (Interestingly, Los Angeles, Philadelphia, and Boston, the three largest urban PHAs of the 10, were among those dropped from the rent burden study.) Thus, the household survey took place only in the high-compliance, "model" PHAs in the sample. Since the number of households comprising the GAO study were derived from a judgmental sample of six PHAs, the results are not statistically representative of the entire PHA or Section 8 populations. *Id.* at 15–16; *see also infra* notes 100–101.

Moreover, these questionnaire responses are not subject to any on-site GAO verification. When on-site, GAO discovered many problems with PHA administration. Although GAO performed extensive verification of income calculations, utility allowances, and rents received from the PHAs, these were not independently certified. Finally, the study encountered some problems with missing data.

[88] GAO, VOL. 1, *supra* note 3, at 2–3.

[89] *Id.* at 3.

[90] *Id.* at 24–25.

[91] *Id.* at 3–4. The 20 worst cases had an average rent burden of 74%. *Id.* at 4.

burden for the approximately 4500 public housing households studied varied from the thirty percent standard by only 0.5 percent.[92] However, the approximately 5000 Section 8 households reviewed averaged an overall rent burden of thirty-six percent, significantly in excess of the statutory rate.[93] Furthermore, because the six PHAs surveyed represent the more cooperative PHAs,[94] and perhaps those most compliant with HUD regulations generally, these findings may overstate the degree to which most PHAs meet the thirty percent standard.

When reviewing these findings, one must realize that Section 8 and public housing populations differ in their average annual incomes. For the households studied by GAO, public housing households averaged $454 in total monthly income, while Section 8 households' income averaged $544 per month.[95] Average utility allowances were $55 per month and $64 per month for public housing and Section 8 households, respectively.[96] After deduction of total rent contributions, public housing households were left on average with $340 per month in disposable income, while generally higher income Section 8 households fared slightly better, with $380 per month in disposable income.[97]

The GAO study indicates that allowances take on greater or lesser significance, depending upon household income and utility services covered by allowances.[98] The lower the household's income and the more the utility allowance framework provides its energy and other services, the greater its reliance on the utility allowance as a fraction of its income and the greater its risk of undue financial hardship if utility assistance proves inadequate.[99]

PHA records do not differentiate household income according to the number of household members supported by that income. Presumably, the larger the household size, the greater its relative hardship when actual utility costs in excess of provided utility expenses deplete disposable income. The GAO study found

---

[92] *Id.* at 3. However, average rent burdens tend to mask differences in actual rent burdens. Deviations from the average can range from 6 to 60% within selected PHAs. *Id.* at 26.

[93] *Id.* at 3, 24.

[94] *See supra* note 87 (discussing survey problems encountered by GAO).

[95] GAO, VOL. 1, *supra* note 3, at 24–25.

[96] *Id.* at 20.

[97] *Id.* at 24–25.

[98] *Id.* at 20, 22.

[99] *Id.* at 27–28.

significant variation in the number of household members occupying similar size units.[100] Therefore, because allowances are differentiated by the number of bedrooms in a unit, they reasonably cover only the *expected* number of occupants. Assisted households' utility consumption logically increases with each additional household occupant, and the allowance-expense disparity grows proportionately with increases in occupancy.[101]

Not suprisingly, the GAO study also shows that households in which tenants are at home during days consume more energy than those whose members leave the home during the day, that PHAs see some tenants as more energy conscious than others, and that some tenants use major appliances that are not included in establishing the utility allowance calculation.[102] All of these factors, in addition to those already discussed, make setting generic allowances for diverse housing stock and populations very difficult and lead to widespread allowance-expense deviations. PHAs and HUD indicated little surprise at the preliminary GAO report results demonstrating significant allowance-expense disparity in the six PHAs studied.[103]

GAO did not study energy consumption or conservation behavior in the households studied, and the GAO report does not address such consumption and practices. However, this behavior is critical in designing effective allowances.

Most PHAs surveyed factored out what they considered extreme energy consumption in calculating their utility allowances.[104] Three of the ten PHAs selected for detailed review deleted the upper and lower ten percent of energy users from their allowance com-

---

[100] *Id.* at 38.

[101] *Id.* This conclusion is less true for heating and cooling utility consumption, which varies primarily with unit efficiency and the volume of space to be conditioned.

[102] *Id.* These additional appliances typically include space heaters, freezers, and air conditioners. *Id.*

[103] Specifically, HUD and PHAs indicated that they believed Section 8 deviations often stemmed from air conditioning and other luxury appliance costs where these energy uses are not included in the households' utility allowance calculation. Some PHAs generally responded that wasteful household consumption practices contributed significantly to allowance-expense deviations in most situations. However, GAO identified no empirical support for either HUD's or the PHAs' attribution of causation. *Id.* at 38–39.

[104] These PHAs made some determination of allowances deemed to be reasonable without regard to "excess" and "under-consumers," usually setting their allowances somewhere between the two energy consumption extremes. The report does not indicate how these PHAs arrive at these judgments. *Id.* at 40.

putations.[105] This deletion tends to lower the resulting allowance.[106]

One of these PHAs, after taking this step, then compared current-year consumption averages with prior-year consumption averages to lessen the effect of the disparity created by yearly weather fluctuations. The GAO study did not determine the prevalence of PHAs that modulate allowances based upon seasonal and yearly temperature fluctuations. HUD regulations provide that public housing allowances may vary over the year to reflect seasonal variations in utility requirements,[107] but none of the six PHAs studied differentiated their allowances over the year.[108]

The problems with provision of reasonable utility allowances heighten the importance of agency oversight. GAO solicited PHA estimates of their own accuracy in establishing utility allowances that closely matched actual utility expenses incurred by assisted households.[109] Overall, surveyed PHAs believed that, for any given month, allowances matched actual expenses in forty-three percent of cases for public housing and fifty-six percent of cases for Section 8 housing.[110] To the contrary, GAO found over the twelve months of its study of six of the best PHAs that, for twenty-seven percent of public housing and sixty percent of Section 8 households reviewed, there was not even a single month where the households' rent burden was thirty percent.[111]

In short, because most PHAs do not collect actual consumption data on their own households, as required by HUD regulations, PHA estimates of compliance are both overly optimistic and unsubstantiated.[112] HUD's lack of oversight of PHA utility practices renders noncompliance more problematic.[113]

---

[105] *Id.*

[106] This is because of the skew to the right of the distribution of utility consumption. *See* Appendix.

[107] 24 C.F.R. § 965.476.

[108] GAO, VOL. 1, *supra* note 3, at 39–40. Similarly, tenants paying bills directly to a utility supplier often have the option of entering a budget agreement with the supplier to spread utility costs evenly across the year. Such a level-payment plan would considerably facilitate effective, even distribution of utility expense payments where utility allowances are annually calculated and spread evenly across the 12 calendar months. *Id.* at 36–37.

[109] *Id.* at 25.

[110] *Id.* at 26.

[111] *Id.* at 30.

[112] *Id.* at 59.

[113] *Id.* at 59–60.

C. *Disparity in Treatment of Checkmetered Households*

One of Congress's concerns leading to the GAO investigation was doubt that households under different metering configurations received similar regulatory treatment for similar consumption or conservation behavior.[114] Congress's directive to GAO asked it to break down its rent burden findings into separate categories for each of these metering configurations.[115] However, because GAO determined that gathering data in this way was not feasible through its questionnaire study and PHA file reviews,[116] the metering impact on utility allowances remains unresolved.

GAO estimates, through its PHA questionnaire responses, that about 319,000 public housing households in 1186 PHAs across the country have checkmeters, representing about fifty percent of all public housing households surveyed.[117] The questionnaire responses demonstrated that checkmeters are almost nonexistent in Section 8 housing.[118] Checkmetered public housing households, therefore, represent approximately fifteen percent of the combined 2.4 million households assisted by the Section 8 existing and public housing programs nationwide.

Of these checkmetered households, the GAO questionnaire responses show that seventy-five percent are treated by their PHAs as if they were individually metered.[119] In these cases, the study gives the household its full allowance amount and no credit or rebate for consumption below that allowance amount.[120] The household, like individually metered households, pays out-of-pocket for excess consumption, but in this instance the PHA, rather than the utility supplier, charges the checkmetered household for that amount.[121] Only six percent of PHAs claim to provide a credit to a checkmetered household for any unused portion of its utility allowance.[122] This credit is applied in-kind

---

[114] *Id.* at 77. This includes allowance levels, adjustments, and conversation incentives. Also, the GAO study indicates that Congress had expressed particular concern over the treatment of checkmetered tenants. *Id.*

[115] *Id.* at 17.

[116] *Id.*

[117] *Id.* at 46 (noting that the total included 176,651 public housing households checkmetered for electricity and 142,355 checkmetered for gas).

[118] *Id.* at 45 n.13.

[119] *Id.* at 46.

[120] *Id.*

[121] *Id.* at 45.

[122] *Id.* at 46.

against future months' shelter rent, utility costs, or unrelated tenant expenses, rather than rebated in cash.[123]

Twenty-one percent of PHAs responding to the GAO questionnaire do not employ a credit or rebate policy for checkmetered households.[124] The household receives no compensation for its utility "under-use" or conservation. Thus, the portion of the checkmetered allowance that the household does not consume accrues to the benefit of the PHA.

HUD regulations do not require a checkmeter rebate policy,[125] and the GAO study does not quantify the effect of this inequitable treatment of the checkmetered segment of the assisted housing population.[126] However, GAO concludes that implementing a mandatory credit-rebate policy for checkmetered households would entail administrative costs that exceed the benefit potentially gained by households from equalizing treatment between individually metered and checkmetered households.[127]

D. *Mistakes and Constraints Associated with the Present System: The Unchecked Utility Allowances*

PHAs do not and cannot effectively evaluate the reasonableness or equity of their own utility allowances because there is no precise regulatory standard against which to measure performance.[128] Because PHAs do not comply with HUD regulations and HUD does not make them do so, PHAs lack collected actual utility consumption data for federally assisted tenant households. This lack of data precluded GAO from assessing the added cost to the federal government were HUD to adopt a policy that shifts more of the burden of assisted households' utility expenses directly to the government.[129]

This shift could occur in two ways. First, Congress or HUD could institute federal policy that requires physically converting assisted housing predominantly to master metering where feasi-

---

[123] *Id.*

[124] *Id.* at 46, 77 (Table II.1, note a).

[125] "HUD regulations and handbooks do not" specify procedures for "providing rebates to check-metered households that consume less than their allowance." *Id.* at 46; *see also* 24 C.F.R. § 965.407.

[126] GAO, VOL. 1, *supra* note 3, at 46.

[127] GAO estimates that each household in question would save only about $30 per year. *Id.* at 47.

[128] *Id.* at 40; 24 C.F.R. § 965.473. *See also supra* note 77 and accompanying text.

[129] GAO, VOL. 1, *supra* note 3, at 65.

ble. This would standardize metering types and financial responsibility. Alternatively, the government could pay assisted households' utility costs directly to retail utility suppliers.[130] This latter approach would alleviate rent burden discrepancies due to different metering configurations in assisted housing. However, it could raise federal administrative costs substantially.[131]

On the other hand, HUD's present utility policy, if more precisely defined, implemented, and enforced, could also perform equitably and promote savings to all parties through greater energy efficiency.[132] The remainder of this Article explores means by which HUD could improve its present utility allowance system to reflect more accurately actual household energy use and to fulfill more completely the thirty percent rent burden standard.

III. METER CONFIGURATIONS

Meter configurations are critical, yet often invisible, factors in the allocation of tenant subsidies and the realization of energy efficiency goals. The three basic types of tenant utility meters are master meters,[133] checkmeters or submeters,[134] and individual meters.[135] The builder or developer usually selects the meter type at the time of construction.

---

[130] *Id.* at 64–65.

[131] *Id.* at 65.

[132] *Id.* at 62.

[133] With master meters, the utility supplier meters utility service to the building through one meter and does not meter individual units. Thus, there is no means to bill tenants for their individual utility consumption. However, some owners allocate a percentage of the total master-metered bill to tenants based on the tenants' percentage of total floor space. *See generally* IBS, ENCOURAGING ENERGY CONSERVATION IN MULTIFAMILY HOUSING: RUBS AND OTHER METHODS OF ALLOCATING ENERGY COSTS TO RESIDENTS (1980). Tenants are not responsible for paying the utility supplier and generally cannot have their services terminated even if they do not pay their rent or if the building owner fails to pay the utility supplier.

[134] In submeter, or checkmeter, systems, the building owner maintains a private, unofficial meter on each individual dwelling unit and bills tenants for individual usage. Utility service to the building is still master-metered by the utility supplier; the sum of submeter usage plus usage for common spaces in the building should equal the master meter billed total.

Although one state differentiates a checkmeter as a private system used for "billing" tenants for energy use but not for "resale" of energy to tenants, this distinction is merely semantic; in either case, the building owner is compensated for energy purchased wholesale and reallocated to tenants.

[135] With individual metering, or retail service, the utility supplier meters each dwelling unit and bills tenants directly. The supplier also meters common areas and bills the building owner for usage there. Otherwise, the building owner bears no responsibility for tenants' energy consumption and bills; instead, each tenant must post

Although HUD originally encouraged the installation of master meters,[136] during the 1970s, rising energy prices prompted more building owners to select individual meters for new construction because such meters shift the risk of rising energy prices from landlord to tenant.[137]

In addition to these market forces, regulation by the federal government also influenced the selection of metering systems for new construction. For example, as part of the Public Utility Regulatory Policies Act of 1978 (PURPA),[138] the federal government required states to consider prohibiting electric master meters in newly constructed multifamily buildings.[139] The statute required the use of individual meters in cases where the occupant of each unit has control over the use of electric energy and the long-run benefits of individual metering appear to exceed the costs of separate meters.[140] Despite this seemingly narrow application, many states have used this standard as a general springboard for individual metering in new construction.

However, the trend towards individual metering and energy efficient designs did not benefit the existing public housing stock. Most of the public housing stock and half the general multifamily stock was built before 1960. Only five percent of the public housing stock and fifteen percent of the multifamily stock was built after 1975.[141]

---

a security deposit, if required, pay any late payment fees, and risk termination for nonpayment.

[136] *See* Ferrey, *Cold Power*, *supra* note 17, at 54 n.101.

[137] The limited data available suggest that the proportion of master-metered units in all new construction is declining each year. For example, of new multifamily units (housing five or more) completed annually, the percentage with master-metered electricity declined from 27% in 1973 to 9% in 1977. *See* Bureau of the Census, *Utilities and Heating Fuel Included in New Apartment Rents: 1973 and 1976*, CURRENT HOUSING REPORTS (July 1978); BUREAU OF THE CENSUS, MARKET ABSORPTION OF APARTMENTS: ANNUAL 978 ABSORPTIONS, REPORT NO. H-130-78-5 (May 1979). For space heating, the percentage of master-metered new construction declined from 61% to 30% between 1973 and 1977. *See* Ferrey, *Cold Power*, *supra* note 17, at 54 n.101.

[138] Public Utility Regulatory Policies Act of 1978, Pub. L. No. 95-617, 92 Stat. 3117 (codified in scattered sections of 15, 16, 30, 42, and 43 U.S.C. (1988)).

[139] *See* 16 U.S.C. § 2623 (1988). After inital legal challenges, this provision was upheld by the Supreme Court in FERC v. Mississippi, 456 U.S. 742, 771 (1982).

[140] 16 U.S.C. § 2627 (1988).

[141] For a more complete discussion of the incidence of various metering types, see generally BUREAU OF THE CENSUS, ANNUAL HOUSING SURVEY 1978–1979 (1981 & 1982); U.S. DEP'T OF ENERGY, RESIDENTIAL ENERGY CONSUMPTION SURVEY: 1979-80 CONSUMPTION AND EXPENDITURES; KATHLEEN M. GREELY ET AL., ANALYZING ENERGY CONSERVATION RETROFITS IN PUBLIC HOUSING 1 (1986) (study performed by the Lawrence Berkeley Laboratory for the Department of Energy).

Meter configurations play a substantial role in the micro-level financial allocations among HUD, PHAs, and tenants. But tenants rarely make a conscious choice of the metering system they prefer and thus face a variety of "hidden" effects. For instance, in a checkmeter system, landlords may be allowed to "mark up" the energy they resell. With individual meters, tenants may gain legal rights not available to them under other metering systems[142] but may face a dollar diversion away from them.[143]

Individual meters may benefit PHAs and HUD at the expense of the tenants. With individual metering systems, the PHA and HUD can provide lower utility allowances regardless of whether or not the system is successful in motivating energy conversation by tenants. Furthermore, the PHAs and HUD are not affected by the legal rights gained by tenants, because those legal rights relate only to the relationship between tenants and the utility supplier.

While these relationships are complicated, three general observations may be useful to policymakers. First, the legal protections that are not present with checkmetering could be restored at the administrative level by HUD or by the PHAs, thus neutralizing the hidden effects on the tenant of metering choices. Second, tenants rarely choose their metering system and may not be aware of their total energy usage. Finally, to the extent that variations in tenants' utility consumption reflects factors beyond their control, individual metering possibly exacerbates, rather than minimizes, unreasonable variations in tenant utility costs.

IV. Creating Meter-Neutral Utility Allowances

A. *The Development of the Current Public Housing Utility Allowance*

The utility allowance determines the percentage division in utility costs between tenants and the government, as well as price-induced utility conservation incentives. Meters have always figured prominently in determining the allowance level for

---

[142] State utility regulatory commissions provided a variety of due process rights to retail consumers. *See, e.g.*, Mass. Gen. L. ch. 164, § 124A–124H (1976 & Supp. 1994); *see generally* Steven Ferrey, Law of Independent Power ¶ 5.01[2] (6th ed. 1994).

[143] GAO, Vol. 1, *supra* note 3, at 47–48.

particular units—explictly in previous federal requirements and implicitly yet today. For purposes of determining whether the thirty percent cap on rent as a fraction of tenant income is met, HUD has always defined "rent" to include a "reasonable" quantity of utility service.[144] At different times, a reasonable amount of utility service was determined by the type of utility meter serving the units and by whether the tenants were elderly households or low-income families.[145]

1. Pre-1980 Allowances

In its 1963 guidance for PHAs, HUD reasoned that where tenants were submetered, they should be "surcharged" for excess consumption above a reasonable quantity of utility service supplied free to the submetered tenants.[146] HUD suggested that the "free" quantity be established at the value representing average consumption plus twenty percent.[147] The rationale for this standard was that this guidance would insure that at least five percent, but no more than twenty-five percent, of these tenants would be surcharged.[148] The surcharges could be collected and retained by the PHA.[149]

In master-metered units, HUD allowed a flat-rate surcharge based on the possession of certain unauthorized appliances.[150] Allowances were revised "whenever any substantial changes are made in the rent schedule."[151]

The results of this system were uneven. In some cases, utility allowances were ample; in others, inadequate. In some PHAs, the utility allowance for certain months left every tenant unreasonably compensated for utility costs.[152] Some utility allow-

---

[144] For the most recent statement of this policy, see 24 C.F.R. § 913.402; 49 Fed. Reg. 31,400 (1984) (describing this policy). *See also* LOCAL PHA HANDBOOK, *supra* note 17, at 14.

[145] *See* Ferrey, *Cold Power*, *supra* note 17, at 40–47.

[146] *See* LOCAL PHA HANDBOOK, *supra* note 17, at 14.

[147] *Id.* at 14.

[148] *Id.* at 15.

[149] *Id.* at 17–20.

[150] *Id.* at 20.

[151] *Id.* at 21. Today, most public housing authorities no longer maintain rent schedules. Rather, they establish tenant rents at 30% of each tenant's adjusted income. *See* 42 U.S.C. § 1437a (1988); GAO, VOL. 1, *supra* note 3, at 21.

[152] JAY M. STEIN, DETERMINING UTILITY ALLOWANCES FOR U.S. PUBLIC HOUSING TENANTS § 4.60 (1979).

ances did not even cover the cost of the electricity necessary to operate the refrigerator the PHA had supplied to tenant units.[153]

2. 1980–1985 Allowances

In 1980, HUD promulgated new utility allowance formulas.[154] The new regulations allowed separate allowances for different types of structures[155] and different sized units.[156] PHAs had the option to establish a distinct allowance category to reflect the unique thermal qualities of a particular unit in the building.[157] Quarterly rather than monthly allowances were preferred, but not required.[158]

This regulation did not affect the previously existing system for master-metered tenants.[159] Tenants paid thirty percent of adjusted income for rent and received without charge all of the utility service that they consumed.[160] The PHA could charge tenants for the possession, but not the use, of household appliances that it deemed "unauthorized." An "unauthorized" appliance could be any energy consuming appliance not supplied by the PHA.[161]

For individual metering, a tenant's contract rent payment was reduced by the cash amount of the utility allowance,[162] which was based on the consumption mean calculated from records of past utility consumption for similarly sized units in the project.[163] Individually metered tenants paid out-of-pocket for utility consumption exceeding the value of their allowances. Energy conservation resulted in direct cash savings to the tenant, while the tenant was penalized for profligate behavior.

---

[153] *Id.*

[154] 24 C.F.R. § 965.474 (1984); *see also* 45 Fed. Reg. 59,505 (1980) (promulgating these utility allowance regulations).

[155] "Separate allowances shall be established for each utility and for each category of dwelling units within structures which are reasonably comparable as to age and construction type, have the same utility combination and the same type of major equipment." 24 C.F.R. § 965.474(a) (1984).

[156] 24 C.F.R. § 965.474(c) (1984).

[157] *Id.* For example, a top-floor corner unit with two walls and a ceiling exposed to the exterior on the north side of the building will consume twice as much winter heating energy to maintain an identical temperature as the same-sized unit that is on the south side of a middle level of the building and not a corner unit, even if each unit's tenant behaves identically. *See* Ferrey, *Cold Power*, *supra* note 17, at 79.

[158] 24 C.F.R. § 965.475(a) (1984).

[159] See GAO, VOL. 1, *supra* note 3, at 47–48.

[160] 24 C.F.R. § 965.470–.480 (1984).

[161] 24 C.F.R. § 965.473(b) (1984).

[162] 24 C.F.R. § 965.475(b) (1984).

[163] 24 C.F.R. § 965.476 (1984).

Submetered tenants paid thirty percent of adjusted family income for rent and received a specified quantity of utility service as their utility allowance. This quantity was established at the ninetieth percentile of actual consumption for dwellings of a given size in each housing project.[164] Checkmetered tenants consumed energy up to the utility allowance amount at no additional cost to themselves. No rebates were provided for conservative consumption;[165] but if a household consumed more than its utility allowance, the PHA could surcharge the tenant for the excess consumption.[166] Under this regulation, about ten percent of consumption was surcharged.[167] The surcharge was computed at the average per unit cost of energy to the PHA through its master meter.[168] The allowance had to be adjusted upward to an appropriate amount if it resulted in more than twenty-five percent of tenants surcharged in any billing period.[169]

The regulations for all three meter types required that allowances be computed based on actual PHA project data from the prior three-year period.[170] If this was not readily available, the prior year or two years were used.[171] With new housing, or after a meter conversion or fuel source change in existing housing, a new estimated utility allowance was established by reference to comparable housing in the locality.[172]

This scheme contained several protections for tenants. First, PHAs could not surcharge tenants or force them to pay out-of-pocket for excess consumption due to management failure to make necessary repairs or to remedy a defect in the structure or heating system.[173] Second, the regulations required PHAs to provide counseling, dwelling efficiency modifications or correc-

---

[164] 24 C.F.R. § 965.477 (1984).

[165] *See* 24 C.F.R. § 865.477 (1983). As originally proposed, the regulations would have provided rebates for consumption; surcharges would be assessed for consumption in excess of 15% of average consumption. *See* 44 Fed. Reg. 1600–03 (1979).

[166] 24 C.F.R. § 965.479 (1984).

[167] 24 C.F.R. § 965.477 (1984).

[168] "The amount of the Surcharge for each block [of excess consumption] shall be computed by applying the Utility Supplier's average rate to the amount of the excess." 24 C.F.R. § 965.479 (1984).

[169] 24 C.F.R. § 965.480(b) (1984).

[170] 24 C.F.R. § 965.476(a) (1984) (further specifying that each year must consist of 12 consecutive months for purposes of determining utility allowances).

[171] *Id.*

[172] 24 C.F.R. § 965.476(b) (1984).

[173] 24 C.F.R. § 965.481(a)(2) (1984).

tive maintenance for any tenants with excessive consumption.[174] Finally, the regulations required PHAs to revise the allowance for individually metered tenants whenever utility rates increased cumulatively by ten percent or more,[175] and PHAs could consider increases for individual tenants whenever consumption exceeded the established allowance by twenty percent or more in any billing period.[176]

3. Post-1985 Allowances

After April 7, 1985, PHAs were permitted to establish utility allowances on any defensible basis.[177] The stated goals of these regulations were (1) to place all tenants on relatively equal footing, (2) to remove the significance of metering type in the determination of allowances, and (3) to provide incentives for energy conservation by tenants.[178] After these regulatory changes, the PHAs, rather than HUD, became totally responsible for data, methodologies, and calculations used to establish tenant utility allowances and surcharges.[179] In addition, the requirement to base allowances on past consumption data was eliminated, although PHAs still can use such data if they choose.[180]

In the preamble to its new regulations, HUD identified nine dwelling- or tenant-specific factors as relevant in establishing allowances.[181] Under the current regulations, PHAs also must set allowances that reflect both appliance efficiency and the physical characteristics of dwellings.[182] Furthermore, the existing standard of a "reasonable allowance" was redefined as an allowance based on a "reasonable consumption of utilities by an energy-

---

[174] 24 C.F.R. § 965.481(c) (1984).
[175] 24 C.F.R. § 965.480(c)(1) (1984).
[176] 24 C.F.R. § 965.481(a)(3) (1984).
[177] 24 C.F.R. §§ 965.473, .476 (1985); *see* 49 Fed. Reg. 31,399 (1984) (making this "deregulation" effective).
[178] 47 Fed. Reg. 35,251 (1982) (giving notice of the proposed rulemaking on "deregulating" utility allowances).
[179] 24 C.F.R. §§ 965.473, .477.
[180] 24 C.F.R. § 965.476(c)(1).
[181] 49 Fed. Reg. 31,401, 31,409 (1985). The nine factors include the equipment and function covered, climatic location, size of dwelling and number of occupants, type of construction and design, the energy efficiency of supplied appliances and equipment, the consumption of reasonable tenant-supplied appliances, the physical condition and energy efficiency of the unit, interior temperature requirements, and the standby temperature of hot water heating equipment. *Id.*
[182] 24 C.F.R. § 965.476(d).

conservative household of modest circumstances."[183] No longer did individually metered tenants necessarily receive lower allowances than other metered tenants.[184]

The current regulations retain several features of the pre-1985 regulations. Individual relief is still allowed to elderly, ill or handicapped tenants on a case-by-case basis.[185] HUD still suggests that allowances be reviewed annually by the PHA,[186] but requires an interim review only if utility rates change by ten percent or more.[187] PHAs must give tenants notice and an opportunity for comment.[188]

Although the new regulations permit all of the changes discussed above, PHAs can continue operating under the prior federal utility allowance system if they so choose.[189] Therefore, a mix of utility allowance methodologies now exists at the local level.

B. *The Section 8 Utility Allowance*

Utility allowances in Section 8 housing, though serving a purpose similar to public housing utility allowances, are fundamentally different in their application because Section 8 stock, unlike PHA housing, is drawn from the free market. The system of controls exercised by PHAs over Section 8 housing works a particular pressure on utility allowances in Section 8 housing.

Every administering housing authority, subject to HUD approval, establishes a fair market rent (FMR)[190] for Section 8

---

[183] 24 C.F.R. § 965.476(a).

[184] 49 Fed. Reg. 31,401–02 (1985).

[185] *See* 24 C.F.R. § 965.479 (requiring that tenants be notified of special relief provisions and the criteria for granting such relief in the notice provided pursuant to 24 C.F.R. 965.473(c)).

[186] 24 C.F.R. § 965.478(a). The failure of this review was noted in the GAO report. *See* GAO, VOL. 1, *supra* note 3, at 54–56; *see also* part II, *supra* notes 54–132 and accompanying text (discussing the GAO study).

[187] 24 C.F.R. § 965.478(b). Any increase in allowances due to rate increases is now retroactive to the effective date of the rate increase. (This provision replaced the old requirement of allowance revision whenever more than 25% of tenants were surcharged. *See* 24 C.F.R. § 965.480(b) (1984); 49 Fed. Reg. 31,406 (1985).).

[188] 24 C.F.R. § 965.473(c) (requiring PHAs to notify tenants at least 60 days prior to any change in allowances and as provided in tenant leases, and, further, to give tenants information on what appliances are included in the calculation of the utility allowance).

[189] 24 C.F.R. § 965.470–.480.

[190] 24 C.F.R. § 888.101.

housing units in its community or region[191] that then acts as a rent ceiling.[192] The difference between the statutorily determined tenant rent payment[193] and the actual rent charged by the property owner, up to the FMR, is covered by a HUD Section 8 voucher subsidy.[194]

Where a unit is master-metered, all utilities are included for the tenant in return for paying the rent.[195] Therefore, assuming other factors are equal, tenants in master-metered units will pay a higher rent than would tenants in equivalent units without master meters. This effect increases rent on some master-metered units above the FMR levels and disqualifies these units from participation in the program, driving tenants with Section 8 entitlements to seek out less expensive, lesser quality housing. Approximately one-fifth of the Section 8 units are master-metered and, therefore, may face this problem.[196]

Individually metered Section 8 units account for the remaining four-fifths of Section 8 housing.[197] (There are few documented submetered units in Section 8 housing.[198]) Unlike master metering, individual metering allows the administering PHA to manipulate or adjust the utility allowance for purposes unrelated to reasonable utility needs.

Certain key differences between Section 8 and public housing arise as a result of the differential prevalence of individual metering in Section 8 housing and master metering in public housing. In most PHA public housing, the tenant pays one sum, statutorily set at thirty percent of income, for both rent and utilities. Thus, the tenant's rent obligation in public housing is easily disaggregated from the utility expenditures necessary to maintain that housing. Tenant utility allowances are not affected by PHA operating deficits or PHA operating expenses. While the utility allowance in public housing may be determined differently based on a variety of factors including metering type, there is no explicit tension between expenditures on the utility allowance and the amount left to subsidize PHA operations.

---

[191] 24 C.F.R. § 888.113.
[192] *See, e.g.*, 24 C.F.R. § 883.305; 24 C.F.R. § 881.204.
[193] 24 C.F.R. § 813.107.
[194] 24 C.F.R. § 887.353.
[195] *Id.*
[196] GAO, VOL. 1, *supra* note 3, at 18.
[197] *Id.*
[198] There is little submetering in Section 8 housing, which reflects the general predominance of multifamily stock. *Id.* at 19 (Table 2.2).

By contrast, only a finite amount of Section 8 subsidy is available for rent and utilities, and these two uses compete for the allocated resources. As a result, Section 8 utility allowances are less likely than public housing utility allowances to relate to actual utility usage or to the energy requirements in a particular building or unit. Administering PHAs decide which competitive factor in the zero-sum allocation—the utility allowance or the rental subsidy—receives preference. Since FMRs are fixed regionally by HUD, the amount of the utility allowance is the primary variable within the control of the PHA, and the PHA's decision determines how much residual subsidy is left to pay private landlords.

PHAs can bring more individually metered units under the designated FMR for the Section 8 program by scrimping on the amount of the utility allowance. One report concluded that some PHAs deliberately suppress the utility allowance component of the FMR[199] in order to effectively raise the rent available to the private market landlord who participates in the Section 8 program. For a PHA experiencing difficulty attracting enough units with rent priced under the FMR ceiling, this technique attracts additional units within program parameters.

The obvious drawback to this technique, however, is that it establishes an inadequate residual utility allowance under the FMR.[200] Tenants with individual metering often pay more for their utilities than they are supposed to, raising their total shelter expenses for rent *plus* utilities above the statutory thirty percent of income for Section 8 housing.[201]

In effect, where individually metered Section 8 utility allowances are suppressed, a largely hidden transfer payment from tenants to private market landlords results. Were this suppression of utility allowances not to occur, the transfer of federal funds would flow normally from HUD to individually metered tenants, and ultimately to their utility suppliers. Unlike master or submetered tenants, individually metered tenants must make up out-of-pocket this lost transfer payment to their individual utility suppliers.

---

[199] URBAN SYSTEMS, *supra* note 17, at 5-4.

[200] GAO found that over half of the Section 8 units it studied did not receive legally adequate utility allowances. GAO, VOL. 1, *supra* note 3, at 3.

[201] *Id.*

Both older and more recent studies conclude that this concern is valid. Survey evidence from 1977 suggested that three-quarters of Section 8 tenants received an inadequate utility allowance.[202] Almost one-third of this group paid twice the utility costs they should have borne according to HUD's Section 8 regulations.[203] The vast majority of Section 8 utility allowance schedules provided less than actual utility requirements,[204] and these values changed little by the time of the GAO's 1991 report.[205]

In short, differences in metering lead to two forms of discrimination in determining utility allowances. First, master-metered tenants are guaranteed reasonable utility quantities as part of their rental payments; individually metered tenants are not. Second, there is a natural tendency to suppress utility allowances for individually metered tenants where not enough units participate at prices equal to or below the fair market rents.[206] This hits particularly hard in certain urban areas, such as New York City, where utility expenses per unit of energy not covered by the utility allowance are particularly expensive.[207]

The abuse of the utility allowance is made easier for the PHAs by the almost total lack of data available to evaluate the utility allowances given to Section 8 tenants. Most PHAs establish the allowance level at a designated community average for similar types of units.[208] Other PHAs differentiate allowances by basic building type or fuel types or by using utility company, prototype, or public housing program data.[209] Limited older data suggest that seventy-three percent of PHAs obtain generic consumption data from local utility companies or public housing utility records.[210] Only a minority of PHAs, between ten and forty percent, use actual data from utility consumption records within their jurisdictions.[211] In part, this lack of empirical data can be attributed to the HUD area offices, which use no consistent meth-

---

[202] URBAN SYSTEMS, *supra* note 17, at 5-13.
[203] *Id.*
[204] *Id.* at 5-18.
[205] GAO, VOL. 1, *supra* note 3, at 3.
[206] *Cf.* U.S. GENERAL ACCOUNTING OFFICE, RENTAL HOUSING: HOUSING VOUCHERS COST MORE THAN CERTIFICATES BUT OFFER ADDED BENEFITS 37–41 (1989).
[207] *Id.*
[208] *See supra* notes 71–75 and accompanying text.
[209] *See supra* notes 71–75 and accompanying text.
[210] URBAN SYSTEMS, *supra* note 17, at 5-9 and 5-10.
[211] *Id.* at 5-10; *see supra* note 75 and accompanying text.

odology to review or approve PHA utility allowances.[212] In its 1991 study, GAO reported that HUD does not even document these allowances.[213]

As a result, there are significant disparities between the allowances afforded individually metered Section 8 and public housing tenants. Section 8 tenants receive an allowance that the limited evidence available suggests is often set to maximize rent subsidies to landlords, rather than to provide sufficient funds for reasonable tenant energy consumption.[214] In contrast, public housing tenants receive an independently established, more realistic allowance, which varies greatly from PHA to PHA and by metering type.[215]

C. *Physical Conformance of Metering Types*

One way to eliminate differences in utility allowance subsidies based on metering types would be to convert all subsidized housing to a common form of metering. HUD has encouraged conversions to individual metering where technically and economically feasible.[216] But with the public and subsidized housing stock split into three different metering types, there is no dominant metering type now evident,[217] nor is there likely to be in the immediate future. For example, in public housing natural gas is master-metered in less than one-half of all projects (even though master metering is the most common metering configuration) while about forty percent of electricity in public housing is master-metered.[218] Moreover, in many jurisdictions, the easy conversions to individual metering have already been made.[219]

---

[212] URBAN SYSTEMS, *supra* note 17, at 5-12; GAO, VOL. 1, *supra* note 3, at 57–60.
[213] GAO, VOL. 1, *supra* note 3, at 59.
[214] URBAN SYSTEMS, *supra* note 17, at 5-4.
[215] 24 C.F.R. § 965.470–.480.
[216] 24 C.F.R. § 965.400.
[217] For a display of the types of metering in public housing, see GAO, VOL. 1, *supra* note 3, at 19 (Tables 2.1 and 2.2).
[218] *Id.*
[219] During the 1970s, individually metered electric heating established a foothold as the preferred heating system for developers of new residential housing. Since the mid-1970s, more than half the newly constructed units have incorporated electric heat. Most electric heating is electric resistance; more efficient heat pumps represent less than 20% of the electric heating in new construction. U.S. DEP'T OF ENERGY, THE FUTURE OF ELECTRIC POWER IN AMERICA: ECONOMIC SUPPLY FOR ECONOMIC GROWTH 3-69 (1983).

Compounding these practical obstacles to conversion, there is no consensus as to which type of metering is best. There are no conclusive studies on the actual consumption savings to be gained from conversions to different types of metering.[220] Indeed, HUD's assumptions about large savings from conversion to individual metering were rejected by the federal courts.[221] Moreover, the advantages of one type of metering over another will vary with the per-unit utility rates and tariffs in different service areas, which are subject to change over time.[222] Tenants and landlords often have differing opinions about metering configurations, and changing rate structures and technologies make the advantages of any particular configuration uncertain.

But even if certainty as to ideal metering types were possible, the capital costs of physically converting the metering and utility distribution systems of the entire public and subsidized housing stock could be prohibitive.[223] It is not only the meters themselves that are expensive, but also necessary rewiring, repiping, venting, cosmetic work, and the cost of complying with applicable codes.[224]

Given that meters themselves save no energy, but merely measure its use, an expensive program of meter conversion in an era of budget constraints and long waiting lists for admissions to

---

[220] For a variety of estimates of such savings, see, e.g., Steven Ferrey, *Fostering Equity in Urban Conservation: Utility Metering and Utility Financing*, *in* 2 OFFICE OF TECHNOLOGY ASSESSMENT, ENERGY EFFICIENCY OF BUILDINGS IN CITIES: WORKING PAPERS (1981) [hereinafter Ferrey, *Fostering Equity*]; BOOZ, ALLEN, & HAMILTON, ALTERNATIVE METERING PRACTICES: IMPLICATIONS FOR CONSERVATION IN MULTI-FAMILY RESIDENCES, 8-25 (1979) (prepared for the U.S. Dep't of Energy under contract number EC-77-C-03-1693); LOU MCCLELLAND, TENANT-PAID ENERGY COSTS IN MULTIFAMILY RENTAL HOUSING: EFFECTS ON ENERGY USE, OWNER INVESTMENT, AND THE MARKET VALUE OF ENERGY (1983) (prepared for the U.S. Dep't of Energy as report number DOE/CS/20050-1).

This data indicates that for space and water heating, a conversion to individual metering can result in savings of 5%, and up to 20% when controlled for weather. When electricity is not used for heating, the savings average 15%.

[221] Massachusetts Union of Public Housing Tenants v. Landrieu, 656 F.2d 899 (D.C. Cir. 1981) (decided without opinion); *see also* Massachusetts Union of Public Housing Tenants v. Pierce, 577 F.Supp. 1499 (D.D.C. 1984) (discussing procedural history and disposition of *Landrieu*); Federal Energy Regulatory Comm. v. Mississippi, 456 U.S. 742 (1982).

[222] *See generally* Ferrey, *Fostering Equity*, *supra* note 220, at 24–25 (documenting that in major urban areas, the factors discussed in this paragraph could increase rates per unit of energy by 33–100% after conversion to individual metering).

[223] At an average cost of about $1,000 per unit, converting one million units would cost $1 billion.

[224] *See, e.g.*, Natural Gas Pipeline Safety Act of 1968, 49 U.S.C. app. § 1671 (1994), which requires cathodic protection of gas lines against corrosion, 49 C.F.R. § 190.463 (1994), as well as inspection of lines and written emergency plans.

subsidized housing should proceed only if shown cost-effective in comparison to alternatives. But even as a long-term strategy, this option is constrained by a variety of regulatory, engineering, and economic concerns.[225]

Given these constraints, HUD efforts to require individual metering may not minimize the overall energy costs to the tenants, PHAs, or HUD. The relatively high delivered cost of electric resistance heating on a BTU basis[226] could increase the cost to HUD and PHAs for that portion of tenant utility expenses that is subsidized by utility allowances, increase the cost to tenants for basic utility service, and alter load factors for electric utilities.

In many units, meter conversion would not be possible at any reasonable cost; in other units, such costs would exceed those of other conservation alternatives. Because an equitable energy system in public and subsidized housing cannot rely on physical

---

[225]For example, local building codes may drive choices of both fuel types and metering systems because they typically prevent the running of natural gas or oil pipes within a building above a few stories in height. *See* Ferrey, *Fostering Equity*, *supra* note 220, at 33. Moreover, local codes may prohibit the placing of individual furnaces above the second or third floor.

Compounding these barriers, heat distribution systems in older buildings often are not engineered to allow conversion to individual metering. Also, because pipe runs in these older buildings often run the entire height of the building, metering one of these pipe runs would not register the consumption for a single unit but rather for all radiators vertically stacked in one portion of the building. *Id.*

Although BTU or flow meters, which measure the warm air or water delivered past a certain monitor (and to a particular unit) from a central boiler, were once quite expensive, their cost has decreased in recent years. *See* BOOZ, ALLEN, & HAMILTON, ALTERNATIVE METERING PRACTICES, *supra* note 220, at 12. Benefit-cost studies by PHAs in 1976–77 indicate that meter conversion often costs $100 per unit, with an overall range between no cost and $4,160 per unit. *Id.* at 13–14, IV-9; Ferrey, *Fostering Equity*, *supra* note 220, at 10. Today, more accurate flow meters and BTU meters are now availible at prices as low as $75 installed and $300 installed respectively. *Id.* However, this cost does not include the costs of replumbing or reducting the heating system of a multi-unit building to accomodate individual meters.

After weighing these concerns, the practical choice for existing high-rise multifamily structures often reduces to either (1) a central hydronic or steam fossil fuel-fired boiler, operating through a master meter, which distributes its energy to upper stories as warm air, steam or hot water, or (2) an individual electric heating system for each apartment, individually metered; the heating source is either a heat pump or the prevalent electric resistance heater.

Although in the short-run electric resistance heating is the cheapest way to convert a building already served by electricity to individual metering, in the long-run it will cost two to three times more than the direct application of fossil fuels or the more-expensive-to-install electric heat pumps. Lou McClelland, *Tenant-Paid Energy Costs in Multifamily Rental Housing: Effects on Energy Use, Owner Investment, and the Market Value of Energy*, *in* 1984 AMERICAN COUNCIL FOR AN ENERGY EFFICIENT ECONOMY PROC. E-9 (1984) [hereinafter 1984 ACEEE PROC.].

[226]KATHLEEN M. GREELY ET AL., BASELINE ANALYSIS OF MEASURED ENERGY CONSUMPTION IN PUBLIC HOUSING 9 (1986).

conformance of metering types in existing housing in the near future, policymakers should seek to render utility allowances meter-neutral.

D. *Minimizing Metering Variables in Allowance Determinations*

Because metering configurations can affect tenant utility consumption, it is appropriate to adjust utility allowances to the extent that a causative connection between metering configuration and reasonable consumption levels exists.[227] However, to the extent that metering configuration does not affect reasonable utility use, it should not influence allowance levels. This principle is particularly important given the financial impacts that flow from meter configurations.[228]

To reflect energy efficiency and individual reality, utility allowances should be empirically based and reflect individual unit and family variables. In order to achieve rational utility allowances, PHAs should account for causative factors in order of their importance.[229] Although the current public housing allowance regulations permit PHAs a wide discretion to consider important causative variables,[230] few PHAs do so.[231] Those PHAs that do so generally do not incorporate correctly the required distinctions or do not distinguish such factors on a project-by-project basis.[232] Section 8 allowances typically are uniform for a PHA or region, taking little account of building-by-building or tenant-based distinctions.[233]

---

[227] *See generally* Ferrey, *Cold Power*, *supra* note 17, at 33.

[228] *See* part III, *supra* notes 133–143 and accompanying text.

[229] More specifically, to set reasonable utility allowances, PHAs should account for the following axioms: (1) meters are unique energy conservation devices because they do not save money themselves but rather lead to differing changes in tenants energy consumption behavior. GREELY ET AL., *supra* note 226, at 9; (2) allowance distinctions for energy conservative behavior modification are unique to a particular unit or group of units. *Id.*; (3) individual data on the actual impact of a metering configuration are reliable; national or aggregated data are suspect when applied to different tenants and units. *Id.*; (4) because the utility allowance is delivered to tenants in dollars, rather than in quantities of energy, empirical data should evaluate the ultimate *cost* of utility service to tenants under alternative metering configurations; and (5) metering configuration should be factored into utility allowance calculations in an order of priority that reflects influences on utility consumption.

[230] 24 C.F.R. § 965.473.

[231] GAO, VOL. 1, *supra* note 3, at 72.

[232] *Id.* at 72–73.

[233] *Id.* at 72.

A study of several representative housing authorities[234] reveals that there are a number of significant causative relationships that PHAs may use to determine reasonable utility allowances:

- Low-rise projects use more energy than high-rise projects on a per square foot basis.
- Oil-heated projects utilize more energy per square foot than gas-heated projects.
- Family projects use more energy per square foot than elderly projects.
- Centrally heated projects consume more energy per square foot than do individually heated projects.
- Centrally heated projects pay less per unit of energy consumed than do individually heated projects.
- Steam hydronic heating systems consume more energy than do water hydronic heating systems.
- Public housing consumes more energy per square foot than does private multifamily housing.
- Position of the unit in the building affects heating requirements per square foot of space.
- The efficiency of major residential appliances affects consumption.

Because each of these relationships is more certain than the relationship between metering and utility usage, HUD should mandate that these higher level variables should be factored into the allowance calculation *before* PHAs make adjustments based on metering configurations.

In fact, under a more rational utility allowance scheme, PHAs would be free to implement allowances that made distinctions on metering configurations only to the degree that the allowances first accounted for all more significant causative variables, only if the metering distinctions were based on data from or directly relevant to the consumption requirements of their particular tenants and units, and only if the empirical data controlled for the influence of multiple correlation of other factors in the evaluation of the independent variable.[235] In addition, HUD area offices should retain final approval over PHAs' allowance-setting methodology. If PHAs cannot meet these criteria, then

---

[234] *See, e.g.*, Appendix; GREELY ET AL., *supra* note 226, at 5, 6, 10, 14.

[235] For a discussion of multiple correlation analysis and coefficients of multiple determination, see HEINZ KOHLER, STATISTICS 599 (1988).

HUD should require that allowances be established at uniform levels that assure tenants of all metering types sufficient utility allowances.

In many ways, the revision of the Section 8 system is more direct than for public housing. This results in part from the dearth of submetering in Section 8. However, individualized tailoring of allowances in the Section 8 program is not a traditionally established practice.[236] And any changes in the Section 8 allowance-setting system must overcome the current incentives for PHAs to understate the utility allowance by setting FMRs more reflective of actual utility costs and the data-gathering problems inherent in the number and variety of Section 8 units by offering PHAs additional monitoring and guidance.

To minimize the influence of the happenstance of metering configurations on utility allowances, two principles are central. First, meters are not themselves efficiency devices; they merely send financial signals to alter human consumption behavior. As such, their impact is quite individual. Second, more important variables that affect reasonable utility requirements should be accounted for first in setting allowances. These involve the application of energy efficiency and energy conservation devices and measures in public housing and Section 8 housing. While both principles seem straightforward, current practice reflects neither.

V. Designing an Equitable and Efficient Utility Allowance

A. *Three Basic Methodologies*

PHAs could use any number of methodologies to determine utility allowances for public and Section 8 housing. In practice, however, three basic types prevail:

- Allowances based on historical practices or rules of thumb ("historical" methodologies);
- Allowances based on engineering models ("top-down" methodologies); and
- Allowances based on consumption data from PHAs, from specific projects within PHAs, or from usage data of general utility service territories ("bottom-up" methodologies).

---

[236] GAO, Vol. 1, *supra* note 3, at 72–73.

Each of these three basic methodologies has several variations, and each is employed currently by some PHAs. While historical methodologies are most easily administered by PHAs, the data gathering required to implement the bottom-up approach imposes the greatest administrative burden on PHAs.

1. Historical Methodologies

Historical, non-empirical methodologies may or may not yield a reasonable allowance. Because the historical methodology relies only on past allowances and rule-of-thumb adjustments, it yields a conclusory number that has no demonstrable relationship to tenants' true utility costs.[237] Moreover, historical methodologies leave no objective basis for determining whether any such method is reasonable or whether the utility allowance and rental payment satisfy the thirty percent of income obligation. Most importantly, these non-empirical allowance methodologies do not allow Congress or any HUD policymaker to determine whether the $1 billion per year in utility subsidies[238] is well allocated. However, some PHAs continue to rely on such methodologies because they are so easy to administer; they involve no calculation, no gathering of field data, and no verification. Unfortunately, their ability to meet basic legal requirements is limited.

2. Top-Down Methodologies

The top-down approach involves compiling engineering data on appliance saturation or end-usage and modeling an apartment unit with typical quantities of energy necessary to maintain a certain level of consumption. This consumption level represents an idealized level of appliance usage or thermal comfort. The top-down methodology does not use data from a specific unit or from actual PHA project utility bills. Thus, it is only as accurate as the model used to determine the consumption pattern.

---

[237] Because the derived number has no empirical basis, HUD and the PHAs employing the historical method have no means of ensuring that the combination of the utility allowance and the rental payment meet the 30% of income standard Congress has established as the maximum permissable rent burden.

[238] R. RITSCHARD ET AL., CUTTING ENERGY COSTS IN PUBLIC HOUSING: TECHNICAL ISSUES, INSTITUTIONAL BARRIERS, AND RESEARCH NEEDS 1 (1986) (prepared for the U.S. Dep't of Energy under contract number DE-AC03-765F00098).

Because it incorporates engineering data, the top-down approach, unlike the historical methodologies, may be objectively verified. (In practice, however, PHAs seldom do verify the applicability of top-down allowances.) If the PHA has the appropriate level of expertise, the top-down methodology can be easy to administer and calculate. The PHA can design a model package of household energy end-uses and compute the quantity of energy necessary to power these end-uses. In addition, this approach makes it possible to derive different models for different kinds of public housing customers.[239]

Despite these advantages, top-down methodologies present two significant potential problems. First, any given model is not necessarily representative of the tenants' actual energy consumption.[240] Thus, a top-down methodology may not reflect tenants' actual consumption patterns or variations in consumption based on factors other than the physical variables of the housing unit and its appliances.

Second, a top-down methodology may not be representative of all units within a given project where there is a wide range of building- or appliance-related differences in consumption. To serve Congressional policy goals fully, a top-down model would have to be tailored to a number of different types of dwellings and situations. Unless a model is subdivided in this way, it will yield a result that, although reasonable in theory, may or may not be reasonable in practice. Engineering data that represents a "typical" family or rental unit may not represent a particular public housing unit because of several distinct characteristics of public housing.[241] The more sensitive a model is to these variables, the more precise and reasonable will be the resulting utility allowance.

However, many PHAs using the top-down approach employ general-population usage data or engineering recommendations

---

[239] For example, the PHA could calculate a different allowance for a tenant who requires a special energy-consuming appliance by adding the estimated consumption of such an appliance to the base utility calculation for that unit. Once the top-down model is in place, it is both flexible and easily administered.

[240] The top-down prototype is only as accurate as the assumptions that underlie it. If actual units show a wide disparity from the assumed prototype, a single allowance derived from the prototype will not reflect tenants' actual utility consumption very well.

[241] Among the characteristics that would make a generalized top-down model inappropriate when applied to low-income or subsidized housing are the following:

- All public housing is rental;
- All tenants are low-income;
- The public housing stock was constructed with minimal insulation or attention to energy efficiency.

see Ferrey, *Cold Power*, *supra* note 17, at 40–43.

that are not designed for rental units, let alone for low-income public housing projects.[242] These sources may miscalculate vital factors in the actual utility situation confronting PHA tenant units. In addition, any model methodology designed to yield a single utility allowance value for a unit of a given size, while perhaps representative and accurate for some typical tenants, will not be appropriate for other tenants.[243]

3. Bottom-Up Methodologies

The bottom-up approach employs actual consumption data, often with various adjustments to the data set, to determine utility allowances. This method reflects the actual consumption experience of tenants. However, it does not control for the amount or type of energy-consuming appliances in a particular unit. HUD has determined that utility allowances should not provide for consumption for certain excessive appliances or uses.[244]

Because a bottom-up methodology uses data reflective of public housing experience, it is more responsive to changes in tenants' utility consumption or rates than either of the two models discussed above. The reasonableness of the resulting bottom-up allowances depends principally on whether the data base employed by the PHA is disaggregated, whether the PHA manipulates the data base, and whether the allowance derived from the data base is flexible.

Many PHAs utilize a bottom-up methodology from an aggregated data base.[245] That is, the PHA may use data from one or

---

More of the public housing population is elderly than the general population (Approximately one-quarter of the public housing stock is designed for elderly or handicapped tenants, while less than 20% of the U.S. population is over 65 or handicapped.);

Because many public housing tenants are elderly, unemployed, or have children of young age, public housing units are occupied for more hours per day than housing in the general stock;

Low-income persons do not possess the same number or vintage of appliances as the general population. *See, e.g.*, EDWARD VINE & CHAIM S. GOLD, LOW INCOME HOUSEHOLDS AND ENERGY USE IN CALIFORNIA (1985).

[242] *See* GAO, VOL. 1, *supra* note 3, at 73.

[243] Even if a prototype allowance is fair for a public housing project generally, it still will not be reasonable for all tenants where there is a range of uncontrollable tenant utility expenses. As the data described in the Appendix show, actual tenant consumption varies substantially, perhaps for reasons beyond tenants' control. Thus, a prototype-based allowance may overcompensate or undercompensate some tenants.

[244] 24 C.F.R. §§ 965.476(a), .477.

[245] A majority of public housing and 40% of Section 8 housing allowances are established using actual data. *See supra* notes 79–85 and accompanying text.

two projects to generalize allowance calculations for all PHA projects within its jurisdiction. Such a procedure will be reasonable only for projects whose data is not used in the calculation if their physical, energy-use, and tenant-occupancy characteristics are similar to those of the projects whose data is used. Although all such projects within a PHA may be similar, it is difficult to verify this fact without specific effort in the field. In fact, it would be just as easy to calculate individually disaggregated allowances for each project because the PHA must collect and analyze data for each project in either case.

When each project's allowances are determined from usage data within that project, the relative disaggregation of the data base renders the allowances more relevant to tenants' actual experiences and makes them more likely to predict tenants' actual costs. However, disaggregation increases the administrative burden of data collection and calculation.

PHAs manipulate the data used in a bottom-up calculation in various ways.[246] The extent to which the resulting utility allowances are affected depends on the values discarded or manipulated,[247] the number of data observations,[248] and the type of calculation performed on the database to derive the allowance values.[249]

Like top-down allowances, bottom-up allowances can be quite flexible. But the flexibility of the latter is a consequence of the disaggregation of the data base to reflect better tenants' experiences, rather than a result of manipulating a model. Again, one

---

[246] For example, some PHAs routinely throw out of the calculation a certain percentage of high and low consumption values. These procedures seem to be ad hoc, lacking any methodological basis. *See* GAO, VOL. 1, *supra* note 3, at 40–41. Such manipulations will influence the calculation of bottom-up allowance values from any database.

[247] If a database has a large variance or standard deviation, an "average" allowance will be significantly affected by discarding extreme values in the survey universe. This effect does not appear if the allowance value is established at a modal point of the distribution, as discussed in more detail in parts V.D–E, *infra* notes 264–285 and accompanying text.

[248] A large database will be less affected than a small database by the exclusion of any particular value. However, large databases do not necessarily produce more reasonable allowances for any particular unit or family, as discussed in more detail in part V.E, *infra* notes 274–285 and accompanying text.

[249] If a PHA sets utility allowances at mean consumption, discarding extreme consumption values from the calculation will have a more profound impact than if the calculation establishes allowances at the modal value because most public housing consumption is positively skewed—more tenants incur extremely high costs relative to the bulk of the distribution than incur extremely low costs. Similarly, the closer in value the discarded data point is to the average or the mode of the distribution, the less its discarding will influence the allowance calculation. *See* part V.E, *infra* notes 274–285 and accompanying text.

should remember that increased flexibility comes at a price of increased data gathering and administrative costs.

In summary, while none of the three basic methodologies is perfectly reasonable, the top-down and bottom-up methodologies at least have an empirical foundation. The bottom-up option also reflects the actual experience of tenants, unlike the other two.

B. *Horizontal or Vertical Equity?*

Which of the three methodologies discussed above best promotes tenant equity? The answer to this question depends on whether one adopts a horizontal or vertical definition of "equity." Is equity treating all tenants equally (horizontal equity)? In that case, an arbitrary but uniform tenant allowance accomplishes this goal. On the other hand, if equity incorporates the concept of treating differently situated tenants distinctly and appropriately to reflect that difference (vertical equity), then any of these allowance methodologies must be disaggregated and individualized. If vertical equity is the goal, the bottom-up allowance methodology, which employs actual PHA data, best measures differences in tenants' situations. However, even the bottom-up methodology may have shortcomings if the data base does not reflect the particular PHA project's utility consumption or individual situations.

Whether policymakers should focus on horizontal or vertical equity depends on whether the deviation in tenant utility consumption reflects factors within or beyond the control of tenants. If the deviation results from factors within tenant control, then a horizontally equitable utility allowance would treat tenants equally, concluding that tenants could conform their utility usage to the allowance rate if they so chose. However, if the factors are beyond the control of tenants, then vertical equity may be more appropriate and an equitable allowance would treat differently situated tenants differently.

The methodologies developed and recommended in the remainder of this Article proceed on the assumption that vertical equity is a more appropriate standard because many variations in tenants' utility usage are beyond tenants' control. There is a significant variance in the energy requirements of identical but differently situated units within the same pro-

ject.[250] Moreover, significant anecdotal evidence also suggests that while variations in consumption reflect factors both within and beyond tenant control, extraneous factors are a significant variable.[251] Unless PHAs can determine that all extraneous factors are accounted for in an allowance methodology, penalizing tenants for consumption that may be beyond their control would be unreasonable.

Thus, in the context of PHA tenant utility allowances, the concepts of "average" and "reasonable" are not identical. The determination of a reasonable allowance, as required by regulation, involves considering whether it affords a tenant enough subsidy to pay for all utility usage that is a function of the thermal integrity of his unit and the requirements of the major energy-consuming appliances furnished to him. In other words, to be reasonable, an allowance must cover uncontrollable energy expenses.

This policy choice to focus on vertical equity dictates only the foundation on which to base an allowance and a method of calculation. Another significant question involves the actual level at which the allowance should be set.

C. *The Limitation of the Mean Value*

Before 1985, HUD urged PHAs to employ a bottom-up methodology utilizing an actual database[252] and to calculate an "average" allowance from the database.[253] "Average" can denote any value that statistically represents the central tendency of the database.[254] But as a practical matter, for individually metered tenants before 1985, PHAs established the allowance value at the mean value for each bedroom-size unit in the data base. The mean value predominated because PHA personnel were familiar with the calculation of mean values, as opposed to other measures of "average" points in a distribution.[255] Moreover, HUD has long assumed that a mean utility allowance value is per se reasonable. This assumption has pervaded the allowance levels both

---

[250] *See supra* note 241 (discussing some of the reasons for this variation).

[251] *See infra* notes 293–317 and accompanying text.

[252] *See* 24 C.F.R. § 965.478 (1984).

[253] *Id.*

[254] KOHLER, *supra* note 235, at 92.

[255] The use of mean value as the "average" dates to the 1963 HUD public housing handbooks. *See* LOCAL PHA HANDBOOK, *supra* note 17, at 34.

for individually metered tenants—where "reasonable" usually translates to the mean value—and for checkmetered tenants—where the free quantity of utilities is calculated at or within a percentage step from the mean allowance value.[256]

Because a mean is merely a derived value, it is possible that a mean value will not correspond to any observed value in the database that the mean represents and from which it is drawn.[257] Moreover, a mean value obscures the very uncontrollable variations that a reasonable, vertically equitable utility allowance may need to address. A single value cannot reflect the reasonable utility requirements of differently situated tenants. Because some tenants may not be "average," an allowance level established at a mean database value may not be reasonable in the sense of treating differently situated tenants differently.[258]

From a statistical perspective, a mean-value allowance represents certain types of databases better than others. A mean calculation may be illustrative of a database if individual utility consumption values represent a normal statistical distribution, observed values have a small standard deviation (i.e., observations are grouped closely about the mean), and the distribution is not skewed abnormally toward either extreme.[259]

Historically, HUD assumed that these three conditions typified public housing nationwide.[260] Actual data,[261] however, reveal that the distribution of public housing utility consumption is very irregular.[262] Because the data do not meet the conditions under which means are most descriptive of a data set,[263] the mean may not be the best method of calculating or reflecting "average" utility allowances.

---

[256] *Id.*

[257] *See* KOHLER, *supra* note 235, at 96–97.

[258] For a more complete discussion of the assumptions underlying this definition of "reasonable," see part V.B, *supra* notes 250–251 and accompanying text.

[259] *See generally* KOHLER, *supra* note 235, at 92–105.

[260] For example, at a 1977 meeting with the author, Lawrence Simons, then HUD Assistant Secretary for Housing, and his staff indicated that HUD assumed that public housing utility consumption was normally and tightly distributed. For evidence rebutting this assumption, see Appendix.

[261] *See* Appendix.

[262] *See id.* A positive skew makes sense because values are bounded by zero on the low end but unbounded on the high end. *Cf.* KOHLER, *supra* note 235, at 116–17.

[263] In some of the 77 data sets, the highest value of consumption is 10 times greater than the modal value (the "hump" in the distribution curve). *See* Appendix.

D. *Selection of Central Tendencies*

Statisticians use three primary measures of central tendency: the mean, the median, and the mode. A mean is an arithmetic summary of the distribution of data that represents a value around which observations tend to cluster.[264] Because the arithmetic mean gives each datum equal meaning and value,[265] where data are abnormal or asymmetric, the mean may nor represent any relevant value.

The median may also be employed to represent a central value in a distribution. The median value, like the mean value, represents a midpoint of the distribution. However, the median is whatever value happens to be found in the middle in a ranked or ordered array of observations in the database.[266] So the median, unlike the mean, always corresponds to an actual observation from the data set. Moreover, it may be more representative than a mean when the data set contains some extreme values.[267] Because the data for public housing utility consumption indicate that there are extreme values, the median may be a more reliable measure of the central tendency of a utility distribution.[268]

A third measure of central tendency is the mode, the most frequently occurring value in a set of data.[269] Like the median, the modal value corresponds to an actual consumption observation. A PHA could use the mode to set a utility allowance at the most common consumption level, but the limitations to such an approach are significant. Many PHA consumption distributions have multiple modes, modes located in an extreme tail of the distribution, or modes that diverge greatly from the mean and median values.[270] Moreover, a modal value "can be quite atypi-

---

[264] For a complete discussion of all three measures of central tendency, see KOHLER, *supra* note 235, at 92–105.

[265] As Professor Kohler states, "Knowing how to calculate the mean is one thing; knowing what it tells us is another. The mean is best viewed as a point of balance in a data set, very much like the fulcrum in a seesaw. Like the fulcrum, the arithmetic mean similarly balances the number of observations on one side of the mean . . . ." *Id.* at 96.

[266] *Id.* at 100.

[267] *Id.*

[268] Like a mean, a median value is limited as to its true representative value where there is not a normal distribution, where there is a large range or variance, or where the data is skewed from the central tendency. *Cf. id.* All of these characteristics are typical in public housing utility consumption. *See* Appendix. Despite these limitations, the median is more representative than the mean.

[269] *See* KOHLER, *supra* note 235, at 101.

[270] *See* Appendix.

cal of the majority of observations."[271] Where there is not a smooth, normal, tightly grouped distribution, the modal value suffers the same deficiencies as the mean and the median.

The lessons to be derived from this discussion are twofold. First, in devising allowance formulas, PHAs must recognize that public housing utility consumption does not fit a textbook description of a standard normal distribution.[272] Where the shape of the utility consumption distribution suggests that non-homogeneous factors underlie the data observed, any single measure of central tendency is likely to be misleading when applied to a particular unit—and is likely to understate tenants' actual utility consumption and need for a utility allowance. Second, variations in consumption are likely the result of both individual factors, such as controllable consumption (about which the allowance system should be less concerned), and factors beyond individual control, such as building characteristics, appliances, and circumstances related to family and work patterns in the household (about which the allowance system should be concerned).[273] In such situations, the derivation of utility allowances from a measure of central tendency alone can lead to unreasonable and inequitable applications to particular units.

E. *Dispersion and Shape of the Utility Distribution*

Current calculations of utility allowances do not account for the dispersion and the shape of the utility consumption distributions, thus ignoring two of the three primary statistical tools available to make meaningful judgments about data sets.

1. Measures of Dispersion

Dispersion is a measure of the spread of data away from the central tendency of a distribution, measured either as distances between different observations or as average deviations of observations from the central tendency of the distribution.[274]

---

[271] KOHLER, *supra* note 235, at 101.

[272] *See* Appendix.

[273] *See* the discussion of equity, part V.B, *supra* notes 250–251 and accompanying text.

[274] For a more complete discussion of dispersion and the measures of dispersion discussed below, see KOHLER, *supra* note 235, at 106.

The range measures the difference between the largest and smallest data points in a set. The PHA data demonstrate a large range within projects.[275] The range's utility is limited because it does not take into account all observations and does not provide a standardized means of comparison among distributions. Thus, it is difficult to use ranges of utilty consumption to compare changes in utility consumption over time and between different units or buildings.

Other more standardized statistical techniques, including the mean absolute deviation, the variance, and the standard deviation, also measure dispersion. The first two measures would be valuable in the construction of a reasonable allowance system but may be more difficult for PHAs to calculate than simple means. The mean absolute deviation calculates the average distance that each data point lies away from the mean.[276] The variance of a distribution is calculated by averaging the squares of each individual deviation from the mean.[277] This process gives much greater weight to the larger deviations,[278] produces a very large number in comparison to the deviation itself, and could be administratively confusing to a PHA official because the units of a variance (such as Kwh^2) are different from the units of the original observations (such as Kwh).

The standard deviation expresses useful statistical information in a very usable, standardized format. This standardization is important for calculation, verification, and assessment at both practical and policymaking levels. And, unlike the variance, standard deviation is expressed in the same units of measurement as the data points in the distribution.[279]

Standard deviation measures could be usefully employed in several ways on public housing data. If a PHA encountered normal distributions, calculating standard deviation could identify those tenants whose utility usage is excessive or extremely conservative. In a non-normal distribution, the ratio of the standard deviation to the mean indicates relative dispersion (often called the coefficient of variation).[280] PHAs could compare coefficients

---

[275] *See supra* notes 173–175 and accompanying text. *See also* Appendix.

[276] KOHLER, *supra* note 235, at 109.

[277] *Id.* at 110.

[278] *Id.*

[279] *Id.* at 111 (explaining that the standard deviation is the positive square root of the variance).

[280] *See id.* (discussing the coefficient of variation).

of variation of different projects' utility distributions, as well as coefficients of the same project's distribution over time, to ensure that the utility allowance provided is reasonable for a particular project.

2. Measures of Shape

As the study presented in the Appendix shows, PHA utility consumption is not normally or symmetrically distributed. To the degree that the peculiar shape of a distribution is not random, but instead a function of a building, appliance, or family requirements, a PHA should account for this abnormality in calculating a reasonable allowance level.

There are two primary statistical tools available to measure skewed data: skewness (symmetry) and kurtosis (peakedness).[281] Skewness measures the degree of distortion from horizontal symmetry.[282] Kurtosis is measured by the coefficient of kurtosis.[283]

3. Dispersion, Shape, *and* Central Tendency: Building a Statistically Robust Toolkit for PHAs

If a PHA plans to craft an allowance scheme that accurately and reasonably reflects actual consumption characteristics, it should use statistical tools that reflect the dispersion and shape of its tenants' utility consumption. It should employ all three basic statistical tools—central tendency, dispersion, and shape—to describe most completely the utility consumption distributions it observes.

These three tools would perform several important functions for different types of tenant distributions. First, they could signal when and where a PHA needs to investigate and distinguish between controllable and uncontrollable tenant utility consumption. Second, given different types of tenant utility consumption distributions, these three statistical tools could define the boundaries between reasonable and unreasonable utility consumption. Third, these tools could help determine reasonable allowance levels with more precision.

---

[281] *Id.* at 116.

[282] *See also* Appendix.

[283] KOHLER, *supra* note 235, at 119. The greater the kurtosis, the more peaked the distribution.

It is not so much that a particular calculation of central tendency alone holds the key to a reasonable allowance. Rather, the key is a more flexible, statistically robust, and disaggregated calculation of an appropriate level of subsidized consumption. An allowance *scheme*, rather than a simple average allowance, will provide a reasonable allowance to each tenant when the consumption distribution is dispersed, abnormal, and asymmetric.

Standardization and the use of allowance-setting methods capable of independent verification would improve oversight of the utility allowance program. A more reasonable allowance scheme, utilizing several statistical tools simultaneously to reflect the reality of PHA utility usage, would be more complex to administer. But left alone, the existing unidimensional "average" utility allowance calculations may not reflect a reasonable allowance level.

As part of any allowance-setting scheme, PHAs should use measures of central tendency, dispersion, and shape to help distinguish among building-, heating system-, and appliance-related consumption factors. PHAs would have to evaluate these factors as they relate to actual buildings and appliances by using unit energy ratings, available data on appliance procurement, thermal modeling, and related techniques.

Most PHAs would need assistance to undertake such detailed analysis. Ideally, HUD would provide PHAs consultants to train their personnel in two key components of a more sophisticated scheme. First, technical assistance would help PHAs identify the variations in utility consumption related to buildings, heating systems, and appliances through the top-down approach discussed earlier.[284] Although such an undertaking would be laborious, it would only need to be performed once for each building and then adjusted in response to major modifications. PHAs could make the original calculations from the architectural and engineering blueprints of the project buildings and data on appliance procurement and consumption. The process could be standardized and computer-modeled.

Second, technical assistance would help PHAs develop allowance formulas that better reflect the variation in tenants' utility usage. These formulas would account for climate, degree days, size of dwellings, number of occupants, and conservation incentives. Here again, PHAs could use standardized approaches or

[284] *See* part V.A.2, *supra* notes 238–243 and accompanying text.

computer-aided models. This exercise would separate controllable from uncontrollable tenant consumption in determining allowances.

It is beyond the scope of this Article to actually derive a utility allowance for a particular PHA project. However, the suggested first step would be to ask PHAs to begin the top-down analysis of housing units and appliances, segregating unit energy requirements and developing a standard deviation from the mean value for specific units. These can be expressed as expected variations and translated into a schedule of additions or subtractions for particular factors from the mean allowance for particular units.

In addition, a reasonable allowance should be unit-specific. A reasonable allowance would disaggregate for these characteristics of the unit and set several utility allowances for groups of units with similar thermal and electric energy use characteristics.[285] Accounting for the expected multi-modal distribution of utility usage, use of the median rather than the mean as the representative measure of central tendency, and sensitivity to the causes of standard deviation are critical elements of a reasonable, objective, and vertically equitable allowance.

VI. Energy Efficiency and Utility Allowance Design

A. *Stakeholder Interests*

Because Congress and HUD make available only a finite amount of utility allowance funds, they create a zero-sum contest between the PHA and tenants for those funds. However, both the PHA and tenants can agree that an allowance program that promotes greater energy efficiency is desirable; greater energy efficiency reduces the PHAs' program costs, allowing them either to serve more eligible persons or improve the quality of housing available to current tenants.

The congressional directive initiating the GAO study of public housing utility allowances expressed an interest in providing tenants a strong incentive to conserve energy and reduce utility costs.[286] By promoting energy conservation, HUD can minimize

---

[285] For a more complete description of this "bottom-up" approach, see part V.A.3, *supra* notes 244–249 and accompanying text.

[286] 1987 Act, *supra* note 56, § 102.

the amount of subsidies necessary to sustain public housing energy consumption.

As the conduit through which federal Performance Funding System (PFS) dollars are passed to tenants as utility allowances,[287] the PHA expands its discretion over total program resources if it can contain energy costs. This consideration is especially important since the delinking of PFS and the utility allowance system in 1985: now the PHA may receive from HUD more or less than is dispersed to tenants in a given year.[288] Faced with these incentives, PHAs generally exhibit the following characteristics:

- Because they rely on HUD for funding, PHAs are extremely sensitive to incentives inherent in the HUD-PHA relationship.
- Because of the significant administrative burdens on PHAs, they are more reactive than innovative with regard to utility allowances.
- PHAs focus on cash flow; incentives for conservation are meaningful if they improve PHA cash flow.

B. *Means of Conserving Energy in Public Housing and Section 8 Housing*

It is important to distinguish between two basic types of energy conservation. Individuals can conserve energy through individual care in the use of utility service.[289] Technical fixes to buildings and heating systems and appliance procurement can also improve the inherent technical efficiency of energy-consuming end-uses.[290]

---

[287] *See* 24 C.F.R. §§ 990.101–.116 (establishing the PFS).

[288] *See* 24 C.F.R. § 965.476; 24 C.F.R. §§ 990.101–.116.

[289] This type of energy conservation can be motivated by financial incentives, assuming that the tenant has a sufficiently elastic demand for utility service, understands the nature and operation of incentives, has control over energy-consuming end-uses, and is able to conform his or her behavior in line with these incentives.

[290] This type of conservation includes weatherization, procurement of efficient energy-consuming appliances, and balancing of heating systems. In rental housing, management may have to make these investments rather than tenants.

1. Incentives for Tenants to Conserve Energy

The 1985 deregulation of public housing utility allowances by HUD allows PHAs to provide incentives for tenant energy conservation.[291] There is, of course, a limit to the number of behavioral changes or lifestyle deprivations that any consumer of utility service can or should endure.[292] In addition, savings accomplished at the expense of reasonable tenant comfort or necessity would conflict with PHAs' obligation to provide reasonable utility allowances.

2. Energy Efficiency Investments

Energy efficiency investments save energy by their very undertaking—the savings are inherent in the technology. They do not depend on tenant behavioral responses, and they are permanent for the life of the efficiency investment. Energy efficiency investments are positively correlated with education and income.[293] Most public housing was constructed prior to 1974, when energy awareness increased in response to the oil boycott.[294] More than two-thirds of all public housing is located in federal regions II-V.[295] These four of the ten federal regions also experience some of the coldest winter climates. Almost two-thirds of all public housing units are in climate zones with more than 4000 heating degree days.[296]

---

[291] 24 C.F.R. § 965.476; *see also* 47 Fed. Reg. 35,251 (1982) (notice of proposed rulemaking). These incentives, where implemented, have assumed the form of lower tenant utility allowances, rather than specific conservation incentives. Lower allowances shift financial responsibility for utility consumption from management to tenants, who may or may not be able to reduce utility consumption in response to lower allowances.

A PHA takes little financial risk by promoting this type of energy conservation. If conservation is successful, the PHA can even profit by retaining a portion of any conservation savings. 24 C.F.R. § 990.107(g).

[292] For a discussion of the human dimension in energy consumption, see PAUL STERN & ELLIOTT ARONSON, ENERGY USE: THE HUMAN DIMENSION 1–31.

[293] U.S. DEP'T OF ENERGY, RESIDENTIAL ENERGY CONSUMPTION SURVEY, *supra* note 35, at 23.

[294] *See* RITSCHARD ET AL., *supra* note 238, at 1–2; *see also* notes 140–42 and accompanying text.

[295] 1 PERKINS & WILL AND EHRENKRANTS GROUP, AN EVALUATION OF THE PHYSICAL CONDITION OF PUBLIC HOUSING STOCK 11 (1980) [hereinafter PHYSICAL CONDITION OF PUBLIC HOUSING] (prepared for U.S. Department of Housing and Urban Development).

[296] *Id.* at 107.

Given these factors, it is vital to provide tenants well-insulated housing. Yet public housing is some of the least well-insulated housing nationwide.[297] At best, the thermal integrity of public housing is irregular. Also, there is an extremely wide variation of energy use in public housing nationwide.[298] The facilities at the highest end of the scale use six times more energy than those at the low end of the scale.[299] Even controlling for differences in weather and heating degree days, there still exists a variation by a factor of three in the BTU energy use among different facilities.[300]

This lack of thermal integrity is remediable. A report commissioned by HUD on the efficiency potential of public housing estimates that a complete cost-justified energy conservation program would reduce annual energy operating costs by thirty to sixty percent (with an average forty-eight percent savings) depending on project characteristics.[301]

These savings could be achieved for an average investment of $800–$2,500 per unit in 1980 dollars,[302] paying back the initial investment in six years and yielding a return on investment of more than fifteen percent annually. The cost of implementing these measures in all public housing would be $2 billion and yield a savings of 1.5 quads of energy.[303]

The Office of Technology Assessment (OTA) estimates that low-income single family housing has the technical potential to conserve about half of its energy use.[304] A body of data in low-income housing shows impressive cost-effective efficiency savings from utility-sponsored programs,[305] from federal government-sponsored conservation demonstration programs,[306] and from the low-income Weatherization Assistance Program.[307] More-

---

[297] *Id.*

[298] SHERWOOD ET AL., *supra* note 63, at 3 (Table 1).

[299] *Id.*

[300] *Id.* at 6 (Table 2).

[301] PERKINS & WILL AND EHRENKRANTS GROUP, ENERGY CONSERVATION FOR HOUSING: A WORKBOOK 1-1, 1-6 (May 1982) (prepared for the Department of Housing and Urban Development under Research Contract H-2850 as a component of PHYSICAL CONDITION OF PUBLIC HOUSING, *supra* note 295).

[302] *Id.* at 1-6.

[303] *Id.* at 1-8.

[304] 1982 OTA REPORT, *supra* note 1, at 155 (Table 53).

[305] *See*, CHARLES A. GOLDMAN, TECHNICAL PERFORMANCE AND COST-EFFECTIVENESS OF CONSERVATION RETROFITS IN EXISTING U.S. RESIDENTIAL BUILDINGS: ANALYSIS OF THE BECA-B DATA BASE, 32 (Oct. 1983).

[306] Steven Ferrey, *Pulling a Rabbit Out of the Hat: Innovative Financing for Low-Income Conservation*, *in* 1984 ACEEE PROC., *supra* note 225, at H-38.

[307] *Id.* at H-39; GOLDMAN, *supra* note 305, at 45, 51.

over, there are many proven examples of successful, cost-justified energy efficiency investments in building shells that show impressive savings with quick paybacks.[308]

In addition to installing efficiency investments in the building shell, there are impressive savings available from improvements to the efficiency of the heating system. For example, in Trenton, New Jersey, elderly public housing units experienced an almost fifty percent reduction in heating energy used solely from the installation of a computerized heating load control system.[309] This savings reflects an investment of about $250 per unit, with a payback of less than one year.[310] In short, the data illustrate

---

[308] Consider the following examples:

In low-income private market multifamily buildings in Roxbury, Massachusetts, conservation retrofits in the first year demonstrated savings of 26–55%. Stephen Morgan, *Sharing Savings in Multifamily Housing: The Incentive Dividend in* 1982 AMERICAN COUNCIL FOR AN ENERGY EFFICIENT ECONOMY PROC. 11 (1982) [hereinafter 1982 ACEEE PROC.]. These savings represent simple paybacks of less than five years on the initial capital cost of the conserving investment; the return on investments is 21–38% per year. *Id.*

Very rudimentary conservation measures installed in San Francisco PHA buildings, as part of Pacific Gas & Electric Company's Zero Interest Loan Program (ZIP), resulted in savings of 7–20%. Charles A. Goldman et al., *Energy Conservation in Public Housing: The San Francisco Experience, in* 1984 ACEEE PROC., *supra* note 225, at H-48–49. These savings showed 4–6 year paybacks of the initial investment in energy savings, conserving energy at the cost of $2.50/MBTU (million BTU). *Id.* The average cost of the conservation materials installed was $150 per dwelling unit. *Id.*

Over 2400 units of multifamily projects managed by Citizens Conservation Corporation achieved savings averaging 35% of previously consumed energy. Charles R. Haun, *Seven Keys to Energy Conservation in Multi-Family Dwellings: Citizens Conservation Corporation's Approach to Energy Conservation in Multi-Family Buildings Housing Low-Income and Elderly Residents, in* 1984 ACEEE PROC., *supra* note 225, at C-54.

[309] Chaim S. Gold, *The Page Homes Demonstration Energy Conservation Computer System, in* 1982 ACEEE PROC., *supra* note 308, at 11.

[310] *Id.* In addition, the Department of Energy estimates that for oil-heated residences, simple improvements to the furnace can achieve impressive heating energy savings. *See generally* U.S. DEP'T OF ENERGY, ENERGY SAVING OPTIONS FOR HOME HEATING EQUIPMENT (1978). In field tests, 20% savings on an initial investment of $500 have been demonstrated from installation of an oil furnace tune-up package in low-income housing. U.S. DEP'T OF ENERGY, INCREASED BENEFITS OF ENERGY ASSISTANCE PROGRAMS THROUGH OIL FURNACE RETROFITS B-5 (1981). Investments in low- and moderate-income building heating system improvement and basic weatherization in Chicago achieved energy savings of 26–42%. Michael Freedberg & Dan Schumm, *New Initiatives in Financing Multifamily Energy Conservation: Recent Developments in Chicago, in* 1986 AMERICAN COUNCIL FOR AN ENERGY EFFICIENT ECONOMY PROC. 4.74 (1986) [hereinafter 1986 ACEEE PROC.]. Use of outdoor reset and cutout controls for hydronically heated buildings in Minnesota demonstrated savings averaging 18%, with a payback of less than one year and an installed cost averaging $450 per unit. Martha Hewitt & George Peterson, *Measured Energy Savings from Outdoor Resets in Modern, Hydronically Heated Apartment Builings, in* 1984 ACEEE PROC., *supra* note 225, at C-135. Balancing single-pipe steam systems in Minnesota multifamily buildings yielded savings of 15–25%, representing about a two-year payback. George Peterson, *Correcting Uneven Heating in Single Pipe Steam Buildings: The Minneapo-*

significant savings in heating system and building shell efficiency available from a variety of measures.

3. More Efficient Household Appliances

Public housing authorities furnish all units with major household appliances.[311] Miscellaneous household appliances and domestic hot water heaters are responsible for about thirty-five percent of primary residential energy consumption but account for almost half of energy expenses.[312] The remainder of use and expense is for space heating.

For nearly all household appliances, there are now alternative models at least fifty percent more efficient than their older counterparts.[313] For example, efficiencies of refrigerators currently in use vary by a factor of more than three.[314] Freezers and air conditioners currently in use demonstrate similar variations in efficiencies.[315] Equally impressive savings are available for lighting, water heaters, furnaces, and stoves.[316]

Investing in efficient household appliances demonstrates an annual rate of return on the investment of between nine percent and fifty-two percent,[317] repaying in energy savings the additional costs of the more efficient appliances in two to eight-and-half years depending on the appliance—or in less than half the period of the appliance's lifetime.[318] Since appliances, unlike

---

lis Steam Control System, in 1984 ACEEE PROC., *supra* note 225, at C-153. Replacement of steam heating systems with hot water systems and new boilers in Minnesota saved 17–39% of energy. Mary Sue Lobenstein et al., *Converting Steam Heated Buildings to Hot Water: Practices, Savings, and Other Benefits, in* 1986 ACEEE PROC., *supra* at 1.14.

[311] *See* 24 C.F.R. §§ 965.601–.605 (requiring PHAs to furnish appliances).

[312] *See* Howard S. Geller, *Efficient Residential Appliances and Space Conditioning Equipment: Current Savings Potential, Cost Effectiveness and Research Needs, in* 1984 ACEEE PROC., *supra* note 225, at E-119. Primary energy accounts for the total raw units of energy that are used to produce the energy service, including losses in the combustion of fossil fuels. As the building shell is made more efficient, miscellaneous household appliances account for a greater share of total household energy consumption. *Id.*

[313] *Id.* at E-118.

[314] David B. Goldstein, *Efficient Refrigerators: Market Availability and Potential Savings, in* 1982 ACEEE PROC., *supra* note 308, at 2; Michael Jaske, *Trends in Residential Appliance Efficiency Choice, in* 1984 ACEEE PROC., *supra* note 225, at E-166.

[315] Goldstein, *supra* note 314, at 11.

[316] Geller, *supra* note 312, at E-128. Between 19% and 29% of natural gas use in public housing is for cooking. Goldman et al., *supra* note 308, at H-53.

[317] Geller, *supra* note 312, at E-123.

[318] *Id.*

buildings, are constantly replaced with wear, a successful appliance efficiency strategy can be implemented incrementally—replacing a fraction of the existing appliances each year. However, because appliances seldom are replaced before their lifetimes are complete, if a PHA misses an opportunity to purchase the most efficient replacement appliances in any year, it may wait as long as twenty years until that appliance is again replaced. Unrealized appliance efficiency savings will be lost during this entire period.

4. Educating Tenants to Conserve Energy

Finally, tenant behavior in using appliances and heating systems is also a factor.[319] For example, some public housing tenants use their gas stoves regularly for supplemental winter heating[320] or use more electricity to run their televisions constantly than to operate their refrigerators.[321] These behavioral factors can cause significant variations in PHA energy use.[322] But they also can be addressed successfully by tenant education efforts. Under the post-1985 utility allowance deregulation,[323] PHAs can use their discretion to structure successful counseling and education programs.

C. *Incentives*

There is a fundamental barrier to energy efficiency in all rental housing: tenants have no incentive to invest in landlords' housing, and landlords would prefer individual metering to investing in energy efficiency.[324] Although efficiency investments offer a win-win situation to all stakeholders, the PFS mechanism masks PHAs' incentives to invest in energy efficiency.[325] In the first

---

[319] *See generally* STERN & ARONSON, *supra* note 292, at 73–88.

[320] Richard Diamond, *Energy Use Among the Low-Income Elderly: A Closer Look*, *in* 1984 ACEEE PROC., *supra* note 225, at F-63.

[321] *Id.* at F-59.

[322] *Id.* at F-52.

[323] 24 C.F.R. § 965.476.

[324] DEBORAH BLEVISS & ALISSA GRAVITZ, ENERGY CONSERVATION AND EXISTING RENTAL HOUSING 13–14 (1984). The average tenure of tenants in rental housing is less than one year for 42% of tenants and less than five years for 75% of tenants. *Id.* at 9.

[325] For a more complete discussion of the PFS system, see Ferrey, *Cold Power*, *supra* note 17, at 70–74.

year of conservation savings, the PHA and HUD split the savings.[326] Thereafter, over a three-year period, the rolling historic base for calculation of PFS utility subsidies progressively reflects the conservation savings realized.[327] By the end of the fourth year after the investment, the PHA share of the savings diminishes to zero, giving HUD the benefit of any savings.[328] Thus, over the long term, the PHA perceives little incentive to undertake energy efficiency investments. Compounding this disincentive, HUD no longer encourages these investments as a primary funding priority despite their savings for the system as a whole.[329]

Public housing tenants perceive little incentive to invest in efficiency investments in their landlords' property. Therefore, while tenants will benefit from energy efficiency investments, they are neither motivated to finance the investment nor, often, able to do so. Several studies reveal factors of low-income household energy consumption that bear on the ability of tenants to make meaningful lifestyle or behavioral changes in response to financial incentives delivered in the form of lower utility allowances. Overall, the energy intensity of residential space is significantly higher for elderly and poor households.[330] In other words, for the same amount of space, low-income tenants utilize more energy than the general population. Therefore, use of a top-down engineering prototype methodology[331] or utility usage data from the general population may tend to underestimate the actual usage of a low-income household.

Elderly and low-income households are less energy efficient and employ fewer energy conservation technologies.[332] These households consume a disproportionate amount of their household energy budget for heating and cooling,[333] likely because low-income households have fewer appliances and because low-income households and the elderly keep their dwellings sig-

---

[326] 42 U.S.C. § 1437g(a)(2) (1988); 24 C.F.R. § 990.107(f).

[327] 24 C.F.R. § 990.107(c).

[328] *Id.* Although HUD has allowed PHAs making certain energy efficiency investments to retain the savings for up to 12 years, as necessary, to repay their costs, few PHAs have used this authority because, under it, the PHAs do no better than break even while experiencing lower PFS operating subsidies.

[329] The priorities for allocation of CIAP funds were revised to remove energy efficiency investments from the list of priority funding items. 24 C.F.R. § 968.201.

[330] Marilyn A. Brown & Paul A. Rollinson, *The Residential Energy Consumption of Low-Income and Elderly Households: A Summary of Findings from Decatur, Illinois, in* 1984 ACEEE PROC., *supra* note 225, at H-5.

[331] *See* discussion parts V.A–B, *supra* notes 240–251 and accompanying text.

[332] Brown & Rollinson, *supra* note 330, at H-7.

[333] *Id.* at H-11.

nificantly warmer than the general population.[334] This could reflect more time spent in the home, or it could represent less attention to conservation possibilities. Another potential explanation for this difference is that elderly households consume less than average amounts of energy for cooking.[335]

Energy consumption is also a function of the stage of one's life. Consumption increases with maturity through the childrearing years; the elderly consume less than single young adults.[336] This life-cycle variation in energy use suggests that the elderly may be fairly energy conservative as a group and that the type of family unit in an apartment, as well as the age of children, must be regarded as a variable in the calculation of a reasonable utility allowance.

PHAs may or may not be able to cut excessive utility consumption by cutting utility allowances. But simply lowering allowance formulas (typically established at the arithmetic mean or arbitrarily at historic levels) across the board as a regulatory inducement risks causing hardship to some tenants in units that do not have the capability to reduce utility consumption without depriving themselves of basic energy needs. Without making such an effort to distinguish building-, unit-, and family-related characteristics, lower across-the-board utility allowances, however implemented, miss the goal. Moreover, because efficiency investments reduce uncontrolled variations in tenant utility usage (the variations that lead to vertical inequities in current allowance-setting schemes), they render *any* system of utility allowances more reasonable. By trimming extreme energy consumption and minimizing building- and appliance-related variations in tenant consumption, efficiency investments lead to a situation in which remaining variations in energy consumption are more likely within the control of the tenant, making utility allowance incentives a more appropriate motivator of conservation.

Thus, energy efficiency investments are prerequisites to a more equitable allowance scheme. Because efficiency renders the consumption distribution more normal, any project-wide allowance

---

[334] *Id.* (reporting that on average low-income households keep their dwellings two degrees Fahrenheit warmer than the general population).

[335] *Id.* at H-9.

[336] *Id.* at H-6.

scheme will be more representative and reasonable after energy efficiency investments are made.

VII. Conclusion

Public housing energy subsidies cost taxpayers more than $1 billion annually, and Section 8 utility subsidies are of equal magnitude. They represent twenty-six to forty-two percent of PHA operating costs, constituting as much as eighty-five percent of HUD operating subsidies.[337] Public housing energy costs are sixty-five percent higher than in comparable private market housing.[338] Each of these two factors raises questions about the current utility allowance system.

The GAO report reveals that HUD does not monitor closely PHA utility allowance administration, as it is required to do. In fact, many tenants pay substantially more for shelter and utilities than thirty percent of their incomes, the maximum allowed under the Brooke Amendment. Most PHAs are not aware of their rent burden shortfalls, which result, in part, from the disparate treatment of public housing and Section 8 tenants and of tenants with differing meter configurations.

PHAs currently set allowances using simplistic historical or "average consumption" models, implicitly assuming that utility consumption in public housing is normally distributed. However, the observed data for public housing and Section 8 housing exhibits more complicated patterns and demands more sophisticated statistical methods. And the current, simplistic models frequently produce utility allowances that give tenants with differing utility needs identical allowances, thus effecting a hidden transfer among tenants.

This Article has developed five basic criteria to evaluate both current allowance methodologies and any proposed changes to those methodologies:

- Accuracy in setting allowances that keep tenants' shelter and utility expenses below the thirty percent statutory rent ceiling.
- Incentives for efficient energy use.
- Administrative ease.

---

[337] Mills, *supra* note 63, at 170.
[338] *Id.*

- Potential for cost savings in implementation.
- Vertical equity, in the sense of treating differently situated tenants differently and appropriately.

A. *Recommendations for Public Housing*

Several measures would help HUD and PHAs set more equitable allowances for public housing. First, they should neutralize allowance disparities based on metering configuration. Since physical conformance of all metering configurations is impractical, HUD should encourage the development of meter-neutral allowance formulas to replace the current, unresponsive models used by PHAs.

Second, using a more robust set of statistical tools to set utility allowances in public housing would help PHAs meet statutory criteria. Better statistical tools would distinguish between controllable and building-related or uncontrollable consumption, allowing PHAs to allocate utility allowances in consideration of these differences. To establish equitable utility allowances, PHAs must conduct a disaggregated evaluation of building- and appliance-related factors before establishing allowances. Thereafter, PHAs can employ several statistical tools to standardize the provision of reasonable allowances.

A unidimensional or aggregated allowance will not provide a tool with which a PHA can distinguish tenant-controllable from building- and appliance-related consumption. Although statistically more complex bottom-up approaches that distinguish building- and appliance-related consumption factors involve more cost to implement, they will set more equitable allowances and may result in savings by allowing the identification of low-cost units where traditional allowances have overcompensated tenants. Moreover, while administrative burden and cost are significant in the first year of implementation of a more complex allowance, burden and cost decrease substantially thereafter because building-related variables need only be determined once. To facilitate the transition to more complex allowance-setting methods, HUD should provide PHAs technical assistance to develop standardized approaches.

Finally, incentives for energy efficiency are important not only because of congressional emphasis, but because their cost savings for all stakeholders make them the only measure with the potential to facilitate more "reasonable" allowances from the

perspective of *all* stakeholders. Efficiency improvements are possible and very cost-effective in public housing. And because efficiency improvements tend to reduce the variation, skew, and dispersion in utility consumption patterns, such improvements would make *any* allowance scheme more accurate.[339]

B. *Recommendations for Section 8 Housing*

The Section 8 utility allowance program has three salient differences from the public housing program. First, unlike public housing, Section 8 has a fair market rent ceiling. Second, explicit regulatory protections for tenants are more extensive for public housing tenants. Third, there is more random variation in the size and characteristics of Section 8 housing than of public housing.[340]

These three differences indicate that PHAs should use even more care in determining Section 8 utility allowances than they do in determining public housing utility allowances. Not only are there more Section 8 tenants, but there is less regulatory guidance, greater PHA discretion and disparity in allowance setting, and greater variation in the energy-consuming characteristics of the housing stock. Moreover, PHAs must resist any incentive to suppress Section 8 utility allowances in order to raise the portion of fair market rent allocable to shelter, thus inducing more landlords to participate in the program. This transfers a greater percentage of the fair market rent to owners at the expense of tenant utility allowances.

The deficiencies of traditional allowance methodologies for public housing are equally true for Section 8 housing. Because participation in the Section 8 program may be periodic for a given building or unit, it becomes a more daunting task to develop the type of unit-specific audit of each unit that this Article recommends for the more stable public housing stock. The administrative burden of such an undertaking would likely be formidable.

---

[339] Because efficiency improvements tend to narrow the variation and skew of utility consumption distributions, more units will converge toward the utility allowance level. With extremes narrowed in this fashion, any allowance system will be more equitable.

[340] URBAN SYSTEMS, *supra* note 17, at 5-35. Bedroom size is not a reliable predictor of utility requirements in Section 8 housing because of variability of the quality and size of the private market housing stock, among other reasons. *Id.*

Instead, PHAs should develop a prototype top-down methodology for determining Section 8 utility allowances. This prototype calculation can be improved by standardizing calculation factors (with guidance from HUD) and by gauging the prototype's accuracy in comparison to actual consumption data in equivalent public housing or private market units.

The energy efficiency of both public housing and Section 8 housing is minimal by design, making both ideal candidates for efficiency investments. Because PHAs tend to react to HUD cash flow incentives rather than innovate on their own, HUD must initiate and promote a more reasonable allowance concept based on energy-efficient dwellings. Properly implemented, the dual incentives of energy efficiency and utility allowance reform could remedy the significant defects in both Section 8 and public housing utility allowances highlighted by the GAO for Congress.

APPENDIX

In 1977, the author and Katherine Bachman performed a statistical analysis of energy consumption in HUD-funded housing projects. *See* STEVEN E. FERREY & KATHERINE BACHMAN, STATISTICAL ANALYSIS OF PROPOSED HUD UTILITY ALLOWANCE REGULATION (1977). The authors gathered seventy-seven sets of summer and winter utility consumption data from PHAs in nine different cities across the country, representing all types of climates. The data were collected from different kinds of buildings with varying bedroom sizes, dates of construction, heating systems, and brands of PHA-supplied appliances. *Id.* at 6.

Fifty-one of the seventy-seven data sets for electricity had mean values in excess of their median values, indicating that the distribution of these sets is positively skewed. *Id.* at 13. In those units, HUD regulations led as many as forty-six percent of tenants being surcharged. *Id.* Data sets for natural gas units showed a similar skew. *Id.* at 14–15. Moreover, in some of the seventy-seven data sets, the highest value of consumption observed was ten times greater than the modal value (the "hump" in the distribution curve). *Id.*

The analysis concluded, "This consumption does *not* exhibit similar mean and median value, does *not* exhibit steep bell-shaped curves, and does *not* exhibit a non-disperse pattern." *Id.* at 16 (emphasis original). In short, the analysis shows that any prototype methodology designed to yield a single allowance value will be reasonable for a few tenants but that many other tenants will be either undercompensated or overcompensated by a unitary allowance.

In fact, many of the consumption curves plotting the data described above are like roller-coasters, indicating the large range and variance of the data. The underlying data and graphical representations of each of the data sets comprising this analysis are on file with the author.

NOTE

SHOULD WE BELIEVE THE PEOPLE WHO BELIEVE THE CHILDREN?: THE NEED FOR A NEW SEXUAL ABUSE TENDER YEARS HEARSAY EXCEPTION STATUTE

ROBERT G. MARKS*

As politicians and courts feel increasing pressure to get tough on crime, the criminal justice system remains ill-equipped to cope with the serious problem of child sexual abuse. Since the strongest evidence in these cases is often technically hearsay under federal and state rules of evidence, the hearsay exclusion rules make prosecution of these cases quite difficult. In this Note, the author analyzes the legal rules under which the states currently admit child hearsay in sexual abuse prosecutions. He concludes that the traditional hearsay exceptions, and even statutory exceptions designed specifically to deal with this subject, do not strike the proper balance between the rights of the accused and the goal of fighting child abuse. Therefore, the author proposes a Model Statute that he argues would allow factfinders to hear crucial pieces of evidence, while still keeping out the least reliable forms of hearsay.

The sexual abuse of children is one of America's most terrifying social problems. In 1990, a U.S. Department of Health and Human Services task force declared the situation a national emergency.[1] The American Humane Association estimates that in 1986 alone there were 132,000 substantiated cases of child sexual abuse.[2] General population surveys suggest that between twelve and

---

* Associate, Wolf, Block, Schorr and Solis-Cohen, Philadelphia, Pennsylvania. A.B., Princeton University, 1991; J.D., Harvard Law School, 1994. The author would like to thank Professor Alan Dershowitz, John Oakes, Jordan Cohen, and Benjamin Uressman for their advice and suggestions throughout this Note and beyond. Comments about this Note may be addressed via Internet to HJOL@HULAW1.HARVARD.EDU.

[1] *See* Henry Weinstein, *Child Sex Abuse Cases Pose Dilemma for Prosecutions*, L.A. TIMES, Sept. 19, 1993, at A1.

[2] *See* AMERICAN HUMANE ASS'N, HIGHLIGHTS OF OFFICIAL CHILD NEGLECT AND ABUSE REPORTING 1986 23 (1988). In 1983, the American Humane Association reported 71,961 cases of child sexual abuse—an 852% increase over the number it reported in 1976. *See* Jason A. Levine, Recent Development, *The Confrontation Clause and Hearsay Statements by Child Victims of Sexual Abuse: White v. Illinois, 112 S.Ct. 736 (1992)*, 15 HARV. J.L. & PUB. POL'Y 1040, 1040 (1992) (citing *Backlash Feared in Child Sex Abuse Case*, WASH. POST, Mar. 23, 1985, at A13).

It is unclear, however, whether the actual incidence of child sexual abuse is increasing or whether only the reporting of child sexual abuse is increasing. *See* John Patrick Grant, *Face—to Television Screen—to Face: Testimony by Closed-Circuit Television in Cases of Alleged Child Abuse and the Confrontation Right*, 76 KY. L.J. 273, 274–75 (1987).

thirty-eight percent of women and three to sixteen percent of men were subjected to some form of sexual abuse in their childhood.[3] Many child sexual abuse victims suffer serious short- and long-term psychological repercussions, such as extreme depression and anxiety, low self-esteem, and suicidal tendencies.[4] A moral and just society should take extraordinary measures to protect its children from the horror of child sexual abuse.[5]

A strong countervailing consideration is the frightening prospect that a person may be falsely accused of and prosecuted for sexually abusing a child.[6] Alleged perpetrators of child sexual

---

[3] *See* DEBRA WHITCOMB, WHEN THE VICTIM IS A CHILD 2–3 (2d ed. 1992) (summarizing seven studies). *See also* Jacqueline Kanovitz, *Hypnotic Memories and Civil Sexual Abuse Trials*, 45 VAND. L. REV. 1185, 1197 (1992) (summarizing studies reported in JOHN BRIERE, THERAPY FOR ADULTS MOLESTED AS CHILDREN: BEYOND SURVIVAL 1 (1989); Phillip M. Coons, et al., *Post-Traumatic Aspects of the Treatment of Victims of Sexual Abuse and Incest*, 12 PSYCHIATRIC CLINICS OF AMERICA 325, 325 (1989); DAVID FINKELHOR, A SOURCEBOOK ON CHILD SEXUAL ABUSE 17–18 (1986) [hereinafter FINKELHOR, SOURCEBOOK]; RUTH S. KEMPE & C. HENRY KEMPE, THE COMMON SECRET: SEXUAL ABUSE OF CHILDREN AND ADOLESCENTS 13 (1984)); AMERICAN BAR ASS'N, GUIDELINES FOR THE FAIR TREATMENT OF CHILD WITNESSES IN CASES WHERE CHILD ABUSE IS ALLEGED 7 (1985) (noting that up to one-third of all female adults were sexually molested during childhood).

The range of these studies' findings may be due in part to the high number of unreported cases, the difficulty in substantiating allegations, *see* FINKELHOR, SOURCEBOOK, at 15, and the lack of a clear definition of child sexual abuse. *See* JILL DUERR BERRICK & NEIL GILBERT, WITH THE BEST OF INTENTIONS 6 (1991).

[4] *See* John E.B. Myers et al., *Expert Testimony in Child Sexual Abuse Litigation*, 68 NEB. L. REV. 1, 53–54 (1989).

> Adult survivors are more depressed, more anxious, have more dissociative and somatic symptoms, and have lower self-esteem . . . . High rates of sexual abuse are found in the histories of patients with conversion reactions, suicide attempts, self-mutilation, multiple personality disorder, chronic pelvic pain, and in women with eating disorders. Childhood sexual abuse is found in a large percentage of adolescent prostitutes and runaways.

Id. (citations omitted).

[5] Several survivors have written books describing their ordeals. Narratives such as these have motivated efforts to prevent sexual abuse of other children. *See, e.g.*, LOUISE ARMSTRONG, KISS DADDY GOODNIGHT (1978); KATHERINE BRADY, FATHER'S DAYS (1979); SUSAN FORWARD & CRAIG BUCK, BETRAYAL OF INNOCENCE: INCEST AND ITS DEVASTATION (1978); SYLVIA FRASER, MY FATHER'S HOUSE: A MEMOIR OF INCEST AND OF HEALING (1987).

[6] Three recent cases serve as illustrations of the dangers of falsely accusing people of child sexual abuse. Virginia Kelly Michaels, a New Jersey preschool teacher, spent five years in prison before her 1988 conviction was reversed. Despite this eventual reversal, she complains, "My name is now associated with this case forever." Weinstein, *supra* note 1, at A1.

Likewise, even though exonerated in court, the defendants in the McMartin Pre-School case suffered long-term consequences as a result of the 1983 allegations of sexual abuse. They spent thousands of dollars on legal bills defending themselves, and many lost their homes and jobs. *Id.*

Finally, William and Kathleen Swan were convicted of raping their three-year-old daughter and her friend; the Ninth Circuit affirmed the dismissal of their habeas corpus petition. *See Swan v. Peterson*, 6 F.3d 1373 (9th Cir. 1993). Media organizations such as "60 Minutes" and *Readers Digest* have reported stories proclaiming the Swans'

abuse are viewed with such moral condemnation that it may take only an accusation to ruin the accused's life.[7] It is unclear how often people are falsely accused or prosecuted for child molestation, and such a figure would be virtually impossible to determine.[8] There is some statistical evidence demonstrating that false accusations are especially common during divorce and custody disputes.[9] As a nation "bottomed on a fundamental value deter-

---

innocence and alleging that their conviction resulted from unreliable hearsay evidence. *See* Richard Seven, *A Persistent Doubt—A Jury Found the Swans Raped Their Child and Another Girl, but Many Dispute the Case*, SEATTLE TIMES, May 23, 1993, at A1.

[7] RICHARD A. GARDNER, TRUE AND FALSE ACCUSATIONS OF CHILD SEX ABUSE xxvii (1992) ("There are hundreds, I am sure, who have committed suicide because of a false sex-abuse accusation . . . . Careers and marriages have been destroyed, reputations ruined, and people are suffering lifelong stigma because of such an accusation.") [hereinafter GARDNER, TRUE AND FALSE ACCUSATIONS].

[8] Experts who argue that there is a high percentage of false allegations often seem reluctant to estimate, even roughly, how high the percentage might be. One such expert is Richard Gardner, a medical doctor and Columbia University professor who has written over 30 books on child psychotherapeutic technique. *See, e.g.*, RICHARD A. GARDNER, SEX ABUSE HYSTERIA: SALEM WITCH TRIALS REVISITED (1991); RICHARD A. GARDNER, THE PARENTAL ALIENATION SYNDROME AND THE DIFFERENTIATION BETWEEN FABRICATED AND GENUINE CHILD SEX ABUSE (1987). Dr. Gardner argues that it is "impossible" to estimate how widespread false accusations of child sex abuse are, and that "anyone who provides specific figures on this subject is 'pulling them out of the sky.'" GARDNER, TRUE AND FALSE ACCUSATIONS, *supra* note 7, at 693. Dr. Gardner is willing to say, however, "I am convinced that there are hundreds (and possibly thousands) of people who are in jail in the United States today who have been convicted of sex crimes that they never committed." *Id.* at xxvii.

Experts who say false accusations are rare, on the other hand, are quick to point to studies or cite figures that support their conclusions. *See, e.g.*, BILLIE WRIGHT DZIECH & CHARLES B. SCHUDSON, ON TRIAL: AMERICA'S COURTS AND THEIR TREATMENT OF SEXUALLY ABUSED CHILDREN 57 (1989) ("The veracity of sexually abused children has been analyzed by researchers, *all* of whom report that false accusations are *extremely* rare.") (emphasis added); WHITCOMB, *supra* note 3, at 6 (arguing that the "most comprehensive" study on the subject, done by the Denver Department of Social Services in 1983 on 576 cases, found that only six percent of allegations are fictitious).

[9] *See, e.g.*, David P.H. Jones & Ann Seig, *Child Sexual Abuse Allegations in Custody or Visitation Disputes*, *in* SEXUAL ABUSE ALLEGATIONS IN CUSTODY AND VISITATION CASES 22–29 (E. Bruce Nicholson & Josephine Bulkley eds., 1988) (finding that 20% of the accusations made during the 20 custody and visitation disputes that they studied were probably fictitious and arguing that, given studies that suggest that 5–10% of all allegations are false, the setting of the divorce and custody dispute increases the likelihood that clinicians will discover false allegations); Jeffrey J. Haugaard, et al., *Children's Definitions of the Truth and Their Competency as Witnesses in Legal Proceedings*, 15 LAW & HUM. BEHAV. 253, 258 (1991) ("If a child who is told to say a certain thing by one parent believes that the parent's instruction makes the child's statement the truth, then the child could speak 'the truth' in a way that could mislead a jury"); Elissa P. Benedek & Diane H. Schetky, *Allegations of Sexual Abuse in Child Custody and Visitation Disputes*, *in* EMERGING ISSUES IN CHILD PSYCHIATRY AND THE LAW 145, 155 (Diane H. Schetky & Elissa P. Benedek eds., 1985) (finding a false-positive incidence of 55% in a study of 18 children evaluated during disputes over custody and visitation); Arthur H. Green, *True and False Allegations of Sexual Abuse in Child Custody Disputes*, 25 J. AM. ACAD. CHILD PSYCHIATRY 449, 449 (1986) (documenting four false allegations in eleven children reported to be sexually abused by the noncustodial parent).

mination . . . that it is far worse to convict an innocent man than to let a guilty man go free,"[10] the states must always be wary of loosening the standards for convicting the accused.[11] Criminal defendants, including those accused of sexually abusing children, represent an insular minority whose rights must be protected against the majority.[12] Courts and legislatures, therefore, must choose the right balance between protecting its young and preserving the rights of the accused.[13]

Unfortunately, the emotional nature of child sexual abuse makes it difficult to have a rational debate over the rules that should govern the prosecutions of alleged abusers.[14] Courts often use

---

[10] In re Winship, 397 U.S. 358, 372 (1970) (Harlan, J., concurring).

[11] Morgan v. Foretich, 846 F.2d 941, 951 (4th Cir. 1988) (Powell, J. concurring in part and dissenting in part).

> Few cases are more difficult to try than one of child abuse where the child is very young and does not testify in court . . . . It must be remembered that, in addition to assuring the fair presentation of a plaintiff's case, the [trial] court has the responsibility of shielding defendants from the admission of unduly prejudicial evidence.

Id. See also Long v. State, 742 S.W.2d 302, 324 (Tex. Crim. App. 1987) ("This area of the law is dominated by emotion, which is understandable in light of the interests society wants to protect—abused children. But . . . of greater concern should be the adherence to our constitutional rights. We cannot ever permit emotion-charged issues to erode our fundamental liberties . . . .").

[12] *See generally* United States v. Carolene Products Co., 304 U.S. 144, 153 n.4 (1938); JOHN HART ELY, DEMOCRACY AND DISTRUST (1980).

[13] *See* South Carolina Dep't of Social Serv. v. John Doe, 355 S.E.2d 543, 546 (S.C. Ct. App. 1987) (citing Commonwealth v. Haber, 505 A.2d 273, 275 n.1 (Pa. Super. Ct. 1986)).

> It is hard for a legislator, who must periodically run for reelection, to vote against a proposed statute that makes it easier to convict persons accused of child abuse, because of the danger that an opportunistic opponent and an unsophisticated electorate may interpret such a vote as being soft on crime in general, or soft on child-abusers in particular. Politically, the safest thing for a legislator to do is to vote in favor of such a statute, regardless of its merits, and leave it to the courts to deal with the ramifications of its enactment.

Id. See also Douglas J. Besharov, *"Doing Something" About Child Abuse: The Need to Narrow the Grounds for State Intervention*, 8 HARV. J.L. & PUB. POL'Y 539, 567–79 (1985) (arguing that America's emotional reaction has led to vague and overbroad laws, excessive intervention into family control of children, and an over-reporting of abuse and neglect).

[14] Several recent books have harshly criticized the legal system's approach to the prosecution of child sexual abuse. In *Accusations of Child Sexual Abuse*, co-author Ralph Underwager, a psychologist who frequently testifies as an expert on behalf of the defense, argues that "[a]n individual accused of sexual abuse of children can expect that the justice system will reflect the society's values and behave in special, unusual, and likely hostile, judgmental fashion from the moment an accusation is made, no matter what the circumstances or merit of the accusation." HOLLIDA WAKEFIELD & RALPH UNDERWAGER, ACCUSATIONS OF CHILD SEXUAL ABUSE 125 (1988) (citation omitted). On the other end of the spectrum, consider this reply: "For the young [alleged victims of sexual abuse], the legal system offers no refuge and little hope for justice." DZIECH & SCHUDSON, *supra* note 8, at 119

See generally John E.B. Myers, The Child Sexual Abuse Literature: A Call for Greater Objectivity, 88 MICH. L. REV. 1709, 1723–32 (1990) (discussing slanted arguments on the American legal system's handling of child sexual abuse prosecutions).

emotional rhetoric in deciding child molestation cases.[15] Prosecutors may pursue child abuse cases when the evidence is insufficient to support a conviction,[16] and they may pour excessive resources into a prosecution only to be forced to drop the charges.[17] The harsh treatment of alleged child abusers and changes in the law favorable to the prosecution have prompted some commentators to question America's approach to prosecuting child sexual abuse.[18]

This Note analyzes the use of hearsay[19] evidence, specifically the out-of-court statements of child declarants, in criminal prose-

---

[15] *See, e.g.*, Sharp v. Commonwealth, 849 S.W.2d 542, 548 (Ky. 1993) (Spain, J., dissenting) ("Regrettably, my brothers of the majority have deemed it necessary to reverse the convictions and sentences of a *child molester* who has raped and sodomized two little girls under eight years of age over whom he held a position of familial authority") (emphasis added). Note the use of the term "child molester" despite the court's decision to reverse the defendant's conviction.

[16] David Fava, a juror who helped acquit Dale Akiki in November 1993 of multiple accounts of sexually abusing and kidnapping nine children, characterized Akiki's trial as "a witch hunt." Michael Granberry, *Case Illustrates Flaws in Child Abuse Trials*, L.A. TIMES, Nov. 29, 1993, at A3. The evidence against Akiki was almost strictly the children's hearsay, which included some rather incredible allegations of ritual abuse. The children were not subjected to cross-examination until four years after they made their original accusations. Akiki was denied bail five times and spent two and one-half years in jail awaiting trial. *See id.*

[17] Prosecutors have been criticized for spending enormous resources to bring cases that produce few results. The most egregious example is the McMartin Pre-School Case. *See supra* note 6. The first trial, lasting two and one-half years and costing $15 million, resulted in no convictions. *See 2 Acquitted of Child Molestation in Nation's Longest Criminal Trial*, N.Y. TIMES, Jan. 19, 1990, at A18. In another example of overzealous prosecution of a child abuse defendant, Delaware prosecutors dismissed a complaint against Michael Gleason in September 1992, in exchange for his signed statement that he raped his daughter repeatedly when she was three. Gleason, an indigent man who was represented by a public defender, said through his lawyer that he had signed the statement confessing that he raped his daughter in order to avoid trial and get out of jail. He had been held for six months on $125,000 bail pending trial. *See* Michael deCourcy Hinds, *Law's Weaknesses: Delaware Case Shows Difficulty Of Proving Child Sex-Abuse*, CHI. TRIB., Nov. 22, 1992, at 11.

In what may be a more typical case, the Commonwealth of Kentucky maintained in its 13-page motion to drop charges against defendant Jack Barry that investigative reports had "substantiated" the allegations. The Commonwealth wrote that it was only dismissing the complaint because the mother wished to "spare her son the trauma of a trial." Barry protested in vain the use of the word "substantiated." He maintained, "My personal feeling is that the prosecution knew that I was innocent, and they gave up because of that." Cary B. Willis, *Sex Abuse Case Against Boy's Father is Dismissed*, COURIER-JOURNAL (LOUISVILLE), Aug. 26, 1993, at 3B.

[18] Commentators such as Dr. Richard Gardner compare the prosecutions of child sexual abuse to the witch trials in Salem, Mass., in 1692. *See* GARDNER, TRUE AND FALSE ACCUSATIONS, *supra* note 7, at xxvi–xxvii; GARDNER, WITCH TRIALS, *supra* note 8. Additionally, an organization known as Victims of Child Abuse Laws ("VOCAL"), has started to lobby legislatures for reforms in the system. VOCAL has over 100 chapters in more than 40 states made up of people who claim to have been wrongfully accused of child abuse. *See* DAVID HECHLER, THE BATTLE AND THE BACKLASH: THE CHILD SEXUAL ABUSE WAR 119 (1988).

[19] "'Hearsay' is a statement, other than one made by the declarant while testifying at the trial or hearing, offered in evidence to prove the truth of the matter asserted."

cutions for the sexual abuse of children.[20] The Note is divided into three parts. Part I discusses the nature of hearsay, focusing on the reliability, desirability, and constitutionality of using child hearsay in criminal prosecutions.

Part II analyzes the legal rules under which states currently admit child hearsay in sexual abuse prosecutions. Over the last twelve years,[21] most states have loosened their evidentiary exclusionary rules regarding admission of hearsay evidence of child declarants. States have expanded two firmly rooted hearsay exceptions, the excited utterance or spontaneous declaration exception[22] and the medical diagnosis or treatment exception,[23] they have increased the use of residual or catch-all hearsay excep-

---

FED. R. EVID. 801(c). For example, a letter to a company complaining of a defective product is hearsay if it is admitted as evidence of the product being defective. On the other hand, the letter is not hearsay if it is admitted solely to show that the customer complained. *See* FED. R. EVID. 801(c) advisory committee's note.

[20] All states forbid many types of lewd conduct with children. For example, Vermont has four such prohibitions with different punishments for each one. Section 2602 provides:

> A person who shall wilfully and lewdly commit any lewd or lascivious act upon or with the body, or any part or member thereof, of a child under the age of sixteen years, with the intent of arousing, appealing to, or gratifying the lust, passions or sexual desires of such person or of such child, shall be imprisoned not less than one year nor more than five years . . . .

VT. STAT. ANN. tit. 13, § 2602 (1993). Under § 3252(a)(4), one who engages in a sexual act with a person under the age of 18 who is either entrusted to the actor's care by law or by virtue of a family relationship can be imprisoned for not more than 20 years, or fined not more than $10,000.00, or both. *See id.* at § 3252(a)(4). Under § 3252(b), a person who engages in a sexual act with another person under the age of 16 where the victim is entrusted to the actor's care will be imprisoned for not more than 35 years, or fined not more than $25,000.00, or both. *See id.* at § 3252(b). Under § 3253, a person shall be punishable by a maximum sentence of life imprisonment or a fine of not more than $ 50,000,00, or both if the person commits sexual assault on a victim under the age of 10 and the actor is at least 18 years of age. *See id.* at § 3253.

[21] In 1982, only four states looked beyond the spontaneous declaration hearsay exception in FED. R. EVID. 803(2) to admit child hearsay: Kansas had enacted a tender years exception in 1982, *see* KAN. STAT. ANN. ch. 60, art. 460 § dd (Supp. 1982); Washington had proposed a tender years exception, *see* 1982 Wash. Legis. Serv. ch. 129, § 2 (West) (modified and enacted as WASH. REV. CODE § 9A.44.120. (1994); Wisconsin courts used Wisconsin's residual hearsay exceptions, *see* WIS. R. EVID. 908.03(24) and 908.045(6); and Michigan courts had developed a common law tender years exception for child victims of sex crimes (similar exception enacted in 1991 as MICH. R. EVID. 803A). For an analysis of the benefits and drawbacks of the systems in effect in 1982, see Judy Yun, Note, *A Comprehensive Approach to Child Hearsay Statements in Sex Abuse Cases*, 83 COLUM. L. REV. 1745 (1983).

[22] "A statement relating to a startling event or condition made while the declarant was under the stress of excitement caused by the event or condition" is not excluded by the hearsay rule even if the declarant is available as a witness. FED. R. EVID. 803(2).

[23] "Statements made for purposes of medical diagnosis or treatment and describing medical history, or past or present symptoms, pains, or sensations, or the inception or general character of the cause or external source thereof insofar as reasonably pertinent to diagnosis or treatment" are not excluded by the hearsay rule even if the declarant is available as a witness. FED. R. EVID. 803(4).

tions,[24] and they have passed legislation codifying tender years hearsay exceptions for statements made by child victims.[25] Part II argues that the expanding scope of the traditional exceptions and the use of the residual exceptions are problematic. There is little reason to believe that statements admitted under the expanded traditional definitions are trustworthy, and the residual exceptions have insufficient guidelines and discourage children from testifying. Part II also evaluates how well the different state tender years statutes, passed by thirty-four states, achieve three objectives: (1) ensuring that only trustworthy hearsay is admitted, (2) encouraging the child to testify, and (3) being as inclusive as possible without violating either of the first two objectives. Part II concludes that none of the current statutes is optimal.

Part III proposes a model tender years hearsay exception statute ("Model Statute") that divides children's out-of-court statements into four categories of trustworthiness.[26] The Model Statute's provisions would: (1) admit the most reliable statements, even if the child is available but refuses to testify, without requiring corroborative evidence of the act; (2) allow the next most reliable statements if the child either testifies or is unavailable and there is corroboration; (3) admit slightly less reliable

---

[24] Statements not specifically covered by another rule are admissible if:

> [They have] equivalent circumstantial guarantees of trustworthiness, if the court determines that (A) the statement is offered as evidence of a material fact; (B) the statement is more probative on the point for which it is offered than any other evidence which the proponent can procure through reasonable efforts; and (C) the general purposes of these rules and the interests of justice will best be served by admission of the statement into evidence.

FED. R. EVID. 803(24), 804(b)(5). Rule 804(b)(5) is used when the declarant is unavailable; Rule 803(24) is used regardless of whether the declarant is available.

[25] *See, e.g.*, WASH. REV. CODE § 9A.44.120 (1991):

> A statement made by a child when under the age of ten describing any act of sexual contact . . . is admissible in evidence in . . . criminal proceedings . . . if:
>
> (1) The court finds, in a hearing conducted outside the presence of the jury, that the time, content, and circumstances of the statement provide sufficient indicia of reliability; and
>
> (2) The child either:
>
> (a) Testifies at the proceeding; or
>
> (b) Is unavailable as a witness: PROVIDED, That when the child is unavailable as a witness, such statement may be admitted only if there is corroborative evidence of the act.
>
> A statement may not be admitted under this section unless the proponent of the statement makes known to the adverse party his intention to offer the statement and the particulars of the statement sufficiently in advance of the proceedings to provide the adverse party with a fair opportunity to prepare to meet the statement.

See infra note 156 for a list of all statutory tender years hearsay exceptions.

[26] The text of the Model Statute is included in the Appendix for ease of reference.

hearsay only if the child testifies and is subject to cross-examination at the proceeding or by means of a videotaped deposition or closed circuit television; and (4) prohibit the admission of the least reliable statements. Part III concludes that the Model Statute achieves all three of the objectives of an ideal hearsay statute. Its strict guidelines would prevent courts from admitting unreliable hearsay, while its allowance of more hearsay if the child testifies would motivate prosecutors and families to encourage children to testify and to make themselves available for cross-examination. Additionally, the Model Statute would allow the admission of a maximum amount of reliable testimony because it would apply regardless of whether the child testifies. The Model Statute, therefore, balances the advantages to the prosecution of admitting trustworthy child hearsay with the need to ensure a fair and accurate trial for the accused.

I. The Value and Problems of Using Hearsay Evidence in Child Sexual Abuse Cases

A. *Nature of Hearsay*

Child sexual abuse is an extremely difficult crime to prosecute.[27] Abusers may leave no physical marks on their victims,[28] and children often do not resist outwardly or physically.[29] Accordingly, there is usually little physical evidence to corroborate the child's allegations,[30] and the child-victim is often the only

---

[27] Less than 10% of allegations of child sexual abuse lead to criminal charges. *See* Weinstein, *supra* note 1, at A1. "Many [of these] cases involve situations in which the prosecutor or others believe abuse has occurred, but criminal charges cannot be filed because of insufficient admissible evidence, unavailability of necessary witnesses" or other factors. *Id.* About 85% of the cases that are filed eventually lead to either convictions or guilty pleas. This is approximately the same percentage as for other crimes. *See id.*

[28] Child sexual abuse often involves petting, exhibitionism, fondling, and oral copulation, activities that generally leave no physical evidence. *See In re* A.S.W. and E.W., 834 P.2d 801, 804 (Alaska 1992) (citations omitted).

[29] There are many theories on why children do not resist sexual abuse: they are conditioned to comply with authority, they are fearful of threats, they are susceptible to bribes and promises of rewards, they are curious, and they confuse sexual contact with feelings of being loved. *See generally* Yun, *supra* note 21; The Sexual Victimology of Youth (Leroy G. Schultz ed., 1980); Kee MacFarlane, *Sexual Abuse of Children*, *in* The Victimization of Women 81, 88 (Jane Roberts Chapman & Margaret Gates eds., 1978).

[30] Physical or laboratory evidence of sexual abuse is found in only 10–50% of cases. *See* Myers et al., *supra* note 4, at 34 n.120. For example, in a study of 311 cases, genital trauma was found in 16% of cases, and non-genital trauma was found in an

witness.[31] Very young children, however, may be unreliable witnesses. It is not unusual for children to recant their prior accusations when cross-examined, whether or not the accusations are true.[32] Further, abused children are often reluctant to identify their attacker in open court,[33] and the child's family may refuse to allow the child to testify at all because they believe testifying would be harmful to the child.[34] State law may compel the court to declare the child incompetent to testify, despite years of contrary precedent.[35] An absolute prohibition against using hearsay

---

additional 16%. Genital trauma was twice as likely if the perpetrator was a stranger. *See id.* Some studies indicate that physical evidence of sexual assault does not result in a statistically significant increase in the likelihood of conviction. *See id.* at 35 n.121 (citations omitted).

[31] *See, e.g.*, State v. J.C.E., 767 P.2d 309, 311 (Mont. 1988).

[32] *See* Commonwealth v. Costello, 582 N.E.2d 938, 941 (Mass. 1991) ("[C]hild abuse victims, susceptible to parental influence, [are] likely to change testimony . . . .") (citations omitted); State v. Storch, 612 N.E.2d 305, 315 (Ohio 1993).

[33] One commentator contends that the state's difficulty in getting children to testify often results in the state's not being able to proceed against the accused. *See* Sharon P. Brustein, Coy v. Iowa*: Should Children Be Heard and Not Seen?*, 50 U. PITT. L. REV. 1187, 1191 n.26 (1989). *See also* State v. Sheppard, 484 A.2d 1330, 1333 (N.J. 1984) (noting that the prosecutor reported that nearly 90% of the child abuse cases dismissed were dismissed because the children could not "deal with the prospect of facing fathers, stepfathers, relatives, and strangers in a courtroom setting . . . .").

[34] Several commentators suggest that parents and victims view testifying as traumatic. *See, e.g.*, BILLIE WRIGHT DZIECH & CHARLES B. SCHUDSON, ON TRIAL: AMERICA'S COURTS AND THEIR TREATMENT OF SEXUALLY ABUSED CHILDREN 12 (2d ed. 1991) ("For *most* victims, confrontation with the legal system is a second and separate trauma, a process of revictimization . . . . Parents who expose their children to the system *overwhelmingly* regret their decision . . . .") (emphasis added).

[35] For over 100 years, it has been widely accepted in America that a court may find a child of tender years competent as a witness. *See, e.g.*, Wheeler v. United States, 159 U.S. 523 (1895); State v. Whittier, 21 Me. 341, 347 (1842); Washburn v. People, 10 Mich. 372 (1862); State v. Edwards, 79 N.C. 648 (1878). *See generally* South Carolina Dep't of Social Serv. v. John Doe, 355 S.E.2d 543, 547 (S.C. Ct. App. 1989).

While the states employ different tests to determine if children are competent to testify, the general trend is to permit children to testify, especially in sexual abuse cases. *See, e.g.*, MINN. STAT. § 595.02, subd. 1(1) (1992), amended by § 595.02, subd. 1(m) (Supp. 1993) (reversing the prior statute by codifying a presumption of competence of children to testify); MO. REV. STAT. § 491.060 (1986) (amended 1993) (providing that a child under ten who is the victim of sexual abuse is allowed to testify without qualification in a judicial proceeding involving that offense); COLO. REV. STAT. § 13-90-106(1)(b)(II) (1985 Cum. Supp.) (amended 1989) (providing that a child under ten can testify in a civil or criminal proceeding for sexual abuse if the child is able to describe or relate the events or facts in language appropriate for a child of the age); IOWA R. EVID. 601 (providing that a child is presumed competent, but if the competency is questioned, then the court must determine if the child (i) is mentally capable of understanding the questions asked, (ii) is able to formulate and communicate intelligent answers, and (iii) understands the moral duty to tell the truth).

Some states, however, still presume children are incompetent to testify unless the court determines that the child is competent by applying specific standards. *See, e.g.*, OHIO R. EVID. 601(A) (Baldwin 1993) (children under ten who appear incapable of receiving and relating truly just impressions of facts and transactions are not competent); People v. Wolfe, 531 N.E.2d 152, 153 (Ill. App. 1988) (holding that a minor can

evidence, therefore, could make crimes of child molestation nearly impossible to prosecute.[36] Many prosecutors prefer to use hearsay evidence, rather than live testimony from the child-victim, to support their cases.[37] Juries may react with skepticism and disbelief when children testify, and they are sometimes more willing to accept children's spontaneous out-of-court statements as true.[38] Juries frequently find the adults to whom children report their allegations to be persuasive witnesses, whether the adult is a sympathetic parent or baby-sitter, a dispassionate doctor or nurse, or a respected police officer.[39] Further, an adult may be able to recount the child's story more coherently and consistently than the child.[40] The use of hearsay testimony also minimizes the chance that the child will recant his or her allegations of sexual abuse once in court.[41]

The Anglo-American legal tradition, however, disfavors hearsay. Since the declarant does not testify under oath in the presence of the fact finder, he or she is unburdened by the solemnity of the trial and the possibility of public disgrace, thereby escaping the influence of these factors and their accompanying effects on witnesses' veracity.[42] Under our legal system, the trier of fact is entrusted with the task of evaluating the perception, memory, narration, and sincerity of witnesses; assessing these traits is more difficult to do with hearsay because the trier of fact cannot observe the demeanor of witnesses.[43] Hearsay is not subject to cross-examination, which is believed to be vital in exposing imperfections in a witness's account.[44]

---

testify if the court determines that he or she is sufficiently mature to (i) receive correct impressions, (ii) recollect those impressions, (iii) understand questions and narrate answers intelligibly, and (iv) appreciate the moral duty to tell the truth).

[36] "Unless some other form of evidence can be presented, those who abuse small children will be at liberty to do so with utter impunity." State v. Storch, 612 N.E.2d 305, 309 (Ohio 1993).

[37] *See generally* Nancy H. Baughan, White v. Illinois*: The Confrontation Clause and the Supreme Court's Preference for Out-of-Court Statements*, 46 VAND. L. REV. 235 (1993).

Other prosecutors hold a contrary view. According to Rebecca Roe, senior deputy prosecutor and head of the special assault unit for Seattle's King County, "[i]t is our strong preference to get the child to testify in court." *See* Seven, *supra* note 6, at A1.

[38] *See* Yun, *supra* note 21, at 1746, 1751.

[39] *See* Myrna S. Raeder, White*'s Effect on the Right to Confront One's Accuser: Are Expanding Interpretations of Firmly Rooted Hearsay Exceptions Collapsing Confrontation Clause Analysis?*, 7 CRIM. JUST. 2, 5 (1993).

[40] *See* South Carolina Dep't of Social Serv. v. John Doe, 355 S.E.2d 543, 549 (S.C. App. 1987).

[41] *See supra* note 32 and accompanying text.

[42] *See* FED. R. EVID. art. VIII advisory committee's note.

[43] *Id. See generally* Universal Camera Corp. v. NLRB, 340 U.S. 474, 495–96 (1951).

[44] *See generally* FED. R. EVID. art. VIII advisory committee's note ("Emphasis on the

B. *Confrontation Clause Analysis*

The use of hearsay in criminal prosecutions also raises constitutional problems.[45] The Confrontation Clause of the Sixth Amendment to the United States Constitution, applicable to the states through the Fourteenth Amendment, provides, "In all criminal prosecutions, the accused shall enjoy the right . . . to be confronted with the witnesses against him."[46] It is unclear what the drafters intended this clause to mean.[47] Read literally, the clause would forbid the use of any hearsay made by a declarant not actually called as a witness at the trial.[48] The Supreme Court, however, rejected this literal approach one hundred years ago in *Mattox v. United States*, 156 U.S. 237 (1895). In admitting a statement by a witness who died before the defendant's retrial, the Court explained that "general rules of law of this kind, however beneficent in their operation and valuable to the accused, must occasionally give way to considerations of public policy and the necessities of the case."[49]

In a series of cases since *Mattox*, the Court has attempted to establish the proper balance between the values of the Confrontation Clause, public policy, and the necessities of criminal prosecutions.[50] Supreme Court precedent suggests that the Confrontation Clause protects two basic rights for the accused: the right

---

basis of the hearsay rule today tends to center upon the condition of cross-examination."); Ohio v. Roberts, 448 U.S. 56, 64–65 (1980). *See also* MORRIS K. UDALL ET AL., LAW OF EVIDENCE § 121, at 235 (3d ed. 1991) ("Without the testing of cross-examination, it is often impossible to assess the weight reasonably to be attached to evidence . . . ."). According to Wigmore, cross-examination "is beyond any doubt the greatest legal engine ever invented for the discovery of truth." 5 JOHN HENRY WIGMORE, EVIDENCE IN TRIALS AT COMMON LAW § 1367 (James H. Chadbourn rev., 1974).

[45] "Although both [the Confrontation Clause and the hearsay rule] protect similar values, each sets independent prohibitions on admissibility." Swan v. Peterson, 6 F.3d 1373, 1379 (9th Cir. 1993).

[46] U.S. CONST. amend. VI. *See* Pointer v. Texas, 380 U.S. 400, 403 (1965) ("[T]he Sixth Amendment's right of an accused to confront the witnesses against him is likewise a fundamental right and is made obligatory on the States by the Fourteenth Amendment . . . .").

[47] There is almost no evidence of the drafters' intent; Congress passed the amendment without discussion. *See* California v. Green, 399 U.S. 149, 176 n.8 (1970) (Harlan, J., concurring).

There are many examples of the confrontation right in both Biblical and Roman history. It was firmly established in England by the year 1200, before the emergence of the right to trial by jury. For a history of the right, *see* State v. Lanam, 459 N.W.2d 656, 662–65 (Minn. 1990) (Kelley, J., dissenting).

[48] *See* White v. Illinois, 112 S. Ct. 736, 747 (1992) (Thomas, J., concurring).

[49] 156 U.S. at 243.

[50] *See, e.g.*, Ohio v. Roberts, 448 U.S. 56, 64 (1980) ("True to the common-law

to physically face those witnesses who testify against him in court and the right to conduct cross-examination.[51] Neither of these rights, however, has been deemed absolute.[52] In *Ohio v. Roberts*, 448 U.S. 56 (1980), the Supreme Court rejected the notion that cross-examination is required of every statement introduced at trial since that would eliminate every exception to the rule against hearsay.[53] Instead, the Court noted that the Confrontation Clause operates in two distinct ways to restrict the range of admissible hearsay. First, the Sixth Amendment establishes a rule of necessity: "[T]he prosecution must either produce, or demonstrate the unavailability of, the declarant whose statement it wishes to use against the defendant."[54] Second, once the witness is proved to be unavailable, his statements are admissible only if they bear adequate "indicia of reliability."[55] Reliability can be automatically inferred if the hearsay falls within one of the firmly rooted hearsay exceptions.[56] In all other cases, courts should exclude hearsay testimony unless there is a showing of "particularized guarantees of trustworthiness."[57]

Three recent Supreme Court cases have modified and clarified the guidelines for admitting hearsay evidence in child sexual assault prosecutions. In *Idaho v. Wright*, 497 U.S. 805 (1990), the Court reversed a conviction based in part on child hearsay because the trial court did not sufficiently evaluate the independent trustworthiness of the child's statement. The Court listed five factors for future trial courts to consider when evaluating whether children's out-of-court statements exhibit "particularized guarantees of trustworthiness:"[58] spontaneity, consistency, the declarant's mental state, the use of terminology unexpected of a child of similar age, and the existence of any motive to

---

tradition, the process has been gradual, building on past decisions, drawing on new experience, and responding to changing conditions.").

[51] *See, e.g.*, Pennsylvania v. Ritchie, 480 U.S. 39, 51 (1987).

[52] A defendant does not have an absolute right to conduct unlimited cross-examination. *See* Rock v. Arkansas, 483 U.S. 44, 55 (1987) ("[T]he right to present relevant testimony is not without limitation. The right 'may, in appropriate cases, bow to accommodate other legitimate interests in the criminal trial process.'") (quoting Chambers v. Mississippi, 410 U.S. 284, 295 (1973)).

[53] *See Roberts*, 448 U.S. at 63.

[54] *Id.* at 65 (citations omitted).

[55] *Id.* at 66 (citations omitted).

[56] *See infra* notes 114–140 and accompanying text.

[57] *Roberts*, 448 U.S. at 66.

[58] *Wright*, 497 U.S. at 818 (quoting *Roberts*, 448 U.S. at 66).

fabricate.[59] Corroborative evidence of the act is not a permissible factor in determining whether a statement is trustworthy.[60]

The Court has recently narrowed *Roberts'* necessity requirement. Under *White v. Illinois*, 112 S. Ct. 736 (1992), out-of-court statements that fall within a firmly rooted exception to the hearsay prohibition are admissible in child sexual abuse cases without cross-examination regardless of whether the prosecution can prove that the hearsay is necessary because the declarant is unavailable.[61] Neither the residual exceptions nor the tender years statutory exceptions are considered firmly rooted exceptions.[62] Some commentators have criticized *Wright* and *White* for creating a discrepancy in a defendant's confrontation rights. *Wright* and *White* permit prosecutors to forgo live testimony entirely for statements within firmly rooted exceptions without proving unavailability, but *Maryland v. Craig*, 497 U.S. 836 (1990), bars them from regulating live testimony unless necessity is shown. This is unfair to defendants since they would prefer cross-examination via a videotaped deposition or closed circuit television to no cross-examination at all.[63]

In *Maryland v. Craig*, the Court changed the rules regarding the testimony of allegedly abused children. Since children may respond to confrontation differently than adults,[64] courts may use

---

[59] *Wright*, 497 U.S. at 821–22 (citations omitted).

[60] *Id.* at 822–23.

[61] *See White*, 112 S. Ct. at 743–44. Under *Roberts*, 448 U.S. 56, 74 (1980), a witness is not "unavailable" for the purpose of the Confrontation Clause unless the prosecutor has made a good-faith effort, evaluated under a reasonableness standard, to obtain the witness' presence at trial.

[62] *See Wright*, 497 U.S. at 817; Swan v. Peterson, 6 F.3d 1373, 1379 (9th Cir. 1993); State v. Buller, 484 N.W.2d 883 (S.D. 1992).

[63] *See Wright*, 497 U.S. at 816–17. In *Wright*, the Court assumed that the three-year-old declarant was incapable of communicating to a jury and hence unavailable as a witness within the meaning of the Confrontation Clause. *See id.*

Since *White* does not explain why it does not apply the *Roberts* necessity requirement, it is unclear how courts will interpret *White*. *See* Levine, *supra* note 2, for a critique of *White*'s ambiguity.

In *United States v. Inadi*, 475 U.S. 387, 394–400 (1986), a case involving co-conspirators' statements, the Court suggested that *Roberts'* unavailability requirement does not apply if (1) the hearsay derives independent evidentiary value from its context, (2) the burden of proving unavailability is significant, and (3) the benefit of such a requirement is slight. Although children's hearsay often derives meaning from its circumstances, the burden of proving the child's unavailability is not heavy and the benefit to the defendant of allowing him to confront the child is potentially great. Therefore, it is unclear whether *Inadi* suggests *Roberts'* necessity requirement should apply to child sexual abuse prosecutions. *See* Note, *The Supreme Court, 1989 Term: Leading Cases*, 104 HARV. L. REV. 129, 138 n.77 (1990).

[64] "To be sure, face-to-face confrontation may be said to cause trauma for the very purpose of eliciting truth . . . [but] where face-to-face confrontation causes significant

special procedures to ease the trauma of testifying if the court finds that the child would suffer "more than de minimis [emotional distress]" and hence could not "reasonably communicate" if forced to testify in front of the accused in court.[65] The issue in *Craig* was testifying via closed circuit television. The Court held that closed circuit television was not "out-of-court," and hence raised neither hearsay nor Confrontation Clause problems.[66] As of 1994, thirty-seven states allowed videotaped testimony by children in child abuse cases,[67] and thirty states permitted alleged child abuse victims to testify via one-way or two-way closed circuit television.[68] In short, the Confrontation Clause has proved

---

emotional distress in a child witness, there is evidence that such confrontation would in fact disserve the Confrontation Clause's truth-seeking goal." *Craig*, 497 U.S. at 856–57.

[65] *Id.* "Emotional distress" has been defined as more than "'mere nervousness or excitement or some reluctance to testify.'" *Id.* at 856 (quoting Wildermuth v. State, 530 A.2d 275, 289 (Md. 1987)).

[66] *Id.* at 851–52. Most state courts have expanded *Craig*'s holding to videotaped depositions. Since videotaped testimony does not take place in a courtroom in front of the fact finder, this testimony should perhaps be classified as hearsay and, accordingly, be subject to the hearsay rules. Most state courts that have considered the matter, however, have held that a videotaped deposition that mirrors trial proceedings is the functional equivalent of live testimony. *See, e.g.*, State v. Thomas, 442 N.W.2d 10, 12–19 (Wis.), *cert. denied*, 493 U.S. 867 (1989) (stating videotaped testimony is the "functional equivalent to live in-court testimony . . . ."); State v. Jarzbek, 529 A.2d 1245, 1252 (Conn. 1987), *cert. denied*, 484 U.S. 1061 (1988) (same holding); State v. Cooper, 353 S.E.2d 451, 454 (S.C. 1987) (upholding use of videotaped testimony of three-year-old child). *But see* State v. Johnson, 729 P.2d 1169, 1173 (Kan. 1986), *cert. denied*, 481 U.S. 1071 (1987) ("[V]ideotaped testimony is hearsay . . . .").

[67] ALA. CODE § 15-25-2 (1994); ARIZ. REV. STAT. ANN. § 13-4253(B) (1994); ARK. CODE ANN. § 16-44-203 (Michie 1993); CAL. PENAL CODE § 1346 (West 1994); COLO. REV. STAT. §§ 18-3-413, 18-6-401.3 (1994); CONN. GEN. STAT. § 54-86g (1994); DEL. CODE ANN. tit. 11, § 3511 (1993); FLA. STAT. ANN. § 92.53 (West 1994); HAW. REV. STAT. § 626-1 (1994); ILL. REV. STAT. ch. 38, para. 106A-2 (1994); IND. CODE § 35-37-4-8 (1994); IOWA CODE § 910A.14 (1994); KAN. STAT. ANN. § 38-1558 (1993); KY. REV. STAT. ANN. § 421.350 (Michie/Bobbs-Merrill 1993); MASS. GEN. LAWS ANN. ch. 278, § 16D (West 1994); MICH. COMP. LAWS ANN. § 600.2163a (West 1994); MINN. STAT. § 595.02(4) (1994); MISS. CODE ANN. § 13-1-407 (1993); MO. REV. STAT. §§ 491.680 (1993); MONT. CODE ANN. § 46-15-402 (1993); NEB. REV. STAT. § 29-1926 (1993); NEV. REV. STAT. § 174.227 (1993); N.H. REV. STAT. ANN. § 517:13-a (Supp. 1993); N.M. STAT. ANN. § 30-9-17 (Michie 1994); OHIO REV. CODE ANN. § 2907.41 (Anderson 1993); OKLA. STAT. tit. 22, § 753(C) (1993); OR. REV. STAT. § 40.460(24) (1993); 42 PA. CONS. STAT. §§ 5982, 5984 (1994); R.I. GEN. LAWS § 11-37-13.2 (1993); S.C. CODE ANN. § 16-3-1530(G) (Law. Co-op. 1993); S.D. CODIFIED LAWS ANN. § 23A-12-9 (1994); TENN. CODE ANN. § 24-7-116 (1994); TEX. CODE CRIM. PROC. ANN. art. 38.071, § 4 (West 1994); UTAH R. CRIM. PROC. 15.5 (1994); VT. R. EVID. 807(d) (1993); WIS. STAT. ANN. §§ 967.04(7)–(10) (West 1994); WYO. STAT. § 7-11-408 (1994). *See also* Maryland v. Craig, 497 U.S. 836, 853 n.2 (1990) (citing statutes).

[68] ALA. CODE § 15-25-2 (1994); ALASKA STAT. § 12.45.046 (1993); ARIZ. REV. STAT. ANN. § 13-4253 (1994); CAL. PENAL CODE § 1347 (West 1994); CONN. GEN. STAT. § 54-86g (1994); FLA. STAT. ANN. § 92.54 (West 1994); GA. CODE ANN. § 17-8-55 (1994); HAW. REV. STAT. § 626-1 (1994); IDAHO CODE § 19-3024A (1994); ILL. REV. STAT. ch. 38, para. 106B-1 (1994); IND. CODE § 35-37-4-8 (1994); IOWA CODE

to be an easily surmountable barrier to the admission of out-of-court statements by child declarants in sexual abuse prosecutions.

C. *Trustworthiness of Child Hearsay*

Relying on hearsay to obtain child abuse convictions is problematic from the perspective of trial court fact finders. Too often courts will reduce the determination of the admissibility of child hearsay to a question of whether a child can differentiate the "truth" from a "lie."[69] This reduction is not particularly helpful, especially for child declarants, because people rarely make statements that are completely true or totally false. Instead, people tend to exaggerate, forget, imagine, and confuse events when they relate them. Hearsay rules are not intended to determine whether statements are true. The fact finder makes those determinations. Instead, hearsay rules act as a preliminary screen to exclude statements that are inherently untrustworthy.[70]

Hearsay rules are designed to ensure that only reliable evidence is used to convict defendants. There are four dangers traditionally associated with courts' reliance on hearsay. Hearsay testimony may be erroneous because it is based upon: (1) an inaccurate impression of objective reality, (2) a false belief at a later time because of memory deficiencies or fantasies, (3) an imperfect communication or interpretation between the declarant and the person to whom the statement was made, or (4) an intentional falsification by either the declarant or the person to whom the statement was made.[71] Children's out-of-court statements, admitted as hearsay, are particularly susceptible to problems of untrustworthiness. Experts assert that young children's

---

§ 910A.14 (1994); KAN. STAT. ANN. § 38-1558 (1993); KY. REV. STAT. ANN. § 421.350 (Michie/Bobbs-Merrill 1993); LA. REV. STAT. ANN. § 15:283 (West 1994); MD. CODE ANN., CTS. & JUD. PROC. § 9-102 (1994); MASS. GEN. LAWS ANN. ch. 278, § 16D (West 1994); MINN. STAT. § 595.02(4) (1994); MISS. CODE ANN. § 13-1-405 (1993); N.J. REV. STAT. § 2A:84A-32.4 (1994); N.Y. CRIM. PROC. LAW § 65.00 (McKinney 1994); OHIO REV. CODE ANN. § 2907.41 (Anderson 1993); OKLA. STAT. tit. 22, § 753(B) (1994); OR. REV. STAT. § 40.460(24) (1993); 42 PA. CONS. STAT. §§ 5982, 5985 (1988); R.I. GEN. LAWS § 11-37-13.2 (1993); TEX. CODE CRIM. PROC. ANN. art. 38.071, § 3 (West 1994); UTAH R. CRIM. PROC. 15.5 (1994); VT. R. EVID. 807(d) (1993); VA. CODE ANN. § 18.2-67.9 (Michie 1994). *See also* Maryland v. Craig, 497 U.S. at 854 nn.3–4 (citing statutes).

[69] *See generally* GARDNER, TRUE AND FALSE ACCUSATIONS, *supra* note 7.

[70] *See* FED. R. EVID. art. VIII advisory committee's note.

[71] *See* Garcia v. State, 792 S.W.2d 88, 92–93 (Tex. Crim. App. 1990) (Clinton, J., dissenting) (citing STEVEN GOODE ET AL., GUIDE TO THE TEXAS RULES OF EVIDENCE: CIVIL AND CRIMINAL § 8.01.1 (1988)).

eyewitness testimony may be inaccurate.[72] Even if children perceive incidents correctly, their memories of incidents may be subject to distortions or fantasies. Many scholars believe that children are more vulnerable to suggestion and more easily influenced by various authority figures than are adults.[73] Accordingly, children may alter their comments to gain or maintain the affection of authority figures.[74]

Vigorous questioning of a child by an adult may produce distorted accounts of the abuse. An adult who suspects sexual abuse may aggressively try to convince the child to tell the "truth." Confronted with suggestive questions and talk of "bad" things, children can figure out what grown-ups want to hear.[75] To children, the "truth" may simply become a code word for sexual abuse. The problems with suggestive interviewing practices have been acknowledged;[76] some commentators and courts have criticized social workers for sloppy investigative techniques and for eliciting fantasies from children.[77] Even doctors, who can

---

[72] *See generally* Matthews v. State, 515 N.E.2d 1105, 1106 (Ind. 1987); Gail S. Goodman & Rebecca S. Reed, *Age Differences in Eyewitness Testimony*, 10 LAW & HUMAN BEHAVIOR 317 (1986); Ronald L. Cohen & Mary Anne Harnick, *The Susceptibility of Child Witnesses to Suggestion*, 4 LAW & HUM. BEHAV. 201 (1980). *But see* Gail S. Goodman & Vicki S. Helgeson, *Child Sexual Assault: Children's Memory and the Law*, 40 U. MIAMI L. REV. 181 (1985).

[73] *See, e.g.*, Maryland v. Craig, 497 U.S. 836, 860 (1990) (Scalia, J., dissenting); Garcia v. State, 792 S.W.2d at 93 (Tex. Crim. App. 1990) (Clinton, J., dissenting); GARDNER, TRUE AND FALSE ACCUSATIONS, *supra* note 7, at 102. Other experts, however, differ on the degree to which children are more suggestible than adults. *See, e.g.*, Myers et al., *supra* note 4, at 100–03 (citing studies showing that children's memory abilities are different, but not necessarily less reliable, than those of adults).

[74] "[Child sexual abuse] cases raise the specter of fabricated or exaggerated charges, either because of pressure from adults in the child's life or from adult investigators." Guam v. Ignacio, 10 F.3d 608, 612 (9th Cir. 1993) (citing Eileen A. Scallen, *Constitutional Dimensions of Hearsay Reform: Toward a Three-Dimensional Confrontation Clause*, 76 MINN. L. REV. 623, 652 (1992)).

[75] For example, Stephen Ceci, a psychologist at Cornell University, completed a study showing that preschool-age children exposed to repeated suggestive questioning can fabricate stories about events that never occurred. *See* Granberry, *supra* note 16. *See also* GARDNER, TRUE AND FALSE ACCUSATIONS, *supra* note 7, at 114 ("In divorce situations children predictably lie and say to each parent what that parent wants to hear, especially regarding criticisms of the other . . . .").

[76] *See* GARDNER, TRUE AND FALSE ACCUSATIONS, *supra* note 7, at 274; John R. Christiansen, *The Testimony of Child Witnesses: Fact, Fantasy, and the Influence of Pretrial Interviews*, 62 WASH. L. REV. 705 (1987); Thomas L. Feher, *The Alleged Molestation Victim, the Rules of Evidence, and the Constitution: Should Children Really Be Seen and Not Heard?*, 14 AM. J. CRIM. L. 227, 230–33 (1988); Paul R. Lees-Haley, *Innocent Lies, Tragic Consequences: The Manipulation of Child Testimony*, TRIAL, Apr. 1988, at 37; Raymond K. Ramella, Casenote, *The Confrontation Clause and Hearsay in Child Abuse Cases*, 25 CREIGHTON L. REV. 1043, 1048 (1992).

[77] *See, e.g.*, HECHLER, *supra* note 18, at 128–29; State v. Jones, 625 So. 2d 821, 825 (Fla. 1993).

testify to statements made by children pertinent to medical diagnosis or treatment under an exception to the hearsay prohibition, are not well trained in interviewing children on sexual abuse. In fact, most medical students receive little or no training on the subject.[78] As a result, many courts require cross-examination of the child regarding any statement made in response to investigative questioning.[79]

Children's memories may become distorted over time.[80] Consequently, children sometimes have difficulty maintaining a distinction in their own minds between real and fantasized sexual abuse.[81] Along with the passage of time, stress may adversely affect the veracity of children's memories.[82] In many cases, the adults to whom the children confess harbor feelings regarding the named perpetrator. Children may know of, and their allegations may be affected by, these feelings. This is especially true when the alleged abuser is a family member or friend.[83]

A child may lie or exaggerate in reporting an incident of abuse. Courts may fail to consider the possibility that children who know what a lie is may still lie, and lie well.[84] Thus, while courts tend to trust the testimony of young children, whom they believe "are unable to practice real deception,"[85] one study indicated that even children as young as three years old can lie

---

[78] *See* HECHLER, *supra* note 18, at 24–25.

[79] *See, e.g.*, State v. R.C., 494 So. 2d 1350 (La. Ct. App. 1986); *In re* Troy, 842 P.2d 742, 745–46 (N.M. Ct. App. 1992); Burke v. State, 820 P.2d 1344, 1348 (Okla. Crim. App. 1991).

Cf. Idaho v. Wright, 497 U.S. 805, 818–19 (1990) (acknowledging that leading questions may make a child's statements unreliable); Maryland v. Craig, 497 U.S. 836, 868 (1990) (Scalia, J.) (cautioning that courts should be particularly insistent in obtaining cross-examination in child sexual abuse cases because of the reliability problems created by children's suggestibility).

[80] *See* GARDNER, TRUE AND FALSE ACCUSATIONS, *supra* note 7, at 101.

[81] Some children's descriptions of abuse include unlikely, incredible allegations. For example, Dale Akiki was accused and eventually acquitted of conspiring with his wife and another baby sitter to subject children in a church nursery school to rituals involving urine, feces, water torture, the slaughter of an elephant, and the mutilation and drinking of the blood of a giraffe and a rabbit. *See* Granberry, *supra* note 16. Tales of ritual abuse have occurred in other cases too. *See, e.g.*, the McMartin Pre-School case, *supra* note 17. Many experts maintain that there has never been a verified case of ritual abuse. *See* Granberry, *supra* note 16.

Overall, however, experts suggest that children do seem to be able to differentiate fact from fantasy. As an example, while children will pretend to sip tea during a "tea party," they do not try to eat plastic cookies. *See* Myers et al., *supra* note 4, at 103.

[82] *See* GARDNER, TRUE AND FALSE ACCUSATIONS, *supra* note 7, at 102.

[83] *See* State v. Storch, 612 N.E.2d 305, 309 (Ohio 1993).

[84] Some courts equate a young child's "inability to lie" with an automatic assurance that the child's hearsay is trustworthy. *See, e.g.*, Broderick v. King's Way Assembly of God Church, 808 P.2d 1211, 1219 (Alaska 1991).

[85] *Id.* (quoting Phillip Kaufman, an expert at the trial).

effectively.[86] Since children do not have an independent concept of a lie,[87] a child may purposely lie to please the adult. For a child, "'truth' can be what pleases the adult."[88]

The adult who recounts a child's allegations of abuse may deliberately fabricate or expand upon the child's allegation to serve his own desires or confirm his own suspicions. In this situation, the child is used in a "power game[]" between adults.[89] This may explain why one study indicates that the incidence of false allegations during divorce or custody litigation might be as high as twenty percent, while the rate of false accusations outside those contexts is more likely between five and seven percent.[90] The accused is somewhat protected from harm from these fabrications, however, since the adult witness is subject to cross-examination at the trial.

Requiring children to testify and to be cross-examined at the criminal proceeding, whether in person or by means of a videotaped deposition or closed circuit television, would reduce the above mentioned hearsay risks. Knowing why the child made the statement and the circumstances surrounding the making of the statement would greatly assist the trier of fact.[91] Through cross-examination, the jury or judge would learn what the child saw and felt, to whom the child talked about the incident, and what

---

[86] *See* Daniel Goleman, *Lies Can Point to Mental Disorders or Signal Normal Growth*, N.Y. TIMES, May 17, 1988, at C1, C6.

> "In one study we just completed with 3-year-olds, we set up an attractive toy behind the child's back and tell him not to look at it while we leave the room . . . . About 10 percent don't peek while we're gone. Of the rest, a third will admit they peeked, a third will lie and say that they did not peek, and a third will refuse to say . . . . Those who say they did not look—who lie—looked the most relaxed. They've learned to lie well. There seems to be a certain relief in knowing how to lie effectively."

Id. (quoting Michael Lewis, a psychologist at Rutgers Medical School and one of the researchers).

[87] One study indicates that for children ages two to five, a "lie" is a "naughty word," and for ages five to eight and one-half, a "lie" is simply "something that isn't true." These definitions change over time, because for a child, "a lie is what an adult says it is." *See* GARDNER, TRUE AND FALSE ACCUSATIONS, *supra* note 7, at 271 (citations omitted).

[88] State v. Storch, 612 N.E.2d 305, 309 (Ohio 1993).

[89] *Id.*

[90] *See* Jones & Seig, *supra* note 9, at 22–29. It is important to note that sexual abuse allegations are not common in divorce and custody disputes. According to one study, data from seven jurisdictions revealed that only 105 out of 6100 custody or visitation cases involved allegations of sexual abuse (less than two percent). *See* WHITCOMB, *supra* note 3, at 7.

[91] *See* State v. Storch, 612 N.E.2d 305, 313 (Ohio 1993) (arguing that cross-examination of child-victims would provide valuable information for the trier of fact).

the child related to the adult. The fact finder needs this information to accurately evaluate the testimony.

This argument should not be taken to extremes, however. Many children tell the truth when they relate stories of sexual molestation, and many adults honestly attempt to inform the court of the abuse. Accordingly, all allegations of child sexual assault should be investigated and acted upon. On the other hand, the problems of false and exaggerated allegations are real and need to be addressed.[92] The purpose of the above discussion, then, is only to caution that child hearsay, especially hearsay not subject to cross-examination, may not be reliable.

A number of researchers have suggested that some children are traumatized by the courtroom experience.[93] In most of these studies, however, the courts did not use any of the innovative techniques, such as closed circuit television or videotaped depositions, that are designed to reduce the problems associated with testifying.[94] Thus, while testifying is clearly distressful for some children, as well as for some adults,[95] judicial discretion should be able to protect child witnesses from undue stress.[96] Additionally, many clinicians believe that confronting the accused in court is beneficial for the child because it shows that people take his or her problems seriously.[97] Therefore, the hearsay rules need to be designed so that prosecutors and families are encouraged to have children testify against their abusers, especially in situations where the reliability of the hearsay is particularly suspect.

---

[92] *See supra* notes 6–13.

[93] "[O]nly a rare child could fail to be traumatized by the experience of testifying in court." Note, *The Testimony of Child Victims in Sex Abuse Prosecutions: Two Legislative Innovations*, 98 HARV. L. REV. 806, 818 (1985). *See also* Myers, *supra* note 14, at 1724–25 n.41.

[94] GAIL GOODMAN ET AL., TESTIFYING IN CRIMINAL COURT: EMOTIONAL EFFECTS ON CHILD SEXUAL ASSAULT VICTIMS 4–8, 126 (1992).

[95] *See* South Carolina Dep't of Social Serv. v. John Doe, 355 S.E.2d 543, 548 (S.C. Ct. App. 1987).

[96] States' rules regarding child testimony, *see supra* notes 64–68 and *infra* notes 106–110 and accompanying text, are "designed to facilitate prosecution of child sexual abuse cases and to reduce the trauma or emotional distress experienced by victims of child sexual abuse as a result of the litigation process itself." Michael H. Graham, *Indicia of Reliability and Face to Face Confrontation: Emerging Issues in Child Sexual Abuse Prosecutions*, 40 U. MIAMI L. REV. 19, 21 (1985).

[97] *See, e.g.*, Diane H. Schetky & Elissa P. Benedek, *The Sexual Abuse Victim in the Courts*, 12 PSYCHIATRIC CLINICS N. AM. 471, 479 (1989); Myers, *supra* note 14, at 1724 n.41.

II. Hearsay Exceptions and the Systems Under Which States Admit Child Hearsay in Sexual Abuse Cases

A. *The Need for Exceptions to the Rule Against Hearsay*

An absolute prohibition against hearsay would result in the exclusion of many trustworthy, probative statements.[98] Far from achieving justice, such a rule would hinder the state's ability to protect its citizens from sexual abuse and other criminal conduct. Accordingly, the common law has evolved to allow certain out-of-court statements to be admissible when the hearsay was made under specific conditions that insure a degree of reliability similar to that provided by the safeguards of cross-examination.[99] These traditional, or "firmly rooted," exceptions include, for example, the excited utterance (or spontaneous declaration), the medical diagnosis or treatment, and the dying declaration[100] exceptions.[101] In 1975, Congress enacted the Federal Rules of Evidence, which codified many of these exceptions.[102] As of 1992, over thirty states had enacted evidence codes based substantially on the Federal Rules of Evidence, and many more were in the process of drafting such codes.[103]

---

[98] "Common sense tells us that much evidence which is not given under the three conditions [oath, cross-examination, and presence of fact finder] may be inherently superior to much that is. Moreover, when the choice is between evidence which is less than best and no evidence at all, only clear folly would dictate an across-the-board policy of doing without it." Fed. R. Evid. art. VIII advisory committee's note.

[99] "The exceptions to the general rule against hearsay testimony have been developed as 'substitutes' for the safeguards provided by cross-examination." State v. J.C.E., 767 P.2d 309, 312 (Mont. 1988).

[100] The common law medical diagnosis exception dates back to at least the late 19th century. *See* Idaho v. Wright, 497 U.S. 805, 820 (1990) (citing Mattox v. United States, 156 U.S. 237, 244 (1895) ("[T]he sense of impending death is presumed to remove all temptation to falsehood, and to enforce as strict an adherence to the truth as would the obligation of oath . . . .") and Queen v. Osman, 15 Cox. Crim. Cas. 1, 3 (Eng. N. Wales Cir. 1881) (Lush, L.J.) ("[N]o person, who is immediately going into the presence of his Maker, will do so with a lie upon his lips . . . .")).

The traditional common law exception required that the statement have been made by the victim and could only be used in the prosecution for criminal homicide. The current exception for statements made under belief of impending death is more expansive than the common law exception since it also allows the hearsay to be admissible in civil proceedings and any criminal proceeding. *See* Fed. R. Evid. 804(b)(2) advisory committee's note.

[101] *See infra* part II.B for a discussion of the spontaneous declaration and medical diagnosis exceptions.

[102] *See* Fed. R. Evid. 803(24) advisory committee's note.

[103] *See* Eric D. Green & Charles R. Nesson, Federal Rules of Evidence: With Selected Legislative History and New Cases and Problems xii (2d ed. 1992).

The traditional hearsay rules and exceptions were designed for adult, not child, testimony. By treating these two sources of testimony identically, valuable evidence may be excluded by hearsay structures designed to counter problems that may not be relevant to child witnesses. For example, one reason hearsay is disliked is because the witness does not testify in front of the fact finder.[104] As *Maryland v. Craig* noted, however, a child witness in sexual abuse prosecutions may not be required to testify in the presence of the defendant anyway, if so testifying would jeopardize their psychological well-being.[105] In these situations, the argument for a hearsay exclusion is moot.

Testifying in court is a much different experience for children than it is for adults. Many courts allow the attorney questioning the child to lead the witness during direct examination.[106] Other states allow a child to testify without swearing an oath.[107] Some courts have even been willing to allow testimony from a child who does not understand the obligation to testify truthfully.[108] At least eight states permit support persons to accompany the child during questioning.[109] Additionally, courts are permitted to close criminal trials to the press and public if the courts find that the state's interest in protecting "minor victims of certain sex crimes from public degradation, humiliation, demoralization, and psychological damage" necessitates such action.[110] Since a child

---

[104] *See* State v. J.C.E., 767 P.2d 309, 312 (Mont. 1988).

[105] *See supra* notes 64–65 and accompanying text.

[106] *See* Theresa Cusick, *Televised Justice: Toward a New Definition of Confrontation Under* Maryland v. Craig, 40 CATH. U. L. REV. 967, 968 (1991); Myers, *supra* note 14, at 1714. *See, e.g.*, State v. Hale, 326 S.E.2d 418 (S.C. Ct. App. 1985); United States v. Iron Shell, 633 F.2d 77, 92 (8th Cir. 1980), *cert. denied*, 450 U.S. 1001 (1981).

Some state statutes actually authorize the asking of leading questions of child sexual assault victims. *See, e.g.*, ALA. CODE § 15-25-3(c) (Supp. 1994).

[107] *See, e.g.*, FLA. STAT. ANN. § 90.605(2) (West Supp. 1994).

[108] *See, e.g.*, People v. District Court, 791 P.2d 682, 685 (Colo. 1990) (en banc); Galindo v. United States, 630 A.2d 202, 207 (D.C. 1993).

[109] ARK. CODE ANN. § 16-42-102 (Michie 1987); CAL. PENAL CODE § 868.5 (West Supp. 1994); HAW. REV. STAT. § 621-28 (1988); IDAHO CODE § 19-3023 (Supp. 1994); IDAHO R. EVID. 615(c) (Michie 1993); MICH. COMP. LAWS ANN. § 24.275a (West 1994); MINN. STAT. ANN. § 631.046 (West Supp. 1994); R.I. GEN. LAWS § 12-28-9(2) (Supp. 1993); WASH. REV. CODE ANN. § 7.69A.030(2), (7), & (8) (West 1992).

[110] Globe Newspaper Co. v. Superior Court, 457 U.S. 596, 607 n.18 (1982) (quoting Globe Newspaper Co. v. Superior Court, 423 N.E.2d 773, 779 (Mass. 1981)).

At least 14 states have legislation permitting the courtroom to be closed during the child sexual assault victim's testimony. CAL. PENAL CODE § 868.7 (West 1985); FLA. STAT. ANN. § 918.16 (West Supp. 1994) (media can remain); GA. CODE ANN. § 17-8-54 (Supp. 1994) (media can remain); ILL. ANN. STAT. ch. 38, para. 5/115-11 (Smith-Hurd 1993) (media can remain); LA. REV. STAT. ANN. § 15:469.1 (West 1992); MASS. GEN. LAWS ANN. ch. 278, § 16A (West Supp. 1994), ch. 278, § 16C (West Supp. 1994); MICH. COMP. LAWS ANN. § 600.2163a(1), (2), (9), & (10)(a) (West Supp. 1994)

witness is treated much differently than an adult witness, the hearsay rules need to be modified for child declarants to reflect the different conditions under which they testify.

B. *The Traditional or Firmly Rooted Hearsay Exceptions*

The two most commonly used exceptions in child abuse prosecutions are the excited utterance and the medical diagnosis or treatment exceptions.[111] These exceptions, however, are quite narrow. Accordingly, many states have chosen to make it easier for prosecutors to admit children's out-of-court statements by applying the traditional hearsay exceptions in a more lax manner when the case involves child sexual abuse.[112]

The expansion of the hearsay exceptions for child declarants, however, creates special problems.[113] For instance, a statement is admissible under the excited utterance exception so long as it was spontaneous and made under circumstances of shock or

---

(applicable only to preliminary examination proceedings); MINN. STAT. ANN. § 631.045 (West Supp. 1994); N.H. REV. STAT. ANN. § 632-A:8 (1986); N.C. GEN. STAT. § 15-166 (1983); S.C. CODE ANN. § 16-3-1530(G) (Law. Co-op. 1985); S.D. CODIFIED LAWS ANN. § 23A-24-6 (1988); VA. CODE ANN. § 19.2-266 (Michie Supp. 1994); WIS. STAT. ANN. § 970.03(4)(a) (West Supp. 1993) (applicable only to preliminary examination proceedings).

[111] Some states also have a "report of rape" or "fresh complaint" rule. These states allow the child's prompt complaint of rape to be admissible, either to corroborate the victim's trial testimony, *see, e.g.*, People v. Manuel M., 278 Cal. Rptr. 853 (Cal. Ct. App. 1991); Wilson v. State, 397 S.E.2d 59 (Ga. Ct. App. 1990); Commonwealth v. Lavalley, 574 N.E.2d 1000 (Mass. 1991); State v. Livingston, C.C.A. No. 01-C-01-9012-Cr-00337, 1991 Tenn. Crim. App. LEXIS 770 (1991), or to prove that the alleged victim complained of the crime. *See, e.g.*, State v. Pollitt, 530 A.2d 155, 163 (Conn. 1987) (en banc); Galindo v. United States, 630 A.2d 202, 209 (D.C. 1993); State v. Bethune, 578 A.2d 364 (N.J. 1990); Commonwealth v. Stohr, 522 A.2d 589 (Pa. Super. Ct. 1987).

The "report of rape" rule, however, is not an exception to the rule against hearsay, since the statements are not being admitted to prove the truth of the matter asserted. This testimony "comes in only to corroborate" other testimony. NEIL P. COHEN ET AL., TENNESSEE LAW OF EVIDENCE § 803(2).6 (2d ed. 1990).

[112] *See, e.g.*, State v. Myatt, 697 P.2d 836, 842 (Kan. 1985) ("Courts have thus tended to stretch existing hearsay exceptions to accommodate a child victim's out-of-court statements because they are deemed uniquely necessary and trustworthy . . . .").

[113] *See, e.g.*, Sharp v. Commonwealth, 849 S.W.2d 542, 546 (Ky. 1993).

> There may be a temptation among judges to let pity for small children who may have been victimized by vicious adults overcome their duty to enforce the rules of evidence . . . . "The rules of evidence have evolved carefully and painstakingly over hundreds of years as the best system for arriving at the truth . . . . Their purpose would be subverted if courts were permitted to disregard them at will . . . ."

Id. (citation omitted). *See also* State v. Myatt, 697 P.2d at 842 (Kan. 1985) ("The problem with 'stretching' the existing exceptions in this manner is the destruction of the certainty and integrity of the exceptions . . . .").

nervous excitement.[114] The rationale for the exception is that excitement "temporarily stills the capacity of reflection and produces utterances free of conscious fabrication."[115] The two basic requirements for admission of the statement are that the declarant was so startled by an event that all normal thought processes stopped functioning and that the statement was made as a spontaneous result of that event.[116]

Children, however, rarely make statements that would normally qualify as excited utterances because fear, loyalty, or a lack of comprehension cause children to delay reporting abuse, especially if the abuse was perpetrated by a parent or close relative.[117] Therefore, in sexual abuse cases, many courts allow a child's statement to qualify as an excited utterance even though considerable time has elapsed between the actual abuse and when the statement was made.[118] Some courts have even admitted spon-

---

[114] *See* FED. R. EVID. 803(2) advisory committee's note. *See supra* note 22 for text of rule.

See also White v. Illinois, 502 U.S. 346, ____, 112 S. Ct. 736, 743 n.8 (1992) ("There can be no doubt that the two exceptions [excited utterance and medical diagnosis or treatment] that we consider in this case are 'firmly rooted.'"). The excited utterance exception is at least two hundred years old, and may even date as far back as the late 1600s. *See id.* (citing 6 JOHN HENRY WIGMORE, EVIDENCE IN TRIALS AT COMMON LAW § 1747 (James H. Chadbourn rev., 1976) and Thompson v. Trevanion, 90 Eng. Rep. 179 (K. B. 1694)). Nearly four-fifths of the states recognize this exception. *Id.*

[115] *See* FED. R. EVID. 803(2) advisory committee's note.

[116] *See* CHARLES T. MCCORMICK, MCCORMICK ON EVIDENCE § 297 (Edward W. Cleary ed., 3d ed. 1984).

The terms "excited utterance" and "spontaneous declaration" are interchangeable. Some older cases also refer to the exception under the phrase of "res gestae," or "things done." *See* Cassidy v. State, 536 A.2d 666, 670–73 (Md. Ct. Spec. App. 1988) (explaining the term "res gestae").

[117] *See* State v. J.C.E., 767 P.2d 309, 314 (Mont. 1988). *See also* DAVID FINKLEHOR, SEXUALLY VICTIMIZED CHILDREN 65 (1979) (reporting that only 26% of women who were sexually molested as children felt "shock" and only 20% felt "surprised").

[118] *See, e.g.*, Morgan v. Foretich, 846 F.2d 941, 947 (4th Cir. 1988) (finding that a four-year-old victim's statements to her mother about molestation that occurred the previous night should have been admitted under the excited utterance exception to hearsay); State v. Parker, 730 P.2d 921, 924 (Idaho 1986) (14-year-old child's taped statement admitted as evidence even though victim waited silently for family members to arrive before relating events); People v. Nevitt, 553 N.E.2d 368 (Ill. 1990) (holding that a statement made in response to mother's questioning and five hours after the alleged incident was spontaneous); Smith v. State, 252 A.2d 277 (Md. Ct. Spec. App. 1969) (admitting statements made four and one-half to five hours after rape); State v. Plant, 461 N.W.2d 253, 264 (Neb. 1990) (upholding admission of "excited utterance" made by four-year-old child two days after alleged incident and after police questioning); State v. Harris, 717 P.2d 242 (Or. Ct. App. 1986) (allowing unexcited hearsay declarations of sexual misconduct if victim incompetent to testify because of age); State v. Sorenson, 421 N.W.2d 77, 84 (Wis. 1988) ("Though not made in immediate temporal relation to incidents which are the focal point of their statements, this court has held statements made by young children concerning sexual assault to be sufficiently contemporaneous and spontaneous to be admissible . . . ."). One court has justified the

taneous utterances made a considerable time after the event when the declarant was suddenly re-exposed to a related excitement,[119] or when the declarant named the defendant immediately upon awaking in the middle of the night.[120] These expansions of the exception are problematic, since spontaneous declarations are usually only deemed trustworthy because they were made immediately after the startling event.[121] Since the child has had time for reflective thought, there is no special reason to believe that these nonspontaneous statements are trustworthy.

A statement is admissible under the medical diagnosis or treatment exception only to the extent that the statement describes medical history, symptoms, or pains, and is "reasonably pertinent to diagnosis or treatment."[122] The underlying rationale of the exception is that doctors will seek, and patients will give, truthful information because both the doctor[123] and the patient want the patient to get better.[124] Under this rationale, any statement that is reasonably pertinent to diagnosis or treatment should

---

expansion of this exception by arguing that children "react to and relate traumatic events somewhat differently than adults." *In re* Troy, 842 P.2d 742, 747 (N.M. Ct. App. 1992) (citations omitted).

But see, e.g., Matthews v. State, 515 N.E.2d 1105, 1107 (Ind. 1987) (refusing to admit a statement that occurred over 30 hours after the alleged incident); State v. Russell, 872 S.W.2d 866, 879 (Mo. Ct. App. 1994) (holding that statements made as a product of the excitement of being examined by a doctor are not excited utterances); People v. Allen, 568 N.Y.S.2d 132, 134 (N.Y. App. Div. 1991) (refusing to apply exception to statements made morning after alleged attack despite state's argument that child's "unusually quiet and still" behavior in morning was evidence the child was still under stress of situation).

[119] *See, e.g.*, United States v. Scarpa, 913 F.2d 993, 1016–17 (2d Cir. 1990); *In re* Troy, 842 P.2d 742 (N.M. Ct. App. 1992) (admitting testimony that child became hysterical and admitted abuse when the mother told the child that she was driving her to her father's home).

[120] *See, e.g.*, George v. State, 813 S.W.2d 792, 795–96 (Ark. 1991), *modified*, 818 S.W.2d 951 (Ark. 1991); State v. Boston, 545 N.E.2d 1220, 1231 (Ohio 1989).

[121] *See supra* note 115 and accompanying text.

[122] FED. R. EVID. 803(4). *See supra* note 23 for full text of rule; FED. R. EVID. 803(4) advisory committee's note. *See also supra* note 114. Nearly four-fifths of the states recognize this exception. *See* White v. Illinois, 502 U.S. 346, ____, 112 S. Ct. 736, 743 n.8 (1992).

[123] *See* JOHN W. STRONG ET AL., 2 MCCORMICK ON EVIDENCE 250 (4th ed. 1992) ("The general reliance upon 'subjective' facts by the medical profession and the ability of its members to evaluate the accuracy of statements made to them is considered sufficient protection against contrived symptoms.").

[124] *See* FED. R. EVID. 803(4) advisory committee's note. *See also* United States v. Iron Shell, 633 F.2d 77, 83–84 (8th Cir. 1980), *cert. denied*, 450 U.S. 1001 (1981); State v. Robinson, 735 P.2d 801, 809 (Ariz. 1987); MCCORMICK, *supra* note 116, at § 292, at 839 (The patient knows that "the effectiveness of the treatment he receives may depend largely upon the accuracy of the information he provides . . . ."). The motivational aspect of Rule 803(4) is often called the patient's "selfish interest" rationale. *See* State v. Dever, 596 N.E.2d 436, 444 (Ohio 1992), *cert. denied*, 113 S. Ct. 1279 (1993).

be reliable.[125] Accordingly, statements made in preparation of having the doctor testify at trial[126] and statements relating to the identity of the alleged perpetrator[127] are traditionally not admissible within the medical diagnosis or treatment exception because they are not related to diagnosis or treatment.

However, several courts have held that when a doctor examines a sexually abused child to determine what happened to the child and what therapy or treatment is needed, the child's statements are admissible regardless of the circumstances surrounding the examination.[128] This expansion is problematic. First, there is no reason to believe that statements made by the child in order to prepare the doctor to testify are any more trustworthy than other out-of-court statements made by the child. Second, the child's relationship with the testifying doctor is not the typical doctor-patient relationship;[129] it may often border on a police detective-witness interaction.[130] Finally, the doctor does not see the child for treatment unless there has been a claim that the

---

[125] *Iron Shell*, 633 F.2d at 84, created a two-part test to decide whether the declarant's statements were reasonably pertinent to medical diagnosis or treatment: (1) was the declarant's motive consistent with receiving medical care, and (2) was it reasonable for the physician to rely on the information in diagnosis or treatment.

[126] FED. R. EVID. 803(4) advisory committee's note ("Conventional doctrine has excluded from the hearsay exception, as not within its guarantee of trustfulness, statements to a physician consulted only for the purpose of enabling him to testify."). *See also* Souder v. Commonwealth, 719 S.W.2d 730, 735 (Ky. 1986) (excluding "information provided as part of a criminal investigation . . . .").

[127] *See, e.g.*, *Iron Shell*, 633 F.2d 77; United States v. Nick, 604 F.2d 1199, 1201–02 (9th Cir. 1979); State v. Robinson, 735 P.2d 801, 810 (Ariz. 1987); State v. Dollinger, 568 A.2d 1058, 1060 (Conn. App. Ct. 1990); State v. Livingston, No. 01-C-01-9012-CR-00337, 1991 Tenn. Crim. App. LEXIS 770, at *29 (Tenn. Crim. App. Sept. 27, 1991). *See generally*, FED. R. EVID. 803(4) advisory committee's note ("Statements as to fault would not ordinarily qualify under this latter language. Thus a patient's statement that he was struck by an automobile would qualify but not his statement that the car was driven through a red light . . . .").

[128] *See e.g.*, Edwards v. Commonwealth, 833 S.W.2d 842, 845 (Ky. 1992) (allowing hearsay told to a doctor to whom the child was referred by the police under the medical diagnosis or treatment exception); State v. Aguallo, 350 S.E.2d 76 (N.C. 1986) (admitting statements made to a doctor several months after the event strictly in anticipation of prosecution, which included identification of the defendant, without any indication that the child sought medical treatment or was even aware that her identification was necessary to aid her treatment).

But see Sharp v. Commonwealth, 849 S.W.2d 542, 544 (Ky. 1993) (noting that "statements made to a physician who lacks treatment responsibility have less inherent reliability than traditional patient history . . ." and holding that the trial court must therefore decide from the totality of the circumstances whether the probative value of the evidence outweighs its prejudicial effect); State v. Zimmerman, 829 P.2d 861 (Idaho 1992); People v. LaLone, 437 N.W.2d 611 (Mich. 1989) (holding that statements are not admissible if the child visited a psychologist strictly for psychological, not medical, treatment).

[129] *See Sharp*, 849 S.W.2d at 544.

[130] *See* State v. Storch, 612 N.E.2d 305, 313 (Ohio 1993).

child has been abused. The doctor, therefore, is predisposed to confirm what she has already been told by other adults.[131]

Moreover, many courts apply an excessively loose interpretation of "reasonably pertinent" and "medical diagnosis or treatment" when the case involves child sexual abuse. For example, the Court of Appeals for the Eighth Circuit held in *United States v. Renville*[132] that the identification of the alleged attacker is admissible if the sexual abuser is a member of the child's immediate family. The court supported the admission of such hearsay by arguing that the emotional and psychological injuries of sexual abuse are magnified when the attacker lives in the child's home and that physicians have an obligation under state law to prevent abused children from being returned to an environment where the abuse may recur.[133]

A majority of the courts that have discussed *Renville* have approved of its reasoning.[134] Indeed, a few courts have expanded

---

[131] *See* State v. Harris, 808 P.2d 453, 459 (Mont. 1991).

[132] 779 F.2d 430 (8th Cir. 1985).

[133] *See Renville*, 779 F.2d at 437–38.

[134] *See, e.g.*, Morgan v. Foretich, 846 F.2d 941 (4th Cir. 1988) (applying the same rules of evidence in a civil case); Conn v. Wells, No. 93-1313, 1994 U.S. App. LEXIS 1803 (6th Cir. Feb. 1, 1994) (identification admissible despite two false identifications made by the child to the doctor); United States v. George, 960 F.2d 97, 99 (9th Cir. 1992); State v. Robinson, 735 P.2d 801 (Ariz. 1987); Stallnacker v. State, 715 S.W.2d 883 (Ark. 1986); State v. Depastino, No. 14695, 1994 Conn. LEXIS 49 (Conn. Feb. 22, 1994); State v. Tracy, 482 N.W.2d 675 (Iowa 1992); Edwards v. Commonwealth, 833 S.W.2d 842 (Ky. 1992); People v. Meeboer, 484 N.W.2d 621, 629 (Mich. 1992) (MICH. R. EVID. 803(4) refers to medical diagnosis *and* treatment, not diagnosis *or* treatment) (emphasis added); Jones v. State, 606 So. 2d 1051 (Miss. 1992); State v. Thompson, No. 93-134, 1993 Mont. LEXIS 407 (Mont. Dec. 22, 1993); State v. Altgilbers, 786 P.2d 680 (N.M. App. Ct. 1989); State v. Aguallo, 350 S.E.2d 76 (N.C. 1986); State v. Dever, 596 N.E.2d 436 (Ohio 1992), *cert. denied*, 113 S. Ct. 1279 (1993); Kennedy v. State, 839 P.2d 667 (Okla. Crim. App. 1992); State v. Moen, 786 P.2d 111 (Or. 1990); State v. Garza, 337 N.W.2d 823 (S.D. 1983); State v. Livingston, No. 01-C-01-9012-CR-00337, 1991 Tenn. Crim. App. LEXIS 770 (Tenn. Crim. App. Sept. 17, 1991); State v. Ashcraft, 859 P.2d 60 (Wash. 1993); State v. Edward Charles L., 398 S.E.2d 123 (W. Va. 1990).

But see United States v. Cherry, 938 F.2d 748, 756 n.14 (7th Cir. 1991); Sluka v. State, 717 P.2d 394, 399 (Alaska App. Ct. 1986); State v. Jones, No. 80,069, 1993 Fla. LEXIS 1344 (Fla. Aug. 26, 1993) (rejecting the adoption of *Renville* into Florida law and instead applying Florida's statutory tender years exception, § 90.803(23)); Wilson v. State, 390 S.E.2d 903 (Ga. App. Ct. 1990); Vest v. State, 621 N.E.2d 1094 (Ind. 1993) (holding that an uncle who was alleged to have intentionally burned a cigarette on his three-year-old niece's foot was not sexual assault); People v. Hall, 601 N.E.2d 883, 896 (Ill. App. Ct. 1992) (refusing to extend ILL. REV. STAT. ch. 38, para. 113–15 (1987), which is similar to FED. R. EVID. 803(4), to include identification "merely because one of the options of treatment might have included removal of [the child] from defendant's home . . . ."), *cert. denied*, 113 S. Ct. 2341 (1993); State v. True, 438 A.2d 460 (Me. 1981); Cassidy v. State 536 A.2d 666, 684-85 (Md. Ct. Spec. App. 1988), *cert. denied*, 541 A.2d 965 (Md. 1988); State v. Russell, No. 18699, 1994 Mo. App. LEXIS 492, at *10 (Mo. Ct. App. Mar. 22, 1994) (ruling that in order to treat or

the exception to allow the identification of the alleged perpetrator regardless of whether he lives in the same household as the child[135] and regardless of whether the crime involves sexual conduct.[136]

Several commentators have questioned whether the loosening of the medical diagnosis or treatment exception is justified.[137] Even if *Renville* is correct that the identification of the perpetrator is relevant to successful treatment, there is little reason to believe that a child would comprehend that connection.[138] Unless the child understands that telling the truth is essential to his or her health, there is no special reason to believe this type of hearsay.[139]

---

diagnose the child, the doctor did not need to know who put a finger inside the child's vagina); State v. Veluzat, 578 A.2d 93, 96 (R.I. 1990); State v. Gorkey, 574 A.2d 766, 772 (Vt. 1990).

[135] *See, e.g.*, Edwards v. Commonwealth, 833 S.W.2d 842 (Ky. 1992) (ruling that by the time the doctor examined the boy he was no longer at risk of future abuse since he did not live with the defendant, nor was he even allowed unsupervised visitation with the defendant). *But see* Jones v. State, 606 So.2d 1051 (Miss. 1992) (refusing to apply the medical diagnosis or treatment exception when defendant did not live in household, even though defendant was the child's natural father).

[136] *See, e.g.*, State v. Thompson, No. 93-134, 1993 Mont. LEXIS 407, at *20 (Mont. Dec. 22, 1993) (ruling that a 10-year-old's statement to her doctor identifying her stepfather as the one who kicked her was reasonably pertinent to diagnosis or treatment, because "[i]t is important to know the specific mechanism of injury to determine, for example, force and to determine further evaluation. In this case [the doctor] thought it was necessary to know the mechanism of injury to find out if [x-rays were necessary] . . . .").

[137] *See, e.g.*, MICHAEL H. GRAHAM, HANDBOOK OF FEDERAL EVIDENCE § 803.4 at 828, n.4 (2d ed. 1986) ("Concern over the recent revelations of child sex abuse have [sic] caused several state courts to expand, if not distort, the concept of diagnosis or treatment . . . ."); Robert P. Mosteller, *Child Sexual Abuse and Statements for the Purpose of Medical Diagnosis or Treatment*, 67 N.C. L. REV. 257, 258 (1989) (Applications of the medical diagnosis or treatment exception in child abuse cases "have tended to expose the thinness of the justification for extending the exception to statements made without any view toward treatment . . . .").

[138] Some courts argue that the connection is irrelevant if the statement is supported by corroborating evidence or other indicia of reliability. *See* State v. Ashcraft, 859 P.2d 60 (Wash. 1993) (arguing that the unlikelihood of fabrication due to the victim's youthfulness can be an indicia of reliability; the alleged perpetrator's implausible story regarding the abuse can be considered corroborating evidence). *See also* State v. Robinson, 735 P.2d 801, 809 (Ariz. 1987) ("Because of their young age, sexually abused children may not always grasp the relation between their statements and receiving effective medical treatment. It is particularly important, therefore, to ask whether the information sought by the treating doctor was reasonably pertinent to effective treatment . . . ."); State v. Dollinger, 568 A.2d 1058, 1061 (Conn. App. Ct. 1990) ("Even if V was too young to formulate consciously the motive of advancing her own health by being truthful with the doctor, the facts and circumstances here are consistent with that purpose.").

[139] *See, e.g.*, People v. Meeboer, 484 N.W.2d 621, 636 (Mich. 1992) (Brickley, J. dissenting) ("I would contend that children do not necessarily possess the required cognitive ability to make the connection between giving truthful information and

Additionally, the physical well-being interest associated with medical diagnosis or treatment, which encourages adults to tell the truth, may have the opposite effect with children who do not "want" diagnosis or treatment. While an adult associates telling a doctor that she is sick with getting better, a child may associate telling a doctor that he is sick only with having to receive painful treatment, such as a shot. Accordingly, children have a greater incentive to lie to a doctor to avoid medical treatment or diagnosis. Further, the stressful environment of speaking to a doctor may encourage the child to desire the reassuring presence of a relative, even a relative who abused him.[140] Thus, a medical examination may motivate the child to lie about the identity of the perpetrator in order to protect that relative. As a result, the child's statement to the doctor, especially the statement identifying the perpetrator, is not necessarily reliable. This suggests that the expansions of the traditional hearsay exceptions are not as trustworthy as the original exceptions.

C. *Residual Hearsay Exceptions*

Finding that the traditional exceptions to the hearsay rule were not always sufficient, many states have looked for additional ways to admit hearsay by child declarants in sexual abuse prosecutions. States have turned to two vehicles: residual exceptions and statutory tender years hearsay exceptions.

The Federal Rules of Evidence and many state rules of evidence include residual or catch-all exceptions that are to be applied in "new and presently unanticipated situations which demonstrate a trustworthiness within the spirit of the specifically stated exceptions."[141] Federal Rule 804(b)(5) is to be used when the declarant is unavailable, while Federal Rule 803(24) is to be

---

receiving proper medical care required for a self-interest motivation component to exist."). *But see* State v. Dever, 596 N.E.2d 436, 444 (Ohio 1992)

> Once the child is at the doctor's office, the probability of understanding the significance of the visit is heightened and the motivation for diagnosis and treatment will normally be present . . . . Everyday experience tells us most children know that if they do not tell the truth to the person treating them, they may get worse and not better . . . .

[140] *See* State v. Jones, 625 So.2d 821, 825 (Fla. 1993) ("Truthful answers as to the identity of its abuser may well wrench a child from the reassuring presence of its mother or father or both. It is highly unlikely that there operates in an infant declarant a compelling desire to bring about such a result.").

[141] FED. R. EVID. 803(24), 804(b)(5) advisory committee's note.

used regardless of whether the declarant is available.[142] Although Congress intended that the residual exceptions be used sparingly,[143] many courts regularly use the residual exceptions to admit hearsay in child sexual abuse cases.[144]

The residual exceptions provide five requirements for admissibility of a statement: (1) it must have circumstantial guarantees of trustworthiness equivalent to those possessed by the specific hearsay exceptions;[145] (2) it must concern a material fact; (3) it must be more probative than any other evidence that the proponent can reasonably procure; (4) admission of the statement must best serve the general purposes of the evidence rules and the interest of justice; and (5) the opponent must have appropriate notice that the statement will be offered.[146] The first requirement is the only one that concerns trustworthiness,[147] and the indeterminacy of the other criteria, such as best serve the interest of justice, provide no real guidance in deciding whether a statement is reliable.[148] While many courts have listed factors

---

[142] *See supra* note 24 for text of rules.

[143] *See* John Doe, et al. v. United States, 976 F.2d 1071, 1074 (7th. Cir. 1992) (civil suit).

[144] *See, e.g.*, United States v. St. John, 851 F.2d 1096, 1098 (8th Cir. 1988) (noting a "formidable line of Circuit precedent that sanctions the use of hearsay testimony in child sexual abuse cases . . . ."); United States v. Nick, 604 F.2d 1199 (9th Cir. 1979); State v. Robinson, 735 P.2d 801, 810–12 (Ariz. 1987); Oldsen v. People, 732 P.2d 1132, 1136–37 (Colo. 1986); State v. Dollinger, 568 A.2d 1058, 1062–64 (Conn. Ct. App. 1990).

But see People v. Bowers, 801 P.2d 511, 517 (Colo. 1990) ("[S]ection 13-25-129 [state's statutory hearsay exception] constitutes the exclusive basis for admitting a child-victim's hearsay statement of a sexual act committed against the child when such hearsay statement is not otherwise admissible under any other specific hearsay exception created by statute or court rule . . . .").

[145] The statement must be independently trustworthy to be admissible under the residual exception. Courts differ, however, on whether they should consider corroborative evidence at all in support of the statements' admissibility. Overall, most courts believe corroborative evidence is useful. *See* State v. Allen, 755 P.2d 1153, 1164 (Ariz. 1988). *See, e.g.*, United States v. Bailey, 581 F.2d 341 (3rd Cir. 1978); United States v. Nick, 604 F.2d 1199 (9th Cir. 1979); State v. Allen, 755 P.2d 1153, 1164 (Ariz. 1988) ("Corroborative evidence makes a statement more reliable because it increases the likelihood that the statement is true . . . ."); State v. Taylor, 704 P.2d 443 (N.M. App. Ct. 1985); State v. Fearing, 337 S.E.2d 551 (N.C. 1985); State v. Hollywood, 680 P.2d 655 (Or. App. Ct. 1984).

Other courts, however, have flatly rejected the use of corroborative evidence. *See, e.g.*, Huff v. White Motor Corp., 609 F.2d 286, 293 (7th Cir. 1979) ("The probability that the statement is true, as shown by corroborative evidence, is not, we think, a consideration relevant to its admissibility under the residual exception to the hearsay rule . . . ."); State v. Ryan, 691 P.2d 197, 204 (Wash. 1984).

[146] *See* FED. R. EVID. 803(24), 804(a)(5). *See supra* note 24 for text of rule.

[147] The second and third requirements deal with necessity, the fifth involves fairness, and the fourth seems too general to be an independent requirement.

[148] *See* Broderick v. King's Way Assembly of God Church, 808 P.2d 1211, 1219 (Alaska 1991) ("Because the twenty-two situations specified as exceptions under Rule 803 [of Alaska's Rules of Evidence] vary greatly in their circumstantial indicia of

in order to assist a trial judge in determining if the hearsay has "equivalent indicia of trustworthiness,"[149] such guidelines often give the trial courts so many factors to consider that in practice they may be the equivalent of granting the judge complete discretion. Such absolute flexibility minimizes the predictability and consistency of rulings and exacerbates the difficulty of preparing for trial. Since one of the primary purposes of the Federal Rules of Evidence and the states' rules of evidence is to reduce unpredictability and arbitrariness,[150] this outcome is highly unsatisfactory.

Excessive discretion is not the only problem with residual hearsay exceptions. An additional problem stems from the fact that, knowing that a court may deem the child's live testimony more probative than the hearsay, a prosecutor who would rather base her case on the child's hearsay than on the child's testimony[151] may discourage the child from testifying so as to avoid the possible exclusion of the hearsay. For these reasons, expanding the use of the residual exceptions is an undesirable method for increasing the admission of child hearsay in sexual abuse prosecutions.

D. *Statutory Tender Years Hearsay Exceptions*

Since the traditional hearsay exceptions were not designed, and hence do not work well, for child witnesses,[152] tender years hearsay exception statutes are the optimal means for governing

---

trustworthiness, this standard is necessarily inexact."). Some courts maintain that this is an advantage of using a residual exception for child sex abuse cases. *See, e.g.*, United States v. Cree, 778 F.2d 474, 478 (8th Cir. 1985) ("Rule 803(24) provides a trial court with some flexibility when it must make a determination as to the admissibility of hearsay evidence, and there is no specific rule governing admissibility . . . ."). Several courts have declined to apply an independent trustworthiness test and have instead held that "a determination that admissibility would not violate the confrontation clause is tantamount to a decision that the evidence is admissible under the 'catch-all' exception." *See, e.g.*, State v. Edwards, 485 N.W.2d 911 (Minn. 1992).

[149] *See, e.g.*, Gregg v. State, 411 S.E.2d 65, 68 (Ga. App. Ct. 1991) (listing 10 factors that "are to be applied neither in mechanical nor mathematical fashion, but in that manner best calculated to facilitate determination of the existence or absence of the requisite degree of trustworthiness."); State v. J.C.E., 767 P.2d 309, 315–16 (Mont. 1988) (listing 20 factors).

[150] *See* FED. R. EVID. art. VIII advisory committee's note.

[151] *See supra* notes 32–41 and accompanying text.

[152] *See* State v. Aguallo, 350 S.E.2d 76, 83 (N.C. 1986) (Billings, J. dissenting) ("[The legislature should] consider the appropriateness of special rules for obtaining evidence in child sexual abuse cases rather than [having the courts] try to fit this testimony into a mold which cannot contain it.").

child hearsay in child sexual abuse cases.[153] Accordingly, in the past twelve years, most states have enacted specific statutes under which child hearsay—not otherwise admissible under the state's other hearsay exceptions—would be admissible. In 1982, only two states, Washington and Kansas, had such statutes. That number grew to eighteen by 1985,[154] and to twenty-eight by 1989.[155] As of February 15, 1994, the legislatures of thirty-four states had passed tender years hearsay exceptions for criminal proceedings.[156] Of these, Arizona,[157] Idaho,[158] Kentucky,[159] and Mississippi[160] have declared their statutes defective on the ground

---

[153] Whether the exception should be legislatively enacted or judicially prescribed depends on the individual state's constitution and laws. *See infra* notes 157–160. For a general analysis of the legislature's authority to enact rules of evidence, see Hall v. State, 539 So.2d 1338, 1349–66 (Miss. 1989) (Hawkins, J., dissenting).

[154] *See* Cassidy v. State, 536 A.2d 666, 681 (Md. Ct. Spec. App. 1988).

[155] *See* WHITCOMB, *supra* note 3, at 166–67 (citing statutes).

[156] ALA. CODE §§ 15-25-31 to 15-25-37 (1993); ALASKA STAT. § 12.40.110 (1992) (pertains only before a grand jury); ARIZ. REV. STAT. ANN. § 13-1416 (1993); ARK. CODE. ANN. § 16-41-101, ARK. R. EVID. 803(25) (Michie 1992); CAL. EVID. CODE § 1228 (West 1993); COLO. REV. STAT. §§ 13-25-129, 18-3-411(3) (1993); DEL. CODE ANN. tit. 11, § 3513 (1992); FLA. STAT. ANN. § 90.803(23) (West 1992); GA. CODE ANN. § 24-3-16 (1993); HAW. REV. STAT. 804(b)(6) (1993); IDAHO CODE §§ 19-809A, 19-3024 (1994); ILL. ANN. STAT. ch. 725, § 5/115-10 (West 1993); IND. CODE ANN. § 35-37-4-6 (Burns 1992); KAN. STAT. ANN. § 60-460(dd) (1992); KY. REV. STAT. ANN. § 421.355 (Baldwin 1993) (repealed); ME. REV. STAT. ANN. tit. 15, § 1205 (West 1992); MD. CODE ANN. CTS. & JUD. PROC. § 9-103.1 (1993); MASS. ANN. LAWS, ch. 233, § 81 (1993); MICH. R. EVID. 803A (1993); MINN. STAT. ANN. § 595.02 (West 1993); MISS. CODE ANN. § 13-1-403 (1993); MO. REV. STAT. § 491.075 (1993); NEV. REV. STAT. ANN. § 51.385 (Michie 1991); N.J. R. EVID. 803(c)(27) (superseding N.J. STAT. ANN. § 2A:84A, N.J. R. EVID. 63(33) (West 1993) (effective July 1, 1993)); OHIO R. EVID. 807 (Banks-Baldwin 1993); OKLA. STAT. ANN. tit. 12, § 2803.1 (West 1993); OR. REV. STAT. § 40.460 (1991), OR. R. EVID. 803(18a); PA. STAT. ANN. § 5985.1 (1993); R.I. GEN. LAWS § 11-37-13.1 (1993) (pertains only in front of a grand jury); S.D. CODIFIED LAWS ANN. § 19-16-38 (1993); TEX. CRIM. PROC. CODE ANN. ¶ 38.072 (West 1993); UTAH CODE ANN. § 76-5-411 (1993); VT. R. EVID. 804a (narrowing scope of VT. R. EVID. 803(24) to pertain only to certain enumerated sexual crimes); WASH. REV. CODE ANN. § 9A.44.120 (West 1991).

[157] *See* State v. Robinson, 735 P.2d 801, 808 (Ariz. 1987) ("If the statute [ARIZ. REV. STAT. ANN. § 13-1416 (1994)] is less or more restrictive than the rules, the legislature is attempting to dictate what is or is not reliable evidence, and when cross-examination may be necessary to test reliability. These are judicial determinations, which, under [Arizona Constitution] art. 6, § 5(5), are the exclusive province of this court.").

[158] *See* State v. Zimmerman, 829 P.2d 861 (Idaho 1992) ("To the extent that this statute [IDAHO CODE § 19-3924 (1994)] attempts to prescribe the admissibility of hearsay evidence and is in conflict with the Idaho Rules of Evidence, it is of no force or effect.").

[159] *See* Drumm v. State, 783 S.W.2d 380, 382 (Ky. 1990) ("[T]he present statute [section 421.355] transgresses established procedure relating to the competency of children to testify as witnesses, usurps the power of the judiciary to control procedure, and violates Section 27 and 28 of the Constitution of Kentucky. These Sections establish the judiciary as 'a separate body of magistracy . . . .'").

[160] *See* Hall v. State, 539 So.2d 1338, 1339 (Miss. 1989) ("The legislature has enacted upon a matter [issuing rules of evidence] at the core of the judicial power. In such

that under their state constitutions and laws, the judicial branch, rather than the legislative branch, controls evidentiary rules. In addition, the Arkansas Supreme Court invalidated that State's tender years statute because it violated the federal Constitution's Confrontation Clause.[161] The other states that have addressed the issue, however, have found their tender years statutes permissible under both state and federal law.[162]

E. *Categorization of the Different Tender Years Statutes*

All of the statutes of the thirty-four states may be placed into six categories based on how they deal with the issues of unavailability and corroborative evidence. The first and largest category consists of sixteen states that have statutes similar to Washington's section 9A.44.120.[163] In these states, hearsay of a child

---

circumstances the statute [MISS. CODE ANN. § 13-1-403] should not be enforced. The integrity of the judicial department of the government of this state demands no less.").

[161] *See* Vann v. State, 831 S.W.2d 126, 128 (Ark. 1992) (Rule 803(25) provides that the hearsay of a child is admissible upon showing only that it possesses a "reasonable likelihood of trustworthiness." This standard is inadequate since the Confrontation Clause requires that statements bear "particularized guarantees of trustworthiness. . . ." (quoting Idaho v. Wright, 497 U.S. 805, 110 S.Ct. 3139, 3147 (1990))).

[162] *See, e.g.*, Fortner v. State, 582 So.2d 581 (Ala. Crim. App. 1990); Murray v. State, 770 P.2d 1131 (Alaska Ct. App. 1989); St. Clair v. State, 783 S.W.2d 835 (Ark. 1990) (separation of powers doctrine does not preclude General Assembly from enacting child hearsay rule); State v. Krick, 643 A.2d 331 (Del. Sup. Ct. 1993) (discussing DEL. CODE ANN. tit. 11, § 3513(b)(2)(a)(8) (1993), but applying the entire section to uphold admission of hearsay against defendant); Perez v. State, 536 So.2d 206, 209 (Fla. 1988), *cert. denied*, 492 U.S. 923 (1989); Rayburn v. State, 391 S.E.2d 780 (Ga. Ct. App.), *cert. denied*, 498 U.S. 969 (1990); Miller v. State, 517 N.E.2d 64 (Ind. 1987); State v. Myatt, 697 P.2d 836 (Kan. 1985) (holding KAN. STAT. ANN. § 60-460(dd) (1992) constitutional since the statute's requirement that the statements are "apparently reliable" implies that the judge must find that the evidence contains "particularized guarantees of trustworthiness."); Cole v. State, 574 A.2d 326 (Md. Ct. Spec. App. 1990); State v. Bellotti, 383 N.W.2d 308 (Minn. Ct. App. 1986); State v. Hester, 801 S.W.2d 695 (Mo. 1991); Lytle v. State, 816 P.2d 1082 (Nev. 1991); State v. Maben, 626 A.2d 63 (N.J. 1993) (applying N.J. R. EVID. 63(33), which has been superseded by Rule N.J. R. EVID. 803(c)(27)); State v. Storch, 612 N.E.2d 305 (Ohio 1993) (OHIO R. EVID. 807 accords with both U.S. and Ohio Constitutions); Spears v. State, 805 P.2d 681 (Okla. Ct. Crim. App. 1991); In re C.L., 397 N.W.2d 81, 84 (S.D. 1986); Buckley v. State, 786 S.W.2d 357 (Tex. Ct. Crim. App. 1990) (TEX. CRIM. PROC. CODE ANN. ¶ 38.072 (West 1993) not unconstitutional on its face); State v. Ramsey, 782 P.2d 480 (Utah 1989) (child hearsay statute not void for vagueness); State v. Gallagher, 554 A.2d 221 (Vt. 1988) (applies VT. R. EVID. 804a); State v. Swan, 790 P.2d 610 (Wash. 1990) (defendant's right to confrontation of witness not violated by child hearsay statute).

[163] WASH. REV. CODE ANN. § 9A.44.120 (West 1991). *See supra* note 25 for text of statute. The 16 states are Alabama, Arizona, Colorado, Florida, Idaho, Illinois, Maryland, Minnesota, Mississippi, New Jersey, Oklahoma, Oregon, Pennsylvania, South Dakota, Utah, and Washington. *See supra* note 156 for citations of the specific statutes. The statutes of Arizona, Idaho, and Mississippi have been declared invalid by their state courts. *See supra* notes 157, 158, 160.

under a specific age[164] is admissible if (1) the court finds the statement reliable, (2) the adverse party has notice, and (3) the child either (a) testifies or (b) is unavailable and there is corroborative evidence of the act.

The second category, comprised of nine states,[165] permit hearsay by children if the statements are reliable and the child either was cross-examined when the statement was made or is available to testify about the statement.[166]

The third category contains four states that admit the hearsay of children if the court determines (1) the statements are trustworthy and (2) the child is unavailable.[167] These states do not require corroborative evidence of the act.

Three states fall into the fourth category.[168] As is the case with states in category three, category four states require corroborative evidence of the act, even if the child is never cross-examined about the statement. The difference between category three and category four is that states falling into category four do not require that a child be available to testify in order to admit that child's hearsay under the tender years statute.

There is only one state in the fifth category. Ohio only admits child hearsay if (1) the statement is as reliable as statements admitted under other hearsay exceptions, (2) the child's testi-

---

[164] Arizona, Idaho, Minnesota, South Dakota, and Washington set the age at under 10; Alabama, Florida, Maryland, Mississippi, New Jersey, and Oregon at under 12; Illinois, Oklahoma, and Pennsylvania at under 13; and Colorado at the age where the child is entitled to protection as specified by the statute used in charging the defendant (usually 15 or younger) COLO. REV. STAT. § 18-3-411(1) (1994). *See supra* note 156 for citations of specific statutes.

[165] The nine states are Georgia, Indiana, Kentucky, Maine, Massachusetts, Michigan, Rhode Island, Texas, and Vermont.

[166] *See* GA. CODE ANN. § 24-3-16 (1993) (admitting statements by children if they are available to testify in the proceedings); R.I. GEN. LAWS § 11-37-13.1 (1993) (same); TEX. CRIM. PROC. CODE ANN. ¶ 38.072 (1994) (same); VT. R. EVID. 804a (1993) (same); IND. CODE ANN. § 35-37-4-6 (1994) (allowing a statement by a child fourteen or younger if the child either testifies or "was available for face-to-face cross-examination" when the statement was made); KY. REV. STAT. ANN. § 421.355 (if a statement is admitted, "the opposing party may cross-examine the child.") *overruled by* Drummer v. State, 183 S.W.2d 380, 382 (Ky. 1990) (*see supra* note 159); ME. REV. STAT. ANN., tit. 15, § 1205 (West 1992) (permitting a statement made under oath, in the presence of a judge or justice, and subject to cross-examination, to be admissible); MASS. ANN. LAWS, ch. 233, § 81 (1993) (if the child is found unavailable, the statement is only admitted if it was made under oath, accurately recorded, and subject to a sufficient opportunity for cross-examination); MICH. R. EVID. 803A (only applying if the declarant testifies during the same proceeding).

[167] The four states are Arkansas, California, Hawaii, and Kansas. *See supra* note 156 for citations of the specific statutes. ARK. R. EVID. 803(25) was declared unconstitutional by its Supreme Court. *See supra* note 161.

[168] The three states are Delaware, Missouri, and Nevada. *See supra* note 156 for citations of the specific statutes.

mony is not reasonably obtainable by the proponent of the statement,[169] and (3) there is "independent proof" of sexual or physical violence.[170] Category five is similar to category one in that it requires corroboration. However, like category three, category five only allows child hearsay if the child is unavailable to testify.

Lastly, there is only one state in category six. In Alaska, the child must testify in front of the grand jury or be available to testify at the trial and there must be additional evidence that corroborates the statement.[171] Category six (like categories one and five) requires corroboration, but only applies when the child testifies (like category two).

F. *Evaluation of the Different Tender Years Exceptions*

The ideal child hearsay statute should achieve three objectives: (1) encourage children to testify and be subject to cross-examination, (2) ensure the trustworthiness of the statements admitted, and (3) maintain a broad application without foregoing the first two objectives. None of the thirty-four state statutes attain all of these goals.

1. Encouraging Children to Testify While Preserving Cross-Examination

Only cross-examination can eliminate the dangers traditionally associated with hearsay.[172] It is therefore important for state laws to encourage children to testify at the proceedings, either in person or by means of videotaped depositions or closed circuit television. To satisfy this objective, a tender years statute should provide incentives for prosecutors to get children to testify and

---

[169] OHIO R. EVID. 807 (Banks-Baldwin 1993) establishes three specific ways a child's testimony is not reasonably obtainable: if (1) the child refuses to testify or claims a lack of memory after a person trusted by the child urges the child to both describe the acts contained in the statement and to testify, (2) the child is absent from the trial, the proponent of the statement has been unable to procure the child's testimony by process or other reasonable means, and it is probable that the proponent would be unable to procure the child's testimony if the trial were delayed, or (3) the child is unable to testify at the trial because of death or then-existing physical or mental illness and the illness would not improve sufficiently to permit the child to testify if the trial were delayed.

[170] *See* OHIO R. EVID. 807 (Banks-Baldwin 1993).

[171] *See* ALASKA STAT. § 12.40.110 (1993).

[172] *See supra* notes 71–91.

to be subject to cross-examination. Only states in categories one, two, and six do this. Categories two[173] and six[174] provide the most encouragement by allowing child hearsay only if the child was cross-examined when the child made the statement or if the child will be available for cross-examination when the statement is introduced. This forces prosecutors to get the child to testify if they want the judge to admit any hearsay. Category one[175] states provide some incentive by making it harder for the states to admit child hearsay if the child does not testify. These states require corroborating evidence of the act if the child is unavailable. Categories three[176] and five,[177] on the other hand, create disincentives by not allowing the hearsay if the child testifies. These states, therefore, encourage prosecutors to make children unavailable to testify. Category four[178] is neutral toward this objective since it allows hearsay regardless of whether the child testifies and does not require cross-examination.

2. Ensuring Trustworthiness of Admitted Statements

In order to survive a Confrontation Clause challenge, the tender years statute must admit only statements with "particularized guarantees of trustworthiness" or "sufficient indicia of reliability."[179] These standards are vague. Many courts, accordingly, have been forced to create guidelines to help decide if a statement is reliable[180] in the same way courts have done when child hearsay is

---

[173] *See supra* notes 165–166 and accompanying text for explanation of the category.

[174] *See supra* note 171 and accompanying text for explanation of the category.

[175] *See supra* note 163 and accompanying text for explanation of the category.

[176] *See supra* note 167 and accompanying text for explanation of the category.

[177] *See supra* notes 169–170 and accompanying text for explanation of the category.

[178] *See supra* note 168 and accompanying text for explanation of the category.

[179] *See* Idaho v. Wright, 497 U.S. 805 (1990); Ohio v. Roberts, 448 U.S. 56 (1980). ARK. R. EVID. 803(25) was found unconstitutional because it only required a "reasonable likelihood of trustworthiness." *See supra* note 161. KAN. STAT. ANN. § 60-460(dd) was found constitutional since "apparent reliability" implies "particularized guarantees of trustworthiness." *See supra* note 162.

[180] *See, e.g.*, Murray v. State, 770 P.2d 1131 (Alaska Ct. App. 1989) (listing five factors: Whether there is an apparent motive to lie, the general character of the declarant, whether more than one person heard the statements, whether the statements were made spontaneously, and the timing of the declaration and the relationship between the declarant and the witness); State v. Ryan, 691 P.2d 197 (Wash. 1984) (providing nine guidelines in applying WASH. REV. CODE ANN. § 9A.44.120 (West 1991)). Some states include guidelines within the statute itself. *See, e.g.*, HAW. R. EVID. 804(b)(6) (1993) (tests of reliability include age and mental condition of the declarant; spontaneity and absence of suggestion; appropriateness of the language and terminology of the statement given the child's age, lack of motive to fabricate, time interval between the event and the statement; and whether or not the statement was recorded);

admitted pursuant to a residual exception.[181] Unfortunately, by their very nature, the criteria used suffer from the same ambiguity that makes admission under a residual exception unpredictable and easily manipulable by sympathetic judges.

Many states have looked for other tests to help ensure that only reliable hearsay is admitted. One common approach is to require that the child be subject to cross-examination. Another method is to require that there be corroborative evidence of the act. Between these two methods, having the declarant testify and be subject to cross-examination is the superior choice.[182] States in categories two[183] and six[184] have chosen this requirement. Apart from requiring cross-examination, mandating that there be corroborative evidence of the act is the next best way to ensure that the child is not fabricating or incorrectly relating the story.[185] Corroborative evidence is any evidence, direct or circumstantial, which could support a logical and reasonable inference that the act occurred.[186]

Corroborative evidence is inferior to cross-examination as a means of ensuring the trustworthiness of the statement, however, because corroboration only tends to show that the act occurred; it does not directly support the child's allegation as to the identity of the perpetrator.[187] Only cross-examination of the child can eliminate the hearsay dangers associated with the accusation itself.

OHIO R. EVID. 807 (Banks-Baldwin 1993) (listing seven factors); MD. CODE ANN. CTS. & JUD. PROC. § 9-103.1 (1993) (listing 12 factors).

[181] *See supra* notes 148–149 and accompanying text.

[182] *See supra* note 43 and accompanying text.

[183] *See supra* note 173.

[184] *See supra* note 174.

[185] *See generally* Miller v. State, 531 N.E.2d 466 (Ind. 1988) (a child's inability to understand the nature and consequences of the oath is a significant negative factor in the statement's reliability; corroboration, however, acts as a safeguard); State v. Hunt, 741 P.2d 566, 571 (Wash. Ct. App. 1987) (corroboration requirement protects against fabricated or imagined allegations when defendant is unable to cross-examine child-victim); Tape of Hearing on Senate Bill 11, Colorado Senate Judiciary Committee, 54th General Assembly, 1st Sess. (Jan. 1983) (debate over COLO. REV. STAT. § 13-25-129).

[186] *See* State v. Hunt, 741 P.2d at 571–72. *See also* State v. Spronk, 379 N.W.2d 312, 314 (S.D. 1985) (corroborative evidence is evidence that "in some substantial degree tends to affirm" commission of the act); BLACK'S LAW DICTIONARY 344 (6th ed. 1990).

Corroborative evidence is distinguishable from cumulative evidence; to be corroborative, the evidence must be separate from that already given and tend to strengthen or confirm the matter, whereas to be cumulative, the evidence must merely augment or tend to establish a point already proven by other evidence. *See* State v. Allen, 755 P.2d 1153 (Ariz. 1988).

[187] *See* Swan v. Peterson, 6 F.3d 1373, 1380 n.4 (9th Cir. 1993).

Category one[188] combines these two techniques for assuring reliability. If the child testifies, then no corroborative evidence is required. If the child is unavailable, then there must be corroboration. This flexible approach is optimal. Corroborative evidence is sometimes difficult to obtain since sexual abuse often leaves no physical marks on the victim.[189] Therefore, necessity requires that some hearsay be admissible without corroboration. On the other hand, the accused has the right to have only reliable evidence brought against him. Accordingly, category one states admit uncorroborated hearsay only if the child testifies and is subject to cross-examination.

There is, however, one significant problem with the structure of category one. These states admit the same evidence if the child testifies as they admit with mere corroborative evidence. This method fails to treat testimony that is subject to cross-examination as more trustworthy than that which is merely corroborated. An optimal hearsay statute needs to separate statements that are considered unreliable unless the child testifies from statements that are still reliable even though the child is unavailable and there is merely corroboration.

Six states currently employ methods by which out-of-court statements are differentiated according to their trustworthiness. These states limit the types of witnesses who may introduce hearsay and the circumstances and time frame of admissible hearsay statements. Maryland admits child hearsay only if the statement was made to either a licensed physician, a licensed psychologist, a licensed social worker, or a teacher where the person was acting in the course of the individual's profession when the statement was made.[190] California's statute allows the hearsay only if the statement was made prior to the defendant's confession, is consistent with that confession, and was spoken to either a law enforcement or welfare department employee who included it in his written report of the incident.[191] Texas insists that the statement be the first statement made to a person, other than the defendant, eighteen or older.[192] In Michigan, the statement must be spontaneous, the first corroborative statement after the offense, and made immediately after the incident; any

---

[188] *See supra* note 175.
[189] *See supra* notes 28–31 and accompanying text.
[190] MD. CODE ANN. CTS. & JUD. PROC. § 9-103.1(b)(2) (Michie 1993).
[191] CAL. EVID. CODE § 1228 (West 1993).
[192] TEX. CRIM. PROC. CODE ANN. § 38.072 (West Supp. 1993).

delay must be excusable on the basis of fear or other equally effective circumstances.[193] Rhode Island insists that the statement not have been made in response to questioning calculated to lead the child to make a particular statement.[194] Vermont requires that the statement not have been taken in preparation for a legal proceeding and that if a criminal proceeding has been initiated, the statement must have been made prior to the defendant's initial appearance before a judicial officer.[195]

These restrictions have different degrees of effectiveness. Allowing only the first corroborative statement, as Texas and Michigan do, is useful. A child's description of an incident may become tainted by the influences of each adult who questions the child. Therefore, subsequent statements are less reliable than the first statement about the abuse.[196] Requiring that the statement be spontaneous and be made immediately after the offense, as Michigan does, is also beneficial. Statements in this context may be more probative than the child's subsequent testimony because they are free of any external influences or pressures.[197] Restricting admission to statements made before the defendant's confession or judicial appearance and not made in preparation for a legal proceeding, as California and Vermont do, is another valuable requirement. As soon as there is a suspect, there is a risk that adults who talk to the child will be predisposed to confirm what they have already been told.[198] Accordingly, the child is more likely to blame the defendant in order to please the adult.[199]

---

[193] MICH. R. EVID. 803A (1993).

[194] R.I. GEN. LAWS § 11-37-13.1(a)(5) (Supp. 1993).

[195] VT. R. EVID. 804a (1993).

[196] *See* Garcia v. State, 792 S.W.2d 88, 93 (Tex. Crim. App. 1990) (Clinton, J., dissenting).

[197] *See* People v. Rocha, 547 N.E.2d 1335, 1343 (Ill. Ct. App. 1989) ("The child's first accounts of an incident of sexual abuse are sometimes more reliable than later testimony given in the often intimidating courtroom setting."); State v. Robinson, 735 P.2d 801, 814 (Ariz. 1987) (a child's statements regarding sexual abuse are "valuable and trustworthy in part because they exude the naivete and curiosity of a small child, and were made in circumstances very different from interrogation or a criminal trial. . . ."); *cf.* United States v. Inadi, 475 U.S. 387, 395 (1986) (noting that an out-of-court statement made by a co-conspirator while the conspiracy was in progress would "provide evidence of the conspiracy's context that cannot be replicated, even if the declarant testifies to the same matters in court. . . .").

[198] *See generally* State v. Harris, 808 P.2d 453, 459 (Mont. 1991) (arguing that a therapist who treats a child for sexual abuse is inclined to confirm what he or she has been told).

[199] *See* Guam v. Ignacio, 10 F.3d 608, 614 (9th Cir. 1993) (noting that discussion concerning suspicion of sexual abuse between social worker and mother in child's presence "arguably provided basis for child to report inaccurately the abuse . . . in an attempt to please her mother. . . .").

Rhode Island's rule that the hearsay not be in response to questioning calculated to induce a particular response is arguably justifiable because children may invent sexual abuse in order to please an aggressive interviewer.[200] However, this prohibition might be overly restrictive since many children only relate tales of sexual abuse in response to leading questions.[201]

Lastly, Maryland's and California's requirement that the statement be made to a disinterested professional is the least useful. This prerequisite is based on the idea that only certain disinterested professionals are qualified to relate truthfully what children have told them. This concern may be exaggerated, however, since cross-examination of the adult should lower the risk of perjury. More importantly, the rule also fails to address the key issue of reliability: whether the child perceived, remembered, or related the event accurately. Still, examining to whom the statement was made and the circumstances and timing of the statement may be valuable in differentiating types of hearsay according to trustworthiness.

3. Maintaining Broad Application

The prototype hearsay statute should be as inclusive as possible, admitting hearsay even when a child is unavailable to testify. The prosecution of child molestation often would be impossible without child hearsay. Currently, only states in categories one and four admit child hearsay both when the child testifies and when the child is unavailable. Again, live or videotaped testimony plus cross-examination is preferable to the admission of hearsay. Some children, however, may be unavailable to testify. In these cases, out-of-court statements may be the only vehicle for prosecution of their attackers. The optimal tender years statute must be inclusive enough to admit child hearsay whether or not the child is available to testify.

Obviously, a tender years hearsay exception statute must only apply to children of tender years. All states limit the applicabil-

---

[200] *See supra* notes 75–79 and accompanying text.

[201] *See generally supra* note 106 and accompanying text. *See also* People v. Alvarez, 607 N.Y.S.2d 573 (N.Y. Sup. Ct. Crim. Term 1993) (denying the suppression of child hearsay pursuant to New York's tender years statute, (N.Y. FAM. CT. ACT § 1046(a)(vi) (McKinney 1994)), holding that even statements that were "the product of suggestive questioning or substandard clinical procedure" are admissible since the jury can decide the ultimate probative value of the statements).

ity of their tender years statutes to children under a certain age. However, for many states, this age is unnecessarily low. For example, ten states limit the applicability of hearsay exceptions to statements made by children under ten years old.[202] To further the goal of inclusiveness, the statute should select the highest age that will maintain the statute's integrity. Colorado matches the definition of child in its tender years statute to the definition of child in the criminal statute the defendant allegedly violated.[203] This approach is optimal because it recognizes a natural correlation between the law protecting a child from sexual contact and the law allowing the child's hearsay to help prosecute the perpetrator.

4. Achieving All Three Goals

Of the six categories of existing state statutes, category one best accomplishes the three goals of the prototype hearsay statute. It encourages children to testify because, when the child is unavailable, it requires corroborating evidence of the act. It ensures trustworthiness by mandating either that the child testify or that the prosecution introduce corroborating evidence. It is inclusive since it permits hearsay both when the child testifies and when the child is unavailable. Its only structural flaw is that it fails to differentiate according to reliability the statements that should be admissible only if the child testifies, from the statements that should be admissible even if there is merely corroboration of the statement, but no child testimony. This differentiation might be achieved, however, by examining the audience, circumstances, and timing of the statement. Therefore, the ideal child hearsay exception statute should combine the organization of category one and specific guidelines to categorize out-of-court statements by their reliability.

III. The Model Statute

The Model Statute[204] establishes a four-tiered structure for the admission of child hearsay in sexual abuse prosecutions. The

---

[202] These 10 states are Alaska, Arizona, Arkansas, Idaho, Massachusetts, Michigan, Minnesota, Nevada, South Dakota, and Washington. *See supra* note 156 for citations of the statutes.

[203] *See* Colo. Rev. Stat. § 13-25-129(1) (1993).

[204] For the full text of the Model Statute, see the Appendix. Subsequent citations to

first tier consists of the most trustworthy statements: these are statements that fit the traditional hearsay exceptions, such as the spontaneous declaration or the medical diagnosis or treatment exceptions.[205] As explained in section (D), the Model Statute does not preempt the admission of this hearsay. Therefore, the admissibility of these statements under *White v. Illinois*, 112 S. Ct. 736 (1992), does not depend on whether the child is able to testify.[206] The Model Statute does not preempt common law hearsay rules; tradition supports the extreme reliability of these exceptions. However, where many state courts have expanded the definitions of these traditional exceptions in cases involving alleged sexual abuse of a child,[207] section (D) prohibits this expansion because statements admitted under the expansions are not sufficiently trustworthy to justify admission without cross-examination or corroborative evidence of the act.[208]

The second tier covers statements admissible if (1) the statement was made by a child below the protected age according to the statute under which the defendant is being charged,[209] (2) the court finds that the statement provides particularized guarantees of trustworthiness,[210] (3) the adverse party is notified sufficiently in advance such that he or she has a fair opportunity to prepare to meet the statement,[211] and (4) the child either testifies or is found to be unavailable and there is independent corroborative evidence of the act. Section (A)(4) precludes a child deemed available under the statute's definition from refusing to testify. In the interests of justice, where the child is available, the child should testify.[212] Accordingly, section (B) requires the state to

---

sections and subsections refer to the Model Statute reprinted in the Appendix unless otherwise specified.

[205] *See supra* notes 114–116, 122–125 and accompanying text.

[206] *See supra* note 61 and accompanying text.

[207] *See supra* notes 118–120, 128–133 and accompanying text.

[208] *See supra* notes 121, 137–140 and accompanying text.

[209] Section (A). This section is based on COLO. REV. STAT. § 13-25-129(1) (West 1994). *See supra* notes 202–203 and accompanying text for argument in support.

[210] Section (A)(1). This section codifies the Confrontation Clause mandate specified in Ohio v. Roberts, 448 U.S. 56 (1980), and Idaho v. Wright, 497 U.S. 805 (1990). *See supra* notes 53–60 and accompanying text.

[211] Section (C). Almost all of the tender years statutes include a notice provision. The two most common are that the defendant be given at least 10 days notice before the trial, *see, e.g.*, FLA. STAT. ANN. § 90.803(23)(b) (West 1993); MISS. CODE ANN. § 13-1-403(2) (1993), and that the defendant be given a fair opportunity to prepare to meet the statement. *See, e.g.*, IDAHO CODE § 19-3024 (1994); WASH. REV. CODE ANN. § 9A.44.120 (West 1993). Section (C) chooses the flexible approach since this best serves the purpose of the notice requirement; that is, for the defendant to have a fair opportunity to object to and defend against the hearsay.

[212] *See generally* State v. Maben, 626 A.2d 63, 69–70 (N.J. 1993) ("'The law is

make reasonable efforts to procure the child's attendance.[213] Current law includes a number of meanings of unavailability, from mere hesitancy to testify to absolute impossibility.[214] Model Statute section (B) chooses a definition between these extreme examples by providing seven grounds on which a child may be found unavailable. These grounds resemble those in Rule 804(a) of the Federal Rules of Evidence,[215] but have been modified and expanded to apply to child declarants.

There are four important features of the Model Statute's unavailability definition. Section (B)(3) includes total failure of memory or complete inability to communicate about the alleged events since these circumstances make cross-examination impossible.[216] Section (B)(4) describes a complete refusal to testify

---

entitled to every person's evidence. A witness cannot be made the final arbiter as to whether he or she will testify . . . ." (quoting State v. Roman, 590 A.2d 686, 689 (N.J. App. Div. 1991))).

[213] *See generally* State v. Maben, 626 A.2d at 71

> Before a court admits hearsay testimony under the "tender years" exception, it must conduct a probing inquiry of the State's efforts to locate the missing witness to ensure that the search was duly diligent. Courts must scrutinize, among other factors, the State's good-faith effort in light of the seriousness of the crime, the necessity of the witness-declarant, the probability of finding the witness based on the knowledge and resources available, the means at the State's disposal, and the manner in which the State chose to undertake its search, including the timing.

Id.

[214] For a general discussion on different definitions of unavailability of child defendants, see People v. Rocha, 547 N.E.2d 1335, 1339 (Ill. App. Ct. 1989).

[215] *See* FED. R. EVID. 804(a).

> "Unavailability as a witness" includes situations in which the declarant (1) is exempted by ruling of the court on the ground of privilege from testifying concerning the subject matter of the declarant's statement; or (2) persists in refusing to testify despite an order of the court to do so; or (3) testifies to a lack of memory of the subject matter of the declarant's statement; or (4) is unable to be present or to testify at the hearing because of death or then existing physical or mental illness or infirmity; or (5) is absent from the hearing and the proponent of a statement has been unable to procure the declarant's attendance [or testimony] by process or other reasonable means. . . .

Id.

[216] Some states include both inability to communicate and total failure of memory as grounds for unavailability in their tender years statutes. *See, e.g.*, ALA. CODE §§ 15-25-32(2) (1994); DEL. CODE ANN. tit. 11, § 3513(b)(2)(a) (1993).

Additionally, several state courts have defined unavailability for child sexual assault victims to include inability to communicate about the offense. *See, e.g.*, State v. Robinson, 735 P.2d 801 (Ariz. 1987) (uncommunicative child is incapable of testifying and thus unavailable); State v. Giles, 772 P.2d 191 (Idaho 1989) (three-year-old child was incapable of communicating to the jury and was thus unavailable); People v. Rocha, 547 N.E.2d 1335 (Ill. App. Ct. 1989); State v. Chandler, 376 S.E.2d 728 (N.C. 1989) (child witness became unresponsive during testimony and was ruled unavailable); State v. McCafferty, 356 N.W.2d 159 (S.D. 1984) (child victim present in courtroom but unable to testify effectively deemed unavailable).

Rule 804(a)(3) of the *Federal Rules of Evidence* includes a lack of memory of the subject matter as a ground for unavailability. There is no provision in the Federal Rules,

despite judicial requests. It does not mandate a judicial order, which seems excessive for a child declarant.[217] Section (B)(6) provides that a child who is incompetent is unavailable as a witness.[218] Nonetheless, a child's incompetency need not negate the reliability of the hearsay; children often are incompetent simply because of their reluctance to answer questions in an unfamiliar environment.[219]

Section (B)(7) allows for unavailability because of potential psychological trauma since testifying may be harmful to some children.[220] By requiring "severe emotional or mental distress,"

---

however, for an inability to communicate about the offense as a basis for unavailability. *See* FED. R. EVID. 804(a)(3).

[217] *Cf.* FED. R. EVID. 804(a)(2).

[218] A few of the states' tender years hearsay statutes include incompetency as a basis for unavailability. *See, e.g.*, DEL. CODE ANN. tit. 11, § 3513(b)(2)(a)(7) (1993). Several courts have interpreted the general term "unavailability" in their states' tender years statutes to include incompetence. *See, e.g.*, *In re* K.G.L., 403 S.E.2d 464 (Ga. Ct. App. 1991); State v. Lanam, 459 N.W.2d 656 (Minn. 1990); State v. Rodgers, 1993 Ohio App. LEXIS 5880 (Ohio Ct. App. 1993); State v. Jones, 772 P.2d 496 (Wash. 1989) (holding that a trial court's independent determination that a child is incompetent to testify is sufficient, but the prosecution may not stipulate to the child's incompetency).

Section (B)(6) follows the current trend in the United States and presumes that children are competent to testify unless a court concludes that they cannot communicate about the incident. *See supra* note 35.

[219] *See* Oldsen v. People, 732 P.2d 1132, 1135 n.6 (Colo. 1986) (finding of incompetence does not automatically render a child's hearsay inadmissible, as long as its reliability is ensured by circumstances bringing it within a hearsay exception); Perez v. State, 536 So.2d 206, 211 (Fla. 1988) ("The fact that a child is incompetent to testify at trial . . . does not necessarily mean that the child is unable to state the truth . . . ."); People v. Cherry, 88 Ill. App. 3d 1048, 1052 (Ill. App. Ct. 1980) ("The reliability, and therefore admissibility, of a spontaneous declaration comes not from the reliability of the declarant, but from the circumstances under which the statement is made . . . .").

[220] *See supra* note 93 and accompanying text. Several of the tender years hearsay statutes include psychological harm in their unavailability determinations. *See, e.g.*, DEL. CODE ANN. tit. 11, § 3513 (1993) (one of eight grounds for unavailability is a "[s]ubstantial likelihood that the child would suffer severe emotional trauma from testifying at the proceeding or by means of a videotaped deposition or closed-circuit television . . . ."); MO. REV. STAT. § 491.075 (1994) (a child will be found unavailable, even if physically available as a witness, if the court finds that "significant emotional or psychological trauma" would "result from testifying in the personal presence of the defendant . . . .").

Additionally, some state courts have defined unavailability to include mental or emotional harm to the child. *See, e.g.*, Perez v. State, 536 So.2d 206, 207 n.1 (Fla. 1988) ("'Unavailability shall include a finding by the court that the child's participation in the trial or proceeding would result in a substantial likelihood of severe emotional or mental harm . . . .'") (citing FLA. STAT. ch. 90.803(23) (1985)); State v. Bratt, 824 P.2d 983, 989 (Kan. 1992).

> [T]he following factors are relevant to the determination of whether a victim witness is unavailable because of psychological trauma or disability: (1) the probability of psychological injury as a result of testifying, (2) the degree of anticipated injury, (3) the duration of the injury, and (4) whether the expected psychological injury is substantially greater than the reaction of the average victim of rape, kidnapping, or other violent act. Other factors may also be relevant . . . .

the Model Statute sets a higher threshold to establish unavailability than *Maryland v. Craig* requires to authorize special courtroom procedures.[221] This is fair to the accused since a defendant would rather conduct cross-examination through a videotaped deposition or closed circuit television than not conduct cross-examination at all. Since section (B) provides the only grounds for unavailability, a court will not be able to excuse a child from testifying out of mere sympathy.

Section (A)(4)(b)(i) provides that if a child is found to be unavailable, the child's statement is only admissible if there is corroborating evidence of the act. This goes beyond the minimum Confrontation Clause requirements.[222] The Model Statute restricts corroborative evidence in two ways. First, the corroboration must be independent of the child's statement.[223] Counting a child's nonverbal assertions or repetition of the statement as corroborative evidence renders the child's statement self-corroborating and robs the corroboration requirement of meaningful content. Second, other hearsay admitted pursuant to the Model Statute cannot count as corroboration. Without this limitation, a prosecutor could simply combine several otherwise inadmissible uncorroborated statements, thus rendering them admissible.

To be included in the second tier, the statement must satisfy five requirements. It must (1) have been made immediately after the offense or have been accompanied by a delay determined by

---

Id.

[221] *See supra* notes 64–65 and accompanying text.

[222] *See* Idaho v. Wright, 497 U.S. 805 at 830–31 (Kennedy, J. dissenting) ("States are of course free, as a matter of state law, to demand corroboration of an unavailable child declarant's statement as well as other indicia of reliability before allowing the statements to be admitted into evidence.").

[223] Jurisdictions are split over whether the corroboration of the alleged abuse must be independent of the child's statements. Some courts refuse to count a child's repetition of the accusation or the circumstances surrounding the statement as corroboration. *See, e.g.*, People v. Bowers, 801 P.2d 511 (Colo. 1990) (finding use of anatomical dolls and child's gestures during interviews with the police officer and the family therapist insufficient corroboration); Beck v. State, 544 N.E.2d 204 (Ind. Ct. App. 1989) (finding that child's complaint to parent and doctor was insufficient corroboration). Other courts count repetition as corroboration. *See, e.g., In Re* Christina F., 548 N.E.2d 1294 (N.Y. 1989) (interpreting corroboration requirement of Family Court Act section 1046(a)(vi) to hold that a child's hearsay describing sexual abuse may be corroborated by the child's later cross-examined, but unsworn, in-court testimony). Additionally, some courts have held that precocious sexual knowledge manifested in a child's statement is admissible to corroborate the hearsay. *See, e.g.*, Murray v. State, 770 P.2d 1131, 1138 (Alaska Ct. App. 1989); State v. Swan, 790 P.2d 610, 615 (Wash. 1990) ("[T]o give any real effect to the child victim hearsay statute, the corroboration requirement must reasonably be held to include indirect evidence of abuse. Such evidence . . . include[s] a child victim's precocious knowledge of sexual activity . . . .").

the court to be consistent with truth,[224] (2) not have been made in preparation for a legal proceeding,[225] (3) have been the first statement about the offense to a person eighteen or older other than the defendant,[226] (4) have been made prior to the defendant's arrest,[227] and (5) not concern a parent or a significant other of the parent[228] if the statement was made when the parent or significant other of the parent was divorced, separated, or in a similar dispute with the child's other parent or significant other of the parent.[229] All five of these requirements pertain to reliability. Since tier two admits hearsay even if the child is never cross-examined about the statement, all five of the criteria must be satisfied for an out-of-court statement to be in tier two.

Tier three accepts any statement that satisfies all of tier two's requirements except that the statement either was not the first statement about the offense, was made after the defendant's arrest, or concerns a parent or a significant other of the parent if the statement was made when the parent or significant other of the parent was divorced, separated, or in a similar dispute with the child's other parent or significant other of the parent. The Model Statute insists on cross-examination of these statements as a prerequisite to admissibility because the circumstances surrounding these statements suggest that the child may have been influenced by an adult.[230] Hearsay in tier three is admissible only if the child testifies and is subject to cross-examination at the proceeding or by means of a videotaped depo-

---

[224] This provision is based on MICH. R. EVID. 803A(3) (1993). *See supra* note 193 and accompanying text.

[225] This provision is based on VT. R. EVID. 804a(a)(2) (1993). *See supra* note 195 and accompanying text.

[226] This provision is based on TEX. CODE CRIM. PROC. ANN. art. 38.072 § 2(a)(2) (West 1994) and MICH. R. EVID. 803A (1993). *See supra* notes 192, 193, 196 and accompanying text.

[227] *Compare* CAL. EVID. CODE § 1228(c) (1994) (insisting that the statement be made before the defendant's confession); VT. R. EVID. 804a(a)(2) (1993) (mandating that if a criminal proceeding has been initiated, the statement must have been made prior to the defendant's initial appearance before a judicial officer). This tier of the Model Statute admits as evidence only statements made prior to the defendant's arrest to avoid the adult's predisposition to confirm what she may have been told about a suspected perpetrator. *See supra* notes 89–90 and accompanying text.

[228] "Significant other of the parent" is deliberately imprecise. The term allows the trial judge the flexibility to decide if the adults' relationship suggests a heightened probability of false allegations.

[229] None of the existing tender years statutes for criminal prosecutions includes this criterion. The Model Statute incorporates this requirement because false accusations are more common during divorce and custody situations. *See supra* notes 9, 90 and accompanying text.

[230] *See supra* note 9 and accompanying text.

sition or closed circuit television. Corroborative evidence of the act does not help determine whether the child was persuaded to make the accusation.[231] Only cross-examination of the child tests the reliability of the child's identification of the perpetrator.[232]

The least reliable statements are those not made immediately after the offense or those made in preparation for a legal proceeding. Children's memories become distorted over time.[233] An adult preparing for a judicial proceeding may prompt a child to confirm her case. Accordingly, these types of hearsay are particularly suspect.[234] Furthermore, if the child related the story long after the event, the child is more likely to be available and competent to provide cross-examined testimony at the proceeding. Therefore, statements in tier four are both the least probative and least necessary. Accordingly, hearsay in tier four is not admissible even if the child testifies at the proceeding.

IV. Conclusion

The Model Statute proposed in this Note will not guarantee that only honest and correct hearsay will be found admissible. Neither will it ensure perfect accuracy with respect to convicting those who are truly guilty. No statute or judicial decree can achieve those goals. Lawyers and others with personal agendas will always attempt to manipulate the rules to achieve their desired ends. A statute needs to be enacted and enforced in order to discover its loopholes and drawbacks.

On the other hand, the Model Statute will provide specific, logical criteria to determine the trustworthiness of out-of-court statements and will require more precautions before admitting statements of lesser reliability. In this way, it balances the prosecutorial necessity of admitting trustworthy hearsay, the values inherent in the Confrontation Clause, and the current hearsay rules and exceptions. States, therefore, should replace some of their recent modifications of the hearsay rules with tender years statutes based on the Model Statute.

---

[231] *See supra* notes 182–187 and accompanying text.
[232] *See supra* note 91 and accompanying text.
[233] *See supra* notes 80–82 and accompanying text.
[234] *See generally supra* part I.C.

APPENDIX

MODEL STATUTE

(A) A statement made by a child, under the protected age in the statute for which the defendant is being charged, describing any act of sexual contact on or with the child, is admissible in evidence in a criminal proceeding if:

(1) The court finds, in a hearing conducted outside the presence of the jury, that the time, content, and circumstances of the statement provide particularized guarantees of trustworthiness;

(2) The statement was made immediately after the offense or the court finds the delay consistent with truth;

(3) The statement was not made in preparation of a legal proceeding; and

(4) The child either:

(a) Testifies and is subject to cross-examination at the proceeding or by means of a videotaped deposition or closed circuit television; or

(b) Is unavailable as a witness: PROVIDED, That when the child is unavailable as a witness, such statement may be admitted only if:

(i) There is independent corroborative evidence of the act. Independent corroboration does not include hearsay admitted pursuant to this statute;

(ii) The statement was the first statement about the offense made by the child to a person eighteen or older, other than the defendant;

(iii) The statement was made prior to the defendant's arrest; and

(iv) The statement does not concern a parent or a significant other of the parent if the statement was made when the parent or significant other of the parent was divorced, separated, or in a similar dispute with the child's other parent or significant other of the parent.

(B) The unavailability requirement of section (A)(4)(b) is satisfied only if the proponent of the statement has been unable, despite reasonable efforts, to procure the declarant's attendance at either the proceeding or a videotaped deposition, and if the child is found by the court to be unavailable to testify on at least one of the following grounds:

(1) The child's death or then existing physical or mental illness or infirmity;

(2) The child's absence from the jurisdiction;

(3) The child's total failure of memory or complete inability to communicate about the offense;

(4) The child's persistent refusal to testify despite judicial requests to do so;

(5) The child's exemption on the ground of privilege from testifying on the subject matter of the declarant's statement;

(6) The child's incompetency (a child is presumed competent to testify, but if the competency is questioned, then the court must determine if the child is able to describe or relate the events or facts in language appropriate for a child of that age); or

(7) A substantial likelihood that the child would suffer severe emotional or mental distress from testifying at the proceeding or by means of a videotaped deposition or closed circuit television.

(C) A statement may not be admitted under this statute unless the proponent of the statement makes known to the adverse party his/her intention to offer the statement, the content of the statement, the approximate time, date, and location of the statement, the person to whom the statement was made, and the circumstances surrounding the statement that indicate its reliability. Notice must be given sufficiently in advance of the proceedings to provide the adverse party with a fair opportunity to prepare to meet the statement.

(D) The statute does not preempt admission of any statement under a traditional hearsay exception. Courts, however, are prohibited from (1) applying more liberal definitions in other hearsay exceptions for statements made by children when under the protected age in the statute for which the defendant is being charged than they do for adults and (2) admitting via a residual hearsay exception statements that satisfy the description in the first paragraph of section (A).

NOTE

BROADENING THE SCOPE OF COUNSELOR-PATIENT PRIVILEGE TO PROTECT THE PRIVACY OF THE SEXUAL ASSAULT SURVIVOR

ANNA Y. JOO*

Sexual assault survivors are encouraged to vent their feelings and explore their sometimes confused memories with counselors. This same information can often be readily used by the defense counsel of the accused in order to impeach the testimony of the survivor. In response, many states have enacted laws that grant some form of privilege to counselor-patient communications.

In this Note, Ms. Joo closely examines the privacy interest of the sexual assault survivor and the confrontation right of the defendant in order to assist courts and legislatures in striking a balance between the two competing rights. She explores the benefits of an absolute privilege; a limited privilege that involves in camera *inspections by judges; and a limited privilege that allows defense counsel to examine the counseling records as an officer of the court. Ms. Joo finally advocates a "semi-absolute" privilege that provides a guarantee of absolute confidentiality to the survivor's counseling communications, except in specific circumstances.*

She said:

Amy and Jonathan were good friends during their sophomore year in college, and that night, they made plans to go out to dinner. After dinner, they went back to her dorm room to sneak a couple of drinks. Neither of them was drunk, but Amy began to feel uncomfortable as Jonathan started making sexual advances towards her. She left her room, hoping that he would go away, but he came out and brought her back. Once they were back inside her room, Jonathan began to undress. He pushed Amy from the chair onto the bed and got on top of her. Holding her down, he raped her.

Amy subsequently sought therapy but was warned by the rape counselor that anything Amy said during their sessions could be made public in a criminal trial. Amy went ahead with the counseling but is reluctant to press criminal charges.

---

* Judicial Law Clerk, the Honorable Herbert Y.C. Choy, U.S. Court of Appeals, Ninth Circuit; B.A. 1991, Carleton College; J.D., Harvard Law School, 1994. The author dedicates this Note to her parents for their love, support, and encouragement throughout her education. The author wishes to thank Judge John Cratsley, Massachusetts Superior Court; Professor Charles Nesson, and Kathleen Sullivan, assistant district attorney of Middlesex County, Massachusetts, for their invaluable assistance in the writing of this Note.

> *She does not want to risk exposing her personal thoughts and fears to a room full of strangers.*

He said:

> *Jonathan has a different version of what happened that night. After dinner, they returned to her room, put on some music, danced, and kissed. He had an errand to run, so he left for a while. When he came back, he found her wearing a bathrobe. The two continued the previous intimacies of earlier that evening, eventually disrobing and engaging in consensual intercourse. They had intercourse on numerous occasions after that night. She even invited him over to her parents' home for a visit.*
>
> *Jonathan is amazed that Amy could have such a different account of what transpired that evening. He believes that she decided to falsely accuse him of rape because she fears her parents' reaction if they discover that she had been sexually active. He is certain that he can demonstrate her emotional instability and her motive to lie if he could access her counseling records, particularly those portions in which she talks about the alleged rape.*[1]

Personal testimonial privilege, which necessarily excludes relevant and nonprejudicial information from the trial process, is antithetical to the truth-finding principle of the U.S. judicial system.[2] Testimonial privileges, however, have existed to protect certain communications from public disclosure and to foster selected relationships.[3] To determine the proper scope of the counselor-patient privilege[4] for sexual assault survivors, the judicial

---

[1] This hypothetical situation is loosely based on the fact pattern of *Commonwealth v. Stockhammer*, 570 N.E.2d 992, 995–96 (Mass. 1991).

[2] ERIC D. GREEN & CHARLES R. NESSON, PROBLEMS, CASES, AND MATERIALS ON EVIDENCE 520–21 (1983) (excerpting from United States v. Nixon, 418 U.S. 683 (1974)). *See* Dennis v. United States, 384 U.S. 855, 870 (1966) (observing that "disclosure, rather than suppression, of relevant materials ordinarily promotes the proper administration of criminal justice.").

[3] Catharina J.H. Dubbelday, Comment, *The Psychotherapist-Client Testimonial Privilege: Defining the Professional Involved*, 34 EMORY L.J. 777, 777 (1985). *See also* CHARLES MCCORMICK, MCCORMICK ON EVIDENCE § 72, at 269 (J.W. Strong, ed., 4th ed. 1992) ("[Privilege protects] interests and relationships which, rightly or wrongly, are regarded as of sufficient social importance to justify some sacrifice of availability of evidence relevant to the administration of justice.").

[4] In the context of privilege, this Note uses the term "counselor" as a short-hand method of referring to the many different types of counselors who may provide immediate emotional assistance to a sexual assault survivor in the aftermath of the violence. These counselors may be sexual assault counselors, social workers, physicians, psychiatrists, psychotherapists or psychologists. In the area of immediate counseling following a sexual assault, this Note argues that the testimonial privilege applies equally to all of these types of counselors. *See infra* part I.D, notes 52–69 and accompanying text.

system must strike a delicate balance between the needs of the defendant to examine the survivor's counseling records[5] and the equally strong privacy interest of the sexual assault survivor in maintaining the confidentiality of her counseling records.[6]

The balancing process is affected by the weight a particular court places upon the value of counseling to the sexual assault survivor and to society.[7] A court is less likely to uphold the privilege if it views it as operating to withhold important evidence that may damage the complainant's case.[8] On the other hand, a particular court is likely to uphold the privilege if it perceives that counseling relationships contribute greatly to the survivor's recovery process and that the defendant's confrontation right is already unnecessarily broad.

This Note closely examines the privacy interest of the sexual assault survivor and the confrontation right of the defendant in order to assist courts and legislatures in striking the optimal balance between the two competing rights. Part I of the Note explores the various aspects of the sexual assault survivor's privacy interest. Section A examines the various justifications for the existence of testimonial privileges, specifically, the counselor-patient privilege as it applies to sexual assault survivors. Section B proposes a constitutional basis for asserting privilege in a counseling relationship. Section C describes the psychological trauma faced by sexual assault survivors in order to gain a fuller appreciation of the interest at stake. Finally, Section D lists the benefits both to the individual survivor and to the general public that immediate counseling provides.

Part II examines the other side of the balance: the right of the accused to obtain access to the sexual assault survivor's confidential counseling records. Section A explores the reach of the Compulsory Process and the Due Process Clauses in accessing the survivor's confidential counseling records. Similarly, Section B investigates how far the Confrontation Clause extends to grant the accused access to confidential records.

---

[5] The need to examine the records arguably may qualify as a constitutional right under the Confrontation Clause of the U.S. Constitution. *See infra* part II.

[6] *See infra* part I.B, notes 24–35 and accompanying text.

[7] Beth Stouder, Note, *Criminal Law and Procedure (Evidence)—Pennsylvania Establishes New Privilege for Communications Made to a Rape Crisis Center Counselor*—In re Pittsburgh Action Against Rape, *494 Pa. 15, 428 A.2d 126 (1981)*, 55 TEMP. L.Q. 1124, 1146 (1982).

[8] *Id.* at 1147.

Part III explores the four types of privilege that courts can grant to a sexual assault survivor's counseling communications. Section A examines the absolute privilege. Section B probes the limited privilege either with a prior *in camera* inspection by the trial judge or with direct access by the defense counsel. Direct access to confidential files presents a dilemma for the defense attorney, who must strike a proper balance between the roles of a zealous advocate and an officer of the court. Finally, Section C proposes the semi-absolute privilege that provides a guarantee of absolute confidentiality to the survivor except in certain well-defined circumstances.

In order to examine the interests of the defendant and the survivor in a concrete context, Part IV surveys the recent turmoil in Massachusetts's judicial and legislative systems over the interpretation of privilege granted to confidential counseling records of sexual assault survivors and contrasts it with the judicial and legislative experiences with absolute privilege in Pennsylvania. Section A outlines the difficult challenges faced by Massachusetts courts in defining the proper scope of access to confidential records. Section B discusses the recent legislative efforts to secure more confidentiality for counseling records of sexual assault survivors. Section C examines the sexual assault counselor privilege, which is written as an absolute privilege, as a possible avenue of obtaining more protection for the survivor. Finally, Section D surveys Pennsylvania's case and legislative history in order to explore how its courts have upheld the constitutionality of an absolute counseling privilege. The Note concludes by calling for a clearly defined scope of privilege that guarantees absolute confidentiality, absent some limited, articulated circumstances.

I. The Privacy Interest of the Sexual Assault Survivor

A. *Justifications for Testimonial Privileges*

Three types of justifications or explanations exist for testimonial privileges: the utilitarian rationale, the privacy justification, and the power theory.[9] The utilitarian rationale balances the so-

[9] Kerry L. Morse, Note, *A Uniform Testimonial Privilege for Mental Health Professionals*, 51 Ohio St. L.J. 741, 742–44 (1990).

cietal benefit derived from protecting the privileged communication against the benefit produced in litigation from disclosure.[10] The privilege is considered to be necessary if the benefit derived from encouraging open communication outweighs the cost of lost information in litigation. The existence of a privilege is justified often on the grounds of how well a privilege meets the four criteria first articulated by Dean Wigmore. The criteria are:

> (1) Does the communication originate in the belief that it will not be disclosed? (2) Is the inviolability of that confidence essential to achieve the purpose of the relationship? (3) Is the relationship one that society should foster? (4) Is the expected injury to the relationship, through fear of later disclosure, greater than the expected benefit to justice in obtaining later testimony?[11]

Although Dean Wigmore never evaluated the counselor-patient privilege for how well it meets these four criteria, a subsequent examination indicates that such a privilege can be justified.[12] Knapp and VandeCreek argue that the counselor-patient relationship satisfies the first criterion because a relationship between a counselor and her patient implies a contract. The counselor promises trust and confidentiality in exchange for the patient's frank communication of her innermost thoughts and fears.[13] The second criterion is satisfied because this implied contract is at the core of the counseling relationship, and successful treatment can only be achieved by the maintenance of confidentiality.[14]

That the public fosters and supports counseling relationships, the third criterion, is evidenced by governmental funds expended for mental health services and the availability of psychotherapy in institutions such as the armed services, prisons, and universities.[15] Finally, although no empirical evidence exists on the subject, Knapp and Vandecreek assert that courts lose little information because of the counselor-patient privilege.[16] If the excluded information is of minimal relevance in a sexual assault trial,[17] or

---

[10] 8 JOHN H. WIGMORE, EVIDENCE IN TRIALS AT COMMON LAW §§ 2191, 2192, 2285 (John T. McNaughton rev. ed. 1961). *See, e.g.*, United States v. Nixon, 418 U.S. 683 (1974) (employing the utilitarian justification to break the presumptive privilege for Presidential communications).

[11] SAMUEL KNAPP & LEON VANDECREEK, PRIVILEGED COMMUNICATIONS IN THE MENTAL HEALTH PROFESSIONS 9 (1987) (quoting WIGMORE, *supra* note 10, § 2285).

[12] *Id.* at 9–11.

[13] *Id.* at 9.

[14] *Id.* at 10.

[15] *Id.* at 10–11.

[16] *Id.* at 11.

[17] *See* People v. Arenda, 330 N.W.2d 814, 817 (Mich. 1982) (upholding the consti-

if the impact of rummaging through the confidential records will cause substantial damage, the privilege should prevail over the defendant's right to confront and cross-examine adverse witnesses.

Whereas the utilitarian rationale views the goal of the counselor-patient privilege as promoting beneficial future relations, the privacy justification perceives the main purpose of the privilege as shielding the patient from the harm that disclosure may cause.[18] According to the privacy justification, some human relationships are fundamental to human dignity and should be free from state interference.[19] This idea is especially appropriate in the context of a sexual assault survivor-counselor relationship where "[t]here appears to be something harsh or even cruel about using the spontaneous words of a trauma victim seeking help against herself in a criminal trial against her assailant."[20]

The power theory takes a different approach than either the utilitarian rationale or the privacy justification. This theory looks to the political reality within legislatures to explain the existence of certain privileges and their scope.[21] Determining the scope of privilege is necessarily a subjective activity because it involves the weighing of two competing interests. In general, privilege remains vulnerable to the manipulations of certain segments of society. For example, professions with powerful and rich clients are more likely to have the ability to lobby the legislature successfully for a greater scope of privilege.[22]

Under Rule 501 of the Federal Rules of Evidence, each state is free to create and define the scope of privilege for certain relationships to reflect the changing social attitudes within that state.[23] Application of privilege can be extended to new professions or groups as their socially beneficial functions gain recog-

---

tutionality of rape shield laws on the basis that prior sexual conduct has minimal relevance in a criminal trial for rape).

[18] Robert Weisberg, Note, *Defendant v. Witness: Measuring Confrontation and Compulsory Process Rights Against Statutory Communications Privileges*, 30 STAN. L. REV. 935, 943 (1978) (referring to the theory as the "nonutilitarian view" and *citing* David W. Louisell, *Confidentiality and Confusion: Privileges in Federal Courts Today*, 31 TUL. L. REV. 101 (1956)).

[19] Morse, *supra* note 9, at 744 (*citing* KNAPP & VANDECREEK, *supra* note 11, at 13). *See also* Thomas Krattenmaker, *Testimonial Privileges in Federal Courts: An Alternative to the Proposed Federal Rules of Evidence*, 62 GEO. L.J. 61, 85–94 (1973).

[20] KNAPP & VANDECREEK, *supra* note 11, at 114.

[21] Morse, *supra* note 9, at 744.

[22] *Id.*

[23] FED. R. EVID. 501 ("However, in civil actions and proceedings, with respect to an element of a claim or defense as to which State law supplies the rule of decision, the

nition, and the scope of existing privilege can be expanded when the need for broader protection becomes necessary. These changes occur as society's views on rights and its understanding of certain issues continue to evolve. According to the power theory, the scope of privilege granted to the counselor-patient relationship will be expanded as awareness of the importance of the privilege increases and as lobbying groups for sexual assault survivors become more influential.

B. *The Constitutional Basis for the Counselor-Patient Privilege*

The constitutional basis for an individual's general right to privacy has been recognized in the area of reproduction and applied within the scope of the physician-patient relationship. In *Griswold v. Connecticut*,[24] the U.S. Supreme Court struck down a statute forbidding the sale of birth control devices to married couples. The majority opinion, written by Justice Douglas, found that several amendments of the Bill of Rights created a "penumbra" of privacy for individuals.[25] In a subsequent decision, *Carey v. Population Services International*,[26] the Court stated, "[r]ead in the light of its progeny, the teaching of *Griswold* is that the Constitution protects individual decisions in matters of childbearing from unjustified intrusion by the State."[27] In *Roe v. Wade*,[28] the Court, believing that a woman has a right to privacy that is "founded in the Fourteenth Amendment's concept of personal liberty and restrictions on state action,"[29] limited the government's right to intrude on a woman's decision to have an abortion in certain situations. The right to privacy articulated in such case law, however, is limited and must yield to narrowly drawn, compelling state interests.[30]

---

privilege of a witness, person, government, State, or political subdivision thereof shall be determined in accordance with State law.").

[24] 381 U.S. 479 (1965).

[25] *Id.* at 484. Douglas examined the First, Third, Fourth, and Fifth Amendments' protection of particular privacy interests and concluded that there was a general right of privacy emanating from these amendments.

[26] 6431 U.S. 678 (1977).

[27] *Id.* at 687.

[28] 410 U.S. 113 (1973), *reh'g denied*, 410 U.S. 959 (1973).

[29] *Id.* at 153.

[30] Morse, *supra* note 9, at 750. *See, e.g.*, Whalen v. Roe, 429 U.S. 589 (1977)

Relying on these decisions, some courts have acknowledged that a counselor-patient privilege falls within the zone of privacy guaranteed by the Bill of Rights.[31] Further, some state courts have determined that a counselor-patient privilege is protected by a right to privacy found in their respective state constitutions.[32] Recognition of a constitutionally based privilege for communications between a sexual assault survivor and her counselor is likely to mean a greater scope of protection for confidential records.[33] However, freezing the counselor-patient privilege in "constitutional ice" may prevent the privilege from meeting the changing needs of society.[34] Statutes provide an alternate and more flexible basis for protecting the counselor-patient relationship.[35]

C. *Psychological Trauma Faced by Sexual Assault Survivors*

Rape has been described as a "total assault on an individual,"[36] with physical, psychological, and social effects.[37] Common reactions to sexual assault are emotional shock, disbelief, embarrassment, shame, guilt, depression, powerlessness, disorientation, retriggering, denial, fear, anxiety, and anger.[38] Sexual assault survivors

---

(upholding a state law requiring pharmacists to furnish information about patients who use particular drugs).

[31] *See, e.g.*, In re Lifschutz, 467 P.2d 557 (Cal. 1970); In re "B", 394 A.2d 419 (Pa. 1978).

[32] KNAPP & VANDECREEK, *supra* note 11, at 19. *See, e.g.*, Falcon v. Alaska Pub. Offices Comm'n, 570 P.2d 469 (Alaska 1977) (invalidating a law, based on the right of privacy found in the Alaska Constitution); Reynaud v. Superior Court, 138 Cal. App. 3d 1 (1st Dist. 1982) (recognizing a limited right to privacy based on the California Constitution).

[33] *Cf. Stockhammer*, 570 N.E.2d at 1002 (quoting Commonwealth v. Two Juveniles, 491 N.E.2d 234 (Mass. 1986), "[I]n appropriate circumstances, even absolute statutory privileges (nonconstitutionally based) must yield to a defendant's constitutional right to use privileged communications in his defense . . . .").

[34] KNAPP & VANDECREEK, *supra* note 11, at 21.

[35] *Id.*

[36] Maxine H. Neuhauser, *The Privilege of Confidentiality and Rape Crisis Counselors*, 8:3 WOMEN'S RIGHTS L. REP. 185, 186 (1985). *See also* In re Pittsburgh Action Against Rape, 428 A.2d 126, 138–40 (Pa. 1981) [hereinafter *PAAR*] (Larsen, J., dissenting), *superseded by statute as stated in* Commonwealth v. Wilson, 602 A.2d 1290 (Pa. 1992); Bridget M. McCafferty, *The Existing Confidentiality Privileges as Applied to Rape Victims*, 5 J.L. & HEALTH 101, 103 (1990–91).

[37] *PAAR*, 428 A.2d at 138–40.

[38] ROCHEL GROSSMAN & JOAN SUTHERLAND, eds., SURVIVING SEXUAL ASSAULT (1983). *See also* JOYCE E. WILLIAMS & KAREN A. HOLMES, THE SECOND ASSAULT: RAPE AND PUBLIC ATTITUDES 83–84 (1981).

experience negative effects of the assault as measured by degrees of crisis, feelings about men, level of functioning, and health concerns.[39]

Burgess and Holmstrom were the first to identify a typical set of emotional reactions commonly experienced by sexual assault survivors, which they called the Rape Trauma Syndrome (RTS).[40] RTS is a two-stage process.[41] In the immediate stage, the survivor's life is completely disrupted, and she[42] experiences a wide range of emotions, such as fear, anger, and anxiety. The survivor either expresses these emotions, through such behavior as crying, restlessness, and tenseness, or attempts to control these emotions by hiding behind a calm mask.[43]

This initial stage is followed by a long-term process in which the survivor reorganizes her lifestyle on the physical, emotional, and behavioral levels.[44] This latter period may be fraught with nightmares[45] and is characterized by a wide range of emotions, with fear of physical violence and death as the primary feelings described.[46] The survivor may also exhibit "strong avoidance responses" to situations that may remind her of the assault.[47]

In addition to these symptoms, a sexual assault survivor's most common response is a belief that she is somehow responsible for the crime.[48] The survivor's feelings of self-blame are often magnified by social attitudes that blame the survivor for the assault. In fact, a common rationalization for rape is that a survivor possesses certain characteristics, such as provocative-

---

[39] Joyce E. Williams, *Secondary Victimization: Confronting Public Attitudes About Rape*, 9 VICTIMOLOGY: AN INTERNATIONAL JOURNAL 67, 69 (1984).

[40] Ann W. Burgess & Lynda L. Holmstrom, *Rape Trauma Syndrome*, 131 AM. J. PSYCHIATRY 981 (1974).

[41] Subsequent research has conceptualized rape trauma in terms of symptoms rather than stages of recovery. Patricia A. Frazier, *Rape Trauma Syndrome: A Review of Case Law and Psychological Research*, 16:3 LAW & HUMAN BEHAVIOR 293, 299 (1992).

[42] This Note typically uses "she" or "hers," rather than "he" or "his," because the vast majority of rape survivors are women.

[43] Burgess & Holmstrom, *supra* note 40, at 982.

[44] Neuhauser, *supra* note 36, at 186; Morrison Torrey, *When Will We Be Believed? Rape Myths and the Idea of a Fair Trial in Rape Prosecution*, 24 U.C. DAVIS L. REV. 1013, 1064 (1991).

[45] Judith V. Becker, et al., *The Effects of Sexual Assault on Rape and Attempted Rape Victims*, 7 VICTIMOLOGY: AN INTERNATIONAL JOURNAL 106, 107 (1982).

[46] Burgess & Holmstrom, *supra* note 40, at 982–85.

[47] Becker, *supra* note 45, at 107.

[48] Rorie Sherman, *Rape Victims' Records Vulnerable: Massachusetts Prosecutors, Therapists See a Chilling Effect*, NAT'L L.J., Dec. 28, 1992, at 1, 27 (quoting Marianne Winters of the Rape Crisis Program of Worcester, Inc.). *See also* McCafferty, *supra* note 36, at 105.

ness or flirtatiousness, which give some men the right to retaliate with rape.[49] Sexual assault survivors also blame themselves, not for the assailant's actions, but for being an "occasion" for rape.[50] The survivor may call her own sense of judgment into question, especially if the assailant was an acquaintance.[51] These feelings of doubt and insecurity, recorded in counseling files, can be exploited easily by a defense counsel looking for exculpatory evidence.

D. *The Private and Public Benefits of Counseling*

A sexual assault survivor takes the first step to recovery by talking about the rape experience in a nonjudgmental atmosphere.[52] This nurturing environment is created through a firmly established relationship with a well-trained counselor in the initial stages of trauma.[53] The very personal nature of this counselor-client relationship necessitates the confidentiality of such communication.[54] Immediate treatment is of utmost importance,[55] and continued counseling may be necessary for a period of three to twelve months.[56]

The importance of the role that sexual assault counselors play in the recovery process is indicated by the fact that when sexual assault survivors were asked to identify the person who was the most helpful in the recovery process, rape crisis workers were

---

[49] Distinguishing the victim from the rest of society makes individuals feel more comfortable with their vision of a just world where there is a rational reason for everything that happens. Florence L. Denmark & Susan B. Friedman, *Social Psychological Aspects of Rape*, *in* VIOLENCE AGAINST WOMEN: A CRITIQUE OF THE SOCIOBIOLOGY OF RAPE 61 (Susan Sunday & Ethel Tobach, eds., 1985).

[50] Patricia A. Resick, *The Trauma of Rape and the Criminal Justice System*, 9:1 JUST. SYS. J. 52, 55 (1984).

[51] *Id.*

[52] LEE MADIGAN & NANCY C. GAMBLE, THE SECOND RAPE: SOCIETY'S CONTINUED BETRAYAL OF THE VICTIM 123 (1989). *See also* Donna St. George, *The Dilemma: A Fair Trial vs. a Victim's Privacy. The Confidentiality Issue Has Set the Stage on a Protracted Legal Battle*, PHILA. INQUIRER, Oct. 29, 1989, at C3 (quoting Karen Kulp, executive director of Women Organized Against Rape, "[t]he role of counseling is to deal with somebody's feelings; it's not to deal with the facts of a case. We don't ask them to reveal in detail what happened during the assault. We deal with their feelings.").

[53] *See* MADIGAN & GAMBLE, *supra* note 52, at 124–25.

[54] *Developments in the Law—Privileged Communications*, 98 HARV. L. REV. 1450, 1543 (1985) [hereinafter *Developments*] (recognizing the importance of confidentiality in a general psychotherapeutic treatment).

[55] The most favorable prognosis for treatment of acute rape trauma can be obtained if the survivor is seen immediately after the sexual assault. Burgess & Holmstrom, *supra* note 40, at 54.

[56] Neuhauser, *supra* note 36, at 196.

ranked the highest, even above significant others, fathers, and clergy.[57] Crisis intervention counselors help the survivor work through her own trauma and give emotional support to the survivor as well as her family and friend.[58] Without an absolute assurance of confidentiality, however, a sexual assault survivor is more apt than the general population to be deterred from seeking counseling.[59] Unlike others who seek counseling, sexual assault survivors face a very real possibility that their records may be disclosed if they choose to prosecute the offender.[60]

The clinician working with a sexual assault survivor immediately after the crime often has only one chance to help with the survivor's recovery. After seeing to immediate needs such as requests for medical intervention, police attention, and psychological support, the clinician attempts to establish a supportive relationship in an atmosphere of tolerance and acceptance, devoid of expressed or implied condemnation.[61] The health care professional then typically encourages the sexual assault survivor to vent her feelings, an important step along the road to assimilating the traumatic experience.[62] In follow-up meetings, the clinician resolves issues involving support for significant others, devises a practical plan for the future safety of the survivor, and arranges for long-term counseling sessions.[63]

---

[57] WILLIAMS & HOLMES, *supra* note 38, at 90. *See also* SEDELLE KATZ & MARY ANN MAZUR, UNDERSTANDING THE RAPE VICTIM: A SYNTHESIS OF RESEARCH FINDINGS 287 (1979).

[58] ELAINE HILBERMAN, THE RAPE VICTIM 41 (1976) (noting that a counselor's help to family and friends is important because sexual assault survivors sometimes displace their anger beyond the assailant onto family members or close friends who have not been as supportive to the survivor during the recovery process). *See also* WILLIAMS & HOLMES, *supra* note 38, at 83–84.

[59] Neuhauser, *supra* note 36, at 195. *See also* Dubbelday, *supra* note 3, at 800–01 ("[I]t is unlikely that people are going to resist going to a doctor for treatment because they fear public disclosure of their ailments . . . the opposite is true of clients needing psychotherapeutic treatment."); Stouder, *supra* note 7, at 1146 ("The need for confidentiality is illustrated by the increase in the number of anonymous calls to PAAR and the number of women seeking to have their files destroyed after the lower court's decision in *PAAR*."). *But see* DANIEL SHUMAN & MYRON WEINER, THE PSYCHOTHERAPIST-PATIENT PRIVILEGE 81 (1987) (Shuman and Weiner conducted a study that found that lay persons were generally unaware of whether their state had a privilege. The survey, however, measured the reactions of only psychotherapy patients and is very different from the situation in which a sexual assault survivor confides in a counselor. In the latter situation, the possibility of an imminent criminal trial is greater, and this element makes it much more likely that the patient will consider the privilege.).

[60] Neuhauser, *supra* note 36, at 195.

[61] MARY KOSS & MARY HARVEY, THE RAPE VICTIM: CLINICAL AND COMMUNITY APPROACHES TO TREATMENT 104–06 (1987).

[62] *Id.* at 106.

[63] *Id.* at 107–11.

The counseling function can be performed by sexual assault counselors, social workers, physicians, psychiatrists, or psychologists. Therefore, in fashioning a privilege, courts should focus on the function of the counseling relationship rather than on the identity of the counselor. Focusing on the purpose of the communication rather than on the occupation of the counselor avoids the social inequality created by granting a privilege to one type of counselors, such as psychiatrists and psychologists whose clients tend to be more affluent, while denying the privilege to another type, such as social workers and sexual assault counselors whose clients tend to be poor.[64] In order to create boundaries for the privilege, however, courts should restrict the privilege to mental health professionals, which include all persons who are trained to "participate in the diagnosis or treatment of mental or emotional conditions."[65]

Counseling relationships do not simply help the survivors in rebuilding their self esteem and their lives but inure as well to the benefit of the general public. Counseling increases the level of general public health by encouraging sexual assault survivors to obtain medical treatment for their injuries and to test for AIDS and other sexually transmitted diseases.[66] Further, sexual assault counseling helps to break the cycle of "survivors-turned-assailants," to prevent future crimes.[67] Finally, active support from counselors encourages the reporting and prosecution of rape, a crime that, with a ten percent reporting rate, is the single most under-reported major crime.[68] Increased reporting of rape is often credited to the establishment of sex crime units by law enforcement agencies, survivor advocates in rape crisis centers, survivor specialists in prosecutor's offices, and survivor counselors in hospitals.[69]

---

[64] Morse, *supra* note 9, at 745–47.

[65] *Id.* at 747.

[66] Attorney General's *Amici Curiae* Brief for the Petitioner at 13, Commonwealth v. Rape Crisis Services of Greater Lowell, Inc., 617 N.E.2d 635 (Mass. 1993) & Commonwealth v. Rape Crisis Program of Worcester, Inc., 617 N.E.2d 637 (Mass. 1993) [hereinafter, *Amici* Brief].

[67] *Id.*

[68] MADIGAN & GAMBLE, *supra* note 52, at 3. The 10% reporting rate is staggering in the light of the fact that the frequency of *reported* rapes is 10 every hour, 16 attempted every hour. "Victims of Rape," *Hearings Before the House Select Comm. on Children, Youth, and Families*, 101st Cong., 2d Sess. 5 (1990) (these statistics are from the Uniform Crime Report, 1989 and the National Crime Survey, 1989).

[69] Burgess & Holmstrom, *supra* note 40, at 46.

II. The Rights of the Defendant

The criminal defense attorney typically begins the search for favorable evidence at a disadvantage against the state's vast resources and earlier investigation.[70] Under the aegis of legitimacy, the state has financial,[71] procedural,[72] political, psychological,[73] and bargaining[74] advantages. The value of those inherent advantages, however, is reduced because the prosecutor's primary duty is not to win the case but to see that justice prevails.[75] The inequality is further balanced by defendant's constitutional right to compulsory process,[76] under which he may exercise the right to bring forth evidence against him. In addition, the Sixth Amendment's Confrontation Clause[77] provides the defendant with the fundamental right to cross-examine witnesses,[78] although a face-to-face meeting is not essential.[79] These rights, which belong to the defendant, bolster public confidence in criminal verdicts by increasing the availability of information surrounding the crime.[80]

---

[70] Daniel J. Capra, Note, *Access to Exculpatory Evidence: Avoiding the* Agurs *Problems of Prosecutorial Discretion Retrospective Review*, 53 Fordham L. Rev. 391, 391 (1984). *But see* William H. Simon, *The Ethics of Criminal Defense*, 91 Mich. L. Rev. 1703, 1707 (1993) ("[T]he image of the lonely individual facing Leviathan is misleading . . . . It is more plausible to portray the typical defendant as facing a small number of harassed, overworked bureaucrats.").

[71] David Luban, *Are Criminal Defenders Different?*, 91 Mich. L. Rev. 1729, 1735–36 (1993).

[72] *Id.* at 1736–40.

[73] *Id.* at 1740–44.

[74] *Id.* at 1744.

[75] United States v. Bagley, 473 U.S. 667, 685 n.6 (1985).

[76] U.S. Const. amend. VI ("In all criminal prosecutions, the accused shall enjoy the right . . . to have compulsory process for obtaining Witness in his favor . . . .").

[77] U.S. Const. amend. VI ("In all criminal prosecutions, the accused shall enjoy the right . . . to be confronted with the witnesses against him . . . .").

[78] *See* Davis v. Alaska, 415 U.S. 308, 315 (1974) (holding that a state statute that prohibited the defendant from impeaching the witness's credibility violated the Confrontation Clause); Chambers v. Mississippi, 410 U.S. 284, 294 (1973) (holding that the right to confrontation is fundamental to the due process of law and to a fair trial); Douglas v. Alabama, 380 U.S. 415, 419 (1965) (holding that there was a constitutional violation when an accomplice could not be cross-examined because he invoked his Fifth Amendment privilege against self-incrimination); Pointer v. Texas, 380 U.S. 400, 406–08 (1965) (holding that the defendant's confrontation right was violated when he was denied an opportunity to cross-examine the witness either at the preliminary hearing or at the trial).

[79] Deborah Clark-Weintraub, Note, *The Use of Videotaped Testimony of Victims in Cases Involving Child Sexual Abuse: A Constitutional Dilemma*, 14 Hofstra L. Rev. 261, 273 (1985).

[80] *Developments*, *supra* note 54, at 1531.

A. *The Compulsory Process and the Due Process Clauses*

> The very integrity of the judicial system and public confidence in the system depend on full disclosure of all the facts, within the framework of the rules of evidence. To ensure that justice is done, *it is imperative to the function of the courts that compulsory process be available* for the production of evidence needed either by the prosecution or by the defense.[81]

The Compulsory Process Clause gives defendants the right to obtain and use potentially exculpatory evidence.[82] Based on this right, the U.S. Supreme Court has struck down state evidence rules that arbitrarily deny defendants access to favorable evidence.[83] For example, in *Washington v. Texas*,[84] the Court held that the accused's right to compulsory process for obtaining relevant and material testimony was sufficient to override a statute that disqualified "principals, accomplices, or accessories in the same crime"[85] from testifying as defense witnesses.[86]

The Compulsory Process Clause grants protection to the defendant similar to that provided by the Due Process Clause of the Fourteenth Amendment.[87] However, the contours of the Due Process Clause in addressing the fundamental fairness of trials have been more clearly defined.[88] Therefore, an examination of the Due Process Clause may be more appropriate for an analysis of a counseling privilege involving a sexual assault survivor.

The U.S. Supreme Court has articulated the requirements of due process in the area of exculpatory evidence contained in prosecutorial files. In *Brady v. Maryland*, the U.S. Supreme Court established that due process requires the prosecution to disclose

---

[81] United States v. Nixon, 418 U.S. 683, 709 (1974) (emphasis added).

[82] Weisberg, *supra* note 18, at 950. *See* Pennsylvania v. Ritchie, 480 U.S. 39, 56 (1987) (stating that criminal defendants have the right to put before a jury evidence that may bear on the determination of guilt).

[83] Weisberg, *supra* note 18, at 952. *See* Chambers v. Mississippi, 410 U.S. 284 (1973) (holding that hearsay rules cannot be mechanically applied to exclude possibly unreliable evidence).

[84] 388 U.S. 14, 23 (1967).

[85] *Id.* at 15.

[86] The Court, however, added a footnote that indicated that this ruling should not be interpreted as a disapproval of testimonial privileges. *Id.* at 23 n.21.

[87] U.S. CONST. amend. XIV ("[N]or shall any State deprive any person of life, liberty, or property, without due process of law . . . .").

[88] *Ritchie*, 480 U.S. at 56 ("Although we conclude that compulsory process provides no greater protections in this area than those afforded by due process, we need not decide today whether and how the guarantees of the Compulsory Process Clause differ from those of the Fourteenth Amendment. It is enough to conclude that on these facts, Ritchie's claims more properly are considered by reference to due process.").

evidence favorable to the accused.[89] A subsequent decision, *U.S. v. Agurs*, clarified that a prosecutor was not required to allow complete discovery of her files as a matter of routine practice.[90] Finally, *U.S. v. Bagley* limited the prosecutor's duty of disclosure to material evidence that creates a "reasonable probability that, had the evidence been disclosed to the defense, the result of the proceeding would have been different."[91] These three decisions grant the defendant the right to put before a jury evidence that might influence the determination of guilt.[92] However, the right is narrowly construed and does not require the disclosure of everything that may influence a jury.[93]

B. *The Reach of the Confrontation Clause*

> [T]he right of confrontation and cross-examination is an essential and fundamental requirement for the kind of fair trial which is this country's constitutional goal.[94]

Courts have interpreted the Confrontation Clause of the U.S. Constitution[95] as giving defendants broad rights to cross-examine adverse witnesses. In fact, evidence "shown to be relevant and likely to be significant" may override a state rule of exclusion.[96] In *United States v. Nixon*, the U.S. Supreme Court determined that the interests of the criminal justice system were sufficient to outweigh the constitutionally based executive privilege protecting certain communications.[97] Similarly, in *Davis v. Alaska*, the Court ruled that a defendant's right to confrontation was sufficiently strong to overcome a state's interest in maintaining the confidentiality of a witness' juvenile records.[98] How-

---

[89] 373 U.S. 83, 87 (1963). *But see* Moore v. Illinois, 408 U.S. 786, 795 (1972) ("We know of no constitutional requirement that the prosecution make a complete and detailed accounting to the defense of all police investigatory work on a case.").

[90] 427 U.S. 97, 109 (1976).

[91] 473 U.S. 667, 685 (1985).

[92] *Ritchie*, 480 U.S. at 56.

[93] *Agurs*, 427 U.S. at 109–10. *See* Tom Stacy, *The Search for Truth in Constitutional Criminal Procedure*, 91 COLUM. L. REV. 1369, 1393 (1991) (criticizing the construction of the Compulsory Process Clause because it enforces the defendant's right in a system that requires an inordinate amount of trust in the good faith of the prosecutor to recognize and willingly produce exculpatory evidence).

[94] Pointer v. Texas, 380 U.S. 400, 405 (1965).

[95] This right is applicable to the states through the Fourteenth Amendment. *Id.* at 403.

[96] *Two Juveniles*, 491 N.E.2d at 238.

[97] 418 U.S. 683, 711–13 (1974).

[98] 415 U.S. 308, 319 (1974).

ever, the Court in *Pennsylvania v. Ritchie* subsequently warned that a statutory privilege need not fall every time "a defendant asserts a need, prior to trial, for the protected information that might be used at trial to impeach or otherwise undermine a witness' testimony."[99]

These cases indicate that a privilege may be challenged if there is a possibility of finding exculpatory evidence or if the privileged material may cast doubt on the truthfulness of the testimony.[100] Courts are especially likely to enforce the defendant's rights if the procedural rules that restrict testimony are mechanically applied without balancing the interests of the defendant and the government.[101] Privilege is also likely to be breached if the evidence sought by the defense is important and not available from other sources.[102]

In *Pennsylvania v. Ritchie*, the U.S. Supreme Court clarified that the Confrontation Clause grants a trial right to the defendant, and not a "rule of pre-trial discovery."[103] In other words, "[t]he ability to question adverse witnesses . . . does not include the power to require the pretrial disclosure of any and all information that might be useful in contradicting unfavorable testimony."[104] The Court stated in *dicta* that the right of confrontation is satisfied if the defense attorney is not restricted in the scope of his questioning.[105]

Even the defendant's trial rights to confront and to cross-examine witnesses are not absolute. They must give way to certain competing legitimate governmental interests.[106] Such interests may

---

[99] *Ritchie*, 480 U.S. at 52.

[100] KNAPP & VANDECREEK, *supra* note 11, at 21–22. *See also* Weisberg, *supra* note 18, at 949 (concluding that the U.S. Supreme Court balances "the general principle of confrontation against other factors" when making a decision regarding "whether communications privileges should succumb to confrontation clause rights").

[101] Daniel Lowery, Note, *The Sixth Amendment, the Preclusionary Sanction, and Rape Shield Laws:* Michigan v. Lucas, 61 U. CIN. L. REV. 297, 300–01 (1992). *See, e.g.*, Chambers v. Mississippi, 410 U.S. 284 (1973); Davis v. Alaska, 415 U.S. 308 (1974); Rock v. Arkansas, 483 U.S. 44 (1987) (holding unconstitutional a *per se* rule excluding all evidence obtained from a hypnotically refreshed witness because there was no individual balancing of the state's interest against that of the defendant).

[102] KNAPP & VANDECREEK, *supra* note 11, at 22.

[103] *Ritchie*, 480 U.S. at 41, 52.

[104] *Id.* at 54.

[105] *Id.* at 53.

[106] *See* Taylor v. Illinois, 484 U.S. 400, 414–15 (1988) (explaining that "the mere invocation" of a defendant's right to offer testimony of favorable witnesses does not "automatically outweigh countervailing public interests"); Delaware v. Fensterer, 474 U.S. 15, 20 (1985) (stating that the Confrontation Clause only guarantees "an *oppor-*

include "preventing harassment of witnesses, jury prejudice, confusion of the issues, danger to witnesses, or repetitive or marginally relevant questioning."[107] Maintaining the confidentiality of the sexual assault survivor's counseling files has been described as a strong public interest.[108]

III. Four Options for the Courts

The courts and legislatures of each state must determine the proper scope of privilege granted to the counselor-patient relationship as it applies to sexual assault survivors. Three readily apparent options are the absolute privilege, the limited or conditional privilege with an *in camera* review, and the limited or conditional privilege with direct access by the defense counsel. A final option, the semi-absolute privilege, is a proposal for an absolute privilege with narrow, well-defined exceptions, similar to the attorney-client privilege.

A. *Absolute Privilege*

An absolute privilege, which in theory grants complete protection against disclosure, enhances the confidence with which sexual assault survivors can confide in their counselors. Further, granting an absolute privilege avoids placing the sexual assault counselor in a "cruel trilemma," posing three undesirable choices: (1) to violate the extraordinary trust imposed upon them by their clients and profession; (2) to lie, and thereby commit perjury; or (3) to refuse to testify and thereby be held in contempt of court.[109] In order to avoid delivering confidential records to court, sexual assault counselors have been known to lie, to plead sudden "amnesia," or to keep two sets of records, one set for counseling and another for a possible subpoena.[110] Some rape counseling centers have been held in contempt of court and fined

---

tunity for effective cross-examination, not cross-examination that is effective in whatever way, and to whatever extent, the defense might wish.") (emphasis in original).

[107] Lowery, *supra* note 101, at 297.

[108] *Ritchie*, 480 U.S. at 57.

[109] William W. Hague, *The Psychotherapist-Patient Privilege in Washington: Extending the Privilege to Community Mental Health Clinics*, 58 Wash. L. Rev. 565, 572 (1983).

[110] *Id.*

daily because they refused to release confidential documents to the court.[111]

Although an absolute privilege would provide the maximum protection for the confidentiality of counseling files, it has potential downfalls. Because all nonconstitutionally based, absolute privileges must yield to a defendant's significant constitutional considerations,[112] labeling the privilege as absolute may mislead the sexual assault survivor into a false sense of security about the confidentiality of her communications. Further, an absolute privilege that blankets all communications between a counselor and a sexual assault survivor may eliminate some necessary exceptions. For example, under an absolute privilege, the counselor could not warn a third party, should the sexual assault survivor present an immediate physical threat.

B. *Limited Privilege*

Under a limited or conditional privilege, patients have the right to prevent the disclosure of any confidential communications made in counseling. The limited privilege can be breached if there is a sufficiently strong countervailing interest. The limited privilege has been perceived as a workable solution because it presents a compromise between the privacy interest of the sexual assault survivor and the defendant's right to confrontation and compulsory process. Such a privilege, however, does not adequately take into account the fact that the routine, or even occasional, breach of counseling records may discourage sexual assault survivors from utilizing counseling services.[113]

Once survivors fully realize that their communications enjoy only limited protection, they may either forego treatment altogether or be less frank in their communications. Either reaction will prevent defense counsel from gleaning the survivor's private thoughts through her counseling files. In such a case, the valu-

---

[111] *See, e.g.*, Commonwealth v. Rape Crisis Services of Greater Lowell, Inc., 617 N.E.2d 635 (Mass. 1993); Commonwealth v. Rape Crisis Program of Worcester, Inc., 617 N.E.2d 637 (Mass. 1993).

[112] Morse, *supra* note 9, at 753. *See also Two Juveniles*, 491 N.E.2d at 238 ("We think it clear that in certain circumstances the absolute privilege expressed in § 20J, a non-constitutionally based testimonial privilege, must yield at trial to the constitutional right of a criminal defendant to have access to privileged communications.").

[113] Neuhauser, *supra* note 36, at 186–87; *See also infra* part IV.A.4, notes 177–187 and accompanying text (discussing the decline in the percentage of women seeking counseling).

able service of counseling the sexual assault survivor back to mental health will be lost without any significant gain for the confrontation rights of the defendant.

If conditional privilege is the solution, as it currently is in Massachusetts,[114] a limited privilege with an *in camera* inspection is preferable to a limited privilege in which the defense attorney can obtain direct access to the survivor's counseling records, because an *in camera* review is less invasive of the survivor's privacy than a direct examination by the defense counsel. The former type of limited privilege would allow an *in camera* inspection upon a showing by the defense counsel that the records may contain exculpatory evidence.[115] Some have argued that the trial judge is incapable of searching through the records for exculpatory evidence with an advocate's eye.[116] Trial judges, however, have long performed this function in an analogous situation in which the defense attorney suspects that the prosecution is hiding materially favorable evidence in its files.[117] In such a context, the trial judge, rather than the prosecutor or the defense attorney, conducts an *in camera* inspection for exculpatory evidence.[118]

A limited privilege in which the defense counsel can have direct access to confidential counseling files is the least desirable option. Although this option fully satisfies the confrontation and compulsory process rights of the defendant, it does so at too high a cost to the sexual assault survivor. The defense attorney, who is in the best position to identify exculpatory evidence, has direct access to the survivor's entire counseling files. However, the survivor's statements, made in an atmosphere of complete

---

[114] *See infra* part IV.

[115] It is difficult to imagine, however, how the defense attorney can make this showing without a significant knowledge of what is already in the files. *See* Capra, *supra* note 70, at 424.

[116] *See infra* part IV.A.2, notes 158–170 and accompanying text.

[117] *See* Maryland v. Brady, 373 U.S. 83 (1963) (holding that the trial judge can conduct an *in camera* review of the prosecution's files if either (1) the prosecution is in doubt as to whether certain evidence is material or (2) if the defense can make a preliminary showing that the prosecution has materially favorable evidence in its custody).

[118] Capra, *supra* note 70, at 423. *See, e.g.*, Commonwealth v. Collett, 439 N.E.2d 1223, 1231 (Mass. 1982) (trial court judge conducted an *in camera* inspection to determine whether possibly privileged information should be disclosed to an adverse party); Commonwealth v. Tucceri, 589 N.E.2d 1216, 1222 (Mass. 1992) (trial judge ascertained what effect an omission of certain exculpatory evidence would have had on a jury); Commonwealth v. White, 565 N.E.2d 1185, 1189–92 (Mass. 1991) (trial judge ascertained whether certain evidence not offered at a rape trial was so crucial that failure to introduce it amounted to ineffective assistance of counsel).

trust and understanding, is exposed to an opposing advocate's critical eye. Some have claimed that the survivor's privacy interest will still be protected because the attorney will view the records as an officer of the court.[119] The validity of this claim will be examined in the following section.

While the defense attorney is called upon to act as an officer of the court, however, she is also expected to advocate zealously for her client at all times. The *ABA Code of Professional Responsibility*, which was later replaced with the *Model Rules of Professional Conduct*, originated the term "zealous advocacy" in Canon 7: "a lawyer should represent a client zealously within the bounds of the law."[120] As a zealous advocate, the lawyer must maximize the likelihood that her client will prevail within legal limits.[121] Some have even gone as far as to suggest that the attorney is not accountable either morally, professionally, or legally for the means by which success is achieved.[122] Such an interpretation of the duty of zealous advocacy, puts pressure on attorneys to "toe the line of illegality" in order to build a reputation as aggressive and successful litigators.[123] This is especially true for those attorneys engaged in criminal defense work where reputation is crucial to attracting future clients.[124]

The ambiguous language of the *Model Code of Professional Responsibility*, which provided that a lawyer "shall not intentionally . . . fail to seek the lawful objectives of his client through reasonably available means permitted by law,"[125] encouraged the belief that attorneys have an ethical obligation to push the lines of legality for the benefit of her client.[126] On the other hand, the *Code* added that a lawyer could "refuse to . . . participate in

---

[119] *Stockhammer*, 570 N.E.2d at 1002.

[120] MODEL RULES OF PROFESSIONAL CONDUCT Rule 1.3 (1993) (citing MODEL CODE OF PROFESSIONAL RESPONSIBILITY Canon 7). *See also* MODEL CODE OF PROFESSIONAL RESPONSIBILITY EC 7-1 (1980) ("As advocate, a lawyer *zealously* asserts the client's position under the rules of the adversary system.") (emphasis added).

[121] Eugene R. Gaetke, *Lawyers as Officers of the Court*, 42 VAND. L. REV. 39, 40 (1989) (citing Murray L. Schwartz, *The Professionalism and Accountability of Lawyers*, 66 CAL. L. REV. 669, 673 (1978)).

[122] *Id.*

[123] Bruce A. Green, *Zealous Representation Bound: The Intersection of the Ethical Codes and the Criminal Law*, 69 N.C. L. REV. 687, 709 (1991).

[124] *Id.* at 711.

[125] MODEL RULES OF PROFESSIONAL CONDUCT Rule 1.3 (1993) (citing MODEL CODE OF PROFESSIONAL RESPONSIBILITY DR 7-101(A)(1)).

[126] Green, *supra* note 123, at 709.

conduct that he believes to be unlawful, even though there is some support for an argument that the conduct is legal."[127]

The attorney's role as an officer of the court is intended to counterbalance her role as a zealous advocate. The Preamble to the current *Model Rules of Professional Conduct* states: "[a] lawyer is a representative of clients, an *officer of the legal system* and a public citizen having special responsibility for the quality of justice."[128] The *Model Rules* explicitly set out the requirements for fairness to opposing and third parties.[129] Specifically, the defense attorney, as an officer of the court, should not unnecessarily embarrass or burden third parties.[130]

In the current judicial system, however, the role of the zealous advocate is stressed over the role of the officer of the court, and attorneys often fail to adequately consider third party interests.[131] In a rape trial, a defense attorney, who has an ethical obligation to be a zealous advocate for her client, may not adequately take into consideration the privacy rights of the sexual assault survivor when examining the survivor's confidential files for exculpatory evidence. Although the trial judge may prohibit the defense counsel from introducing in evidence certain information obtained from the counseling files, the attorney is not prevented from tracking down independent sources based on leads obtained from the files.[132] The defense counsel can also use the information gained from inspection to structure the content of the cross-examination.[133]

---

[127] MODEL CODE OF PROFESSIONAL RESPONSIBILITY DR 7-101 (1980).

[128] MODEL RULES OF PROFESSIONAL CONDUCT Preamble (1993) (emphasis added).

[129] Catherine J. Lanctot, *The Duty of Zealous Advocacy and the Ethics of the Federal Government Lawyer: The Three Hardest Questions*, 54 S. CAL. L. REV. 951, 966–67 (1990–91) (For example, a lawyer may not file suit to merely harass or maliciously injure another. Further, a lawyer may not misstate issues of fact or law and must deal fairly with unrepresented third parties.).

[130] MODEL RULES OF PROFESSIONAL CONDUCT Rule 4.4 (1993).

[131] Paul L. Haines, Note, *Restraining the Overly Zealous Advocate: Time for Judicial Intervention*, 65 IND. L.J. 445, 448 (1990). *See also* Gaetke, *supra* note 121, at 87 ("The roles of zealous advocate and officer of the court conflict on [the issue of clients' interest], for an expansion of one role necessitates a restriction on the other."); David B. Wilkins, *Who Should Regulate Lawyers?*, 105 HARV. L. REV. 799, 815 (1992) (defining the distinct professional duties generated by the roles of the zealous advocate and an officer of the court).

[132] Interview with a Massachusetts superior court judge who wishes to remain unidentified (May 10, 1993).

[133] Diane M. Kottmyer & Martin F. Murphy, *Developments in Criminal Law: The Changing Face of Rape Prosecutions*, BOSTON B.J. 31, 31–32 (May/June 1992).

In states where a sexual assault survivor's counseling records can be accessed directly by the defense counsel, courts should use their inherent powers to serve as "governor[s] of the trial for the purpose of assuring its proper conduct"[134] to increase supervision of defense attorneys.[135] Close supervision by the trial courts will encourage defense attorneys to take seriously their roles as officers of the court. In this way, courts maximize the protection of the sexual assault survivor's privacy.[136]

C. *Semi-Absolute Privilege*

Establishing a semi-absolute privilege that grants complete confidentiality in all except certain, enumerated, extraordinary circumstances will prevent the harm caused by foregone counseling. The scope of the semi-absolute privilege for communications between a sexual assault survivor and her counselor can be similar to the scope of the attorney-client privilege.[137] In other words, confidentiality is *guaranteed* except in four specified circumstances. First, confidentiality can be breached when the communication involves a future criminal act likely to cause substantial harm to either the survivor or to a third party. Confidentiality can also be breached when the patient uses the counselor in furtherance of a criminal act or when the counselor has to defend herself in a suit for malpractice. Finally, the privilege can be pierced if the complainant waives the privilege or introduces any part of the counseling records in evidence.

The term, "semi-absolute privilege," gives adequate warning to the survivor that communications may be disclosed in certain circumstances. If those specific circumstances are clearly outlined, and if the sexual assault survivor is further reassured that confidentiality will not be breached except in those circumstances, the survivor is likely to accept counseling with full confidence.

---

[134] Haines, *supra* note 131, at 463 (quoting Quercia v. United States, 289 U.S. 466, 469 (1933)).

[135] Massachusetts trial judges are encouraged to take a more active role in protecting confidential records. *See infra* part IV.A.5 & IV.A.6, notes 188–211 and accompanying text.

[136] Haines, *supra* note 131, at 462 (arguing that an increased role of the judiciary in controlling its officers will insure that truth and justice will be the products of the adversary system).

[137] For a more complete discussion of the semi-absolute privilege, *see infra* part IV.D, notes 250–277 and accompanying text.

The semi-absolute privilege closes off one avenue of discovery for the defense attorney; she will not be able to search the survivor's counseling files for bias, motive to lie, and inconsistent statements. The privilege, however, does not limit opportunities during pre-trial discovery and cross-examination at trial to probe the complainant, her family, and her friends for such exculpatory evidence. Therefore, the semi-absolute privilege does not violate the defendant's constitutional rights to confrontation and compulsory process. When one weighs the *mere possibility* of finding exculpatory evidence in confidential counseling records against the *certainty* of the harm that survivors of sexual assault will suffer, the balance tips towards establishing the semi-absolute privilege.

IV. THE DEVELOPMENT OF THE COUNSELOR-PATIENT PRIVILEGE IN MASSACHUSETTS

The legislature and courts of Massachusetts have considered the first three options in their efforts to define the proper scope of privilege as it applies to communications between a sexual assault survivor and her counselor. Supporters of defendant's rights and survivor advocacy groups have lined up on either side of the political battle in an attempt to influence the scope of protection granted to the sexual assault survivor.

The criminal defense advocates rely on the Confrontation Clause of the Sixth Amendment and its Massachusetts equivalent, article 12 of the Massachusetts Declaration of Rights,[138] to urge the courts to release the sexual assault survivor's counseling records for full review by the defense counsel. They argue that an *in camera* inspection by the judge does not satisfy the defendant's confrontation rights because the trial judge, as a neutral party, is incapable of stepping into the shoes of an advocate to search for exculpatory evidence in counseling files.[139]

---

[138] MASS. CONST. pt. 1, art. XII (1990) ("And every subject shall have a right to produce all proofs, that may be favorable to him; to meet the witnesses against him face to face, and to be fully heard in his defense by himself, or his counsel, at his election.").

[139] *See, e.g.*, Commonwealth v. Fitzgerald, 590 N.E.2d 1151 (Mass. 1992) (allowing the defense counsel to introduce complainant's counseling statement from counselor's notes to support the theory that the complainant fabricated the rape accusation in order to avoid parental disapproval of her conduct); Commonwealth v. O'Brien, 536 N.E.2d 361 (Ma. App. 1989), *review denied*, 542 N.E.2d 601 (Mass. 1989) (stating that

On the other side, sexual assault survivors and their advocates emphasize the important role that counseling plays in the recovery process. They argue that a guarantee of confidentiality is absolutely necessary for effective counseling[140] and that routine disclosure of confidential records will deter survivors from seeking help.[141] Survivor advocates argue that the certain social benefit derived from counseling outweighs the possible benefit that may be gained from exposing the counseling records to the defense counsel.

A. *The Massachusetts Case Law*

1. *Commonwealth v. Two Juveniles* Pierced an Absolute Privilege by Allowing an *In Camera* Review

In the 1986 case, *Commonwealth v. Two Juveniles*,[142] the Massachusetts Supreme Judicial Court (SJC) established a precedent for limiting the protection of the sexual assault survivor's privacy. *Two Juveniles* presented the court with an opportunity to rule on the constitutionality of the state's statutory sexual assault counselor privilege, that was enacted as an absolute privilege.[143] The SJC passed on this opportunity, refusing to rule on the constitutionality of the absolute privilege in the abstract.[144] Nevertheless, the SJC breached the absolute privilege by allowing the defendant to request an *in camera* inspection of the counseling records.[145] The SJC justified the use of an *in camera* inspection by stating that:

---

inconsistencies, however minor, between the complainant's statements made at trial and statements found in the social worker's files might have affected the survivor's credibility); Commonwealth v. Fayerweather, 546 N.E.2d 345 (Mass. 1989) (holding that the exclusion of the complainant's psychiatric report prejudiced the defendant. The report indicated that the complainant had claimed to hear the defendant's voice telling her to do things six weeks prior to the sexual assault charge.).

[140] *See supra* parts I.C & I.D, notes 36–69 and accompanying text.

[141] Commonwealth v. Collett, 439 N.E.2d 1223, 1226 (Mass. 1982) ("If it becomes known that confidences are violated, other people may be reluctant to use social work services and may be unable to use them to maximum benefit.").

[142] 491 N.E.2d 234 (Mass. 1986).

[143] MASS. GEN. L. ch. 233, § 20J (1992) ("A sexual assault counselor shall not disclose such confidential communication, without the prior written consent of the victim . . . . Such confidential communications shall not be subject to discovery and shall be inadmissible in any criminal or civil proceeding without the prior written consent of the victim to whom the report, record, working paper or memorandum relates.").

[144] *Two Juveniles*, 491 N.E.2d at 237.

[145] *Id.* ("Use of the device of an *in camera* inspection would derive not from an

> in certain circumstances the absolute privilege expressed in § 20J, a nonconstitutionally based testimonial privilege, must yield at trial to the constitutional right of a criminal defendant to have access to privileged communications.[146]

Prior to an *in camera* inspection, however, the SJC required the defendant to demonstrate a legitimate need for the survivor's privileged records.[147] The SJC established that after the defense counsel made a successful showing of legitimate need, the trial court, in its discretion, could conduct an *in camera* inspection to determine which part of the record contained helpful information to the defense.[148]

To define "legitimate need," the SJC offered several examples of what would not constitute legitimate need but refrained from outlining what would be sufficient. In demonstrating "legitimate need," the defense counsel had to assert more than a mere possibility that the confidential records contained a possible attack on credibility.[149] Further, the defense counsel could not simply declare that the "very circumstances of the communications indicate that they are likely to be relevant and material to the case."[150] Even the fact that the information is not available from any other source was declared insufficient to establish legitimate need.[151]

Because the SJC failed to give a concrete definition of "legitimate need," the lower courts were left to decide independently whether such a showing had been made. In addition, the court failed to address how the defense attorney could make a showing of "legitimate need" without prior access to the files. The SJC created the *in camera* standard without fully considering either the significant interest of the survivor in maintaining the abso-

---

interpretation of § 20J but rather from a determination that the juveniles have a constitutional right which transcends the statute and requires the courts to fashion an exception to the statute.").

[146] *Id.* at 238.

[147] *Id.* at 239.

[148] *Id.* at 239–40.

[149] *Id.* at 239. Other jurisdictions have different requirements for establishing a greater showing of necessity. *See, e.g.*, State v. Siel, 444 A.2d 499 (N.H. 1982) (holding that information sought solely to show prior inconsistent statement was not material); People v. Pena, 487 N.Y.S.2d 935 (1985) (denying access to rape crisis center records because the defendant failed to make a specific showing that the records contained exculpatory material); United States v. Valenzuela-Bernal, 458 U.S. 858 (1982) (holding that the respondent must at least show how the testimony he was prevented from obtaining would have been "both material and favorable to his defense" in order to establish a violation of the Compulsory Process Clause of the Sixth Amendment).

[150] *Two Juveniles*, 491 N.E.2d. at 239.

[151] *Id.*

lute confidentiality of her counseling records[152] or the societal harms that will be caused by decreased reporting of crime and increased mental health problems of the survivors.[153]

Even though the rights of the survivor's privacy had been significantly compromised, some defense advocates complained that the *Two Juveniles* standard created a "double hurdle" for the defendant. First, the defendant was required to make a preliminary showing for the court to exercise its *in camera* inspection power over privileged communications. Second, the court alone had the discretion to decide which materials, if any, would be admitted into evidence.[154] Perhaps in response to this criticism, the SJC in *Commonwealth v. Clancy*,[155] which applied the *Two Juveniles* standard to the limited psychotherapist-patient privilege, hinted that it might have reached a more pro-defendant result if the accused had claimed that his state, rather than federal, constitutional right to confrontation had been violated.[156] The SJC had an opportunity to address this issue in *Commonwealth v. Stockhammer*.[157]

2. *Commonwealth v. Stockhammer* Allows Defense Counsel Direct Access to Couseling Records

In *Stockhammer*, the SJC held that a trial judge's *in camera* inspection of a survivor's conditionally privileged psychological records did not adequately satisfy the defendant's state right to confrontation. In order to satisfy the state confrontation right, the court gave the defense counsel direct access to the survivor's counseling files to search for exculpatory evidence.[158]

Without explanation, the SJC interpreted the state constitutional right to confrontation more broadly than the U.S. Supreme Court's interpretation of the Confrontation Clause in *Pennsylvania v. Ritchie*.[159] In *Ritchie*, the U.S. Supreme Court concluded

---

[152] *See* Neuhauser, *supra* note 36, at 196 (advocating that an *in camera* review is an unacceptable breach of privacy. "[I]t is the mere knowledge of private, personal matters by another which is offensive . . . .").

[153] *See supra* part I.D, notes 52–69 and accompanying text.

[154] Kathryn A. O'Leary, Case Comment, *Evidence—Defendant's Sixth Amendment Right to Confrontation Becomes Discretionary under Sexual Assault Counselor-Victim Privilege*—Commonwealth v. Two Juveniles, 21 SUFFOLK U. L. REV. 1222 (1987).

[155] 524 N.E.2d 395 (Mass. 1988).

[156] *Id.* at 399.

[157] 570 N.E.2d 992 (Mass. 1991).

[158] *Id. at 1002.*

[159] 480 U.S. 39 (1987).

that a trial judge was capable of acting as both an impartial adjudicator during the trial and as an advocate while conducting an *in camera* inspection of the survivor's treatment records.[160] In contrast, the SJC believed that a trial judge could not effectively act as both an adjudicator and an advocate because "the trial court's judgment as to the utility of material for impeachment . . . would [not] exhaust the possibilities."[161]

The SJC rejected the assertion that an *in camera* inspection was the only effective means of protecting the survivor's privacy. The court suggested alternate means of protection such as allowing defense counsel access to privileged records only in their capacity as officers of the court,[162] issuing protective orders to prevent further disclosure of the record,[163] and conditioning the admission of the information in evidence upon a determination that the information is relevant, nonprejudicial, and not available from any other source.[164]

The SJC reasoned that a direct examination of counseling records by the defense attorney "need not be any more intrusive or harmful than those attending *in camera* review of the records by the judge alone."[165] The SJC, however, failed to recognize the significant difference between an inspection by a trial judge, acting in the capacity of a neutral fact-finder, and an examination by the defense counsel, whose duty it is to challenge the credibility of the witness.

Stockhammer did not overrule *Two Juveniles* and was distinguished as a case involving a limited rather than an absolute privilege. The court reasoned that the psychologist[166] and the social worker[167] privileges involved in *Stockhammer* contained specifically carved out exceptions and derived from a "less firmly based legislative concern."[168] Although *Stockhammer* did not ex-

---

[160] *Id.*

[161] *Stockhammer*, 570 N.E.2d at 1001–02 ("[W]hen a judge undertakes to decide if [evidence] benefits the defendant's case he is assuming vicariously and uncomfortably the role of counsel.") (quoting Commonwealth v. Liebman, 446 N.E.2d 714, 717–18 (Mass. 1983)). *See also* Commonwealth v. Clancy, 524 N.E.2d 395 (Mass. 1988) (holding that an *in camera* review confuses the roles of the trial judge and the defense counsel).

[162] *See supra* part III.B, notes 113–136 and accompanying text.

[163] *Id.*

[164] *Id.*

[165] *Id.*

[166] MASS. GEN. L. ch. 233, § 20B (1992).

[167] MASS. GEN. L. ch. 112, § 135 (1992).

[168] *Stockhammer*, 570 N.E.2d at 1002 (quoting *Two Juveniles*, 491 N.E.2d at 234).

plicitly eliminate the *Two Juveniles* "legitimate need" requirement to access privileged records, the requirement, in practice, became less strict. After *Stockhammer*, very few barriers remained between the defense attorney and the sexual assault survivor's psychological, psychiatric, and medical records.[169] In fact, the defense counsel was often able to gain direct access to counseling records without making a showing of absolute necessity and without a prior *in camera* inspection for relevancy.[170]

3. *Commonwealth v. Figueroa* Abolished the Requirement for a Showing of Legitimate Need

A year later, *Commonwealth v. Figueroa*[171] further eroded the sexual assault survivor's right to privacy. The SJC affirmed *Stockhammer* and allowed the defense counsel direct access to conditionally privileged counseling records. By examining the counseling records, the defense counsel could search for bias, prejudice, motive to lie, and could evaluate the mental capacity of the survivor.[172] In addition, *Figueroa* took *Stockhammer* one step farther by explicitly abolishing the requirement for a showing of legitimate need.[173] The court stated:

> [s]ince this is a sexual assault case in which the credibility of the alleged victim is critical, such review must be made

---

[169] These records can be extensive according to Veronica Reed Ryback, director of the Rape Crisis Intervention Program at Beth Israel Hospital in Boston, Mass., and a clinical instructor in psychiatry at Harvard Medical School. For example, the records may include "past psychiatric history, history of drugs and alcohol, past trauma history, past sexual abuse; while at the same time [they] note everything that the [sexual assault survivor] is talking about." Telephone interview with Veronica Reed Ryback (May 13, 1993) [hereinafter "Ryback Conversation"]. The extensive record is necessary so that the counselor can determine if the survivor is at risk of committing suicide. Sherman, *supra* note 48, at 1, 27. *See also* David Lauter, *Assault on Shelter Data Heats Up*, NAT'L L.J., Feb. 28, 1983, at 5, (quoting Peggy Langhammer of the Rhode Island Rape Crisis Center in Providence, "The kinds of information we have are the intimate conversations which go on between a victim [and a therapist.]").

[170] Sherman, *supra* note 48, at 1, 27.

[171] 595 N.E.2d 779 (Mass. 1992).

[172] The privileged records included records protected by the psychotherapist/patient privilege, MASS. GEN. L. ch. 233, § 20B (1992) and the social worker/client privilege, MASS. GEN. L. ch. 112, § 135A (1992).

[173] This interpretation was later challenged by the Attorney General's office in its *amici* brief, *supra* note 66. The Attorney General argued that the *Figueroa* decision was not meant to apply to all survivors of sexual assaults. Rather, the statement was fact-specific to the *Figueroa* case where the credibility of a mentally retarded person was at issue. The Attorney General maintained that the defendant was still required to demonstrate a particularized need for the survivor's records before the records can be disclosed.

> available to counsel without any prior showing by the defendant of special circumstances demonstrating a particularized need for access to communications.[174]

The survivor's privacy continued to be protected by the trial judge's discretion in issuing rules and orders regulating the use of the records at trial.[175] The trial judge's rules and orders, however, varied with the individual judge and did not provide consistent protection for each survivor. A divisional counsel for the Department of Social Services observed that some judges have "thrown open the door to all his department's records" in allowing the defense counsel access to the records.[176]

4. The Aftermath of *Stockhammer* and *Figueroa*

After *Stockhammer* and *Figueroa*, defense attorneys began to file disclosure motions routinely, hoping to obtain contradictory statements made in counseling records.[177] Such statements could be used to harass the complainant. With very few exceptions,[178] Massachusetts courts interpreted the *Stockhammer-Figueroa* decisions as giving defense attorneys broad access to confidential counseling records protected by limited privilege.[179] In fact, "[*Stockhammer* became] a symbol for the notion that any party to any

---

[174] *Figueroa*, 595 N.E.2d at 785.

[175] *Id.* at 781.

[176] Sherman, *supra* note 48, at 27.

[177] *Id. See also* Lauter, *supra* note 169, at 5. According to Joanne Schulman of the New York-based National Center on Women and Family Law, "Defense attorneys are learning from rape cases that [requesting counseling records] is a pretty effective harassment tool." *Id.* Peggy Langhammer from the Rhode Island Rape Crisis Center states, "Frequently a traumatized victim will say things like 'I feel guilty' or 'I think maybe I asked for it,' or 'I want to get back at him,' and those things a defense attorney could have a field day with." *Id.*

[178] *See, e.g., Commonwealth v. Jones*, 615 N.E.2d 207 (Mass. App. 1993) (refusing to allow defendant access to the sexual assault survivor's counseling records because the treatment records did not relate to the rape).

[179] *See* Commonwealth v. Gauthier, 586 N.E.2d 34 (Mass. App. Ct. 1992), *review denied*, 597 N.E.2d 1371 (1992) (allowing access to pre-incident, special education records although the records were deemed to be in the "semi-absolute or limited" privilege). *See also* Commonwealth v. Hrycenko, 578 N.E.2d 809 (Mass. App. Ct. 1991), *rev'd on other grounds*, 630 N.E.2d 258 (Mass. 1994) (granting access to sexual assault survivor's mental health treatment records from 1985 through 1988, despite the fact that the rape occurred in September 1988); Commonwealth v. Arthur, 575 N.E.2d 1147 (Mass. App. Ct. 1991) (allowing access to victim's Department of Social Services records); Commonwealth v. Simcock, 575 N.E.2d 1137 (Mass. App. Ct. 1991), *review denied*, 579 N.E.2d 1360 (Mass. 1991) (granting defense counsel access to pre-incident medical records that were "potentially relevant" in a case involving indecent assault and battery).

litigation has full license to access and search materials which are protected by statutory or common law privileges."[180] Nancy Gertner, who defended Jonathan Stockhammer, commented that the SJC had gone too far by not sufficiently emphasizing that access to counseling records be given only in extraordinary cases.[181] She also added that the SJC should have stressed the continual role that trial judges must play in guarding the survivor's privacy.[182]

Following *Stockhammer* and *Figueroa*, rape counseling clinics in Massachusetts reported a decline in the percentage of women who sought counseling.[183] Sexual assault survivors were apparently less willing to seek counseling because they feared that the counseling records could later be exposed to a demanding court.[184] Some feared that those survivors who agreed to counseling did so without understanding the full implications of the change in the scope of privilege accorded to their communication with their counselors.[185] Even if the sexual assault survivor agreed to counseling, she became less likely to pursue legal action once she realized that her counseling records may be revealed in court.[186] It was likely that the survivor viewed foregoing legal action as a tradeoff for receiving effective counseling treatment.[187]

5. Judicial Attempts to Protect the Survivor's Privacy

Stockhammer assigned significant responsibilities to judges to monitor confidential records throughout the trial process.[188] Accept-

---

[180] *Amici* Brief, *supra* note 66, at 45.

[181] Kevin Cullen, *Liberals Split on Issues in Rape Cases*, BOSTON GLOBE, Mar. 7, 1993, at 29.

[182] *Id.*

[183] In her remarks at the Educational Conference of Superior Court Justices, Braintree, Massachusetts (May 8, 1992), Veronica Reed Ryback reported that when informed of the consequences of *Stockhammer*, 30% of the survivors raised concerns about counseling, avoiding full disclosure about their personal history, and 10% refused counseling outright. In addition, there has been a 20% drop in clients reporting to the police. *See also Amici* Brief, *supra* note 66, at 11.

[184] Ryback Conversation, *supra* note 169. *See also* Stephen A. Saltzburg, *Privileges and Professionals: Lawyers and Psychiatrists*, 66 VA. L. REV. 597, 620 (1980) (It is generally recognized that "[p]sychiatric communications are uniquely sensitive, and successful treatment requires a degree of self-revelation by the patient which can only be accomplished in an atmosphere of inviolate privacy.").

[185] *See* Ryback Conversation, *supra* note 169.

[186] *Id.*

[187] *Id.*

[188] *Stockhammer*, 570 N.E.2d at 1002–03. On remand, trial judges were charged with the responsibility of conducting an *in camera* hearing to determine the admissibility of

ing this responsibility, Massachusetts trial judges independently began to issue written protective orders to alleviate the combined effects of *Stockhammer* and *Figueroa*. For example, Judge Peter Lauriat of the Middlesex Superior Court in Cambridge formulated a model order that incorporated many of the suggestions outlined in *Stockhammer*.[189]

Judge Lauriat's order allowed defense counsel to view the survivor's records only during specified times and only with a court clerk present. In addition, the defense counsel was forbidden from photocopying any part of the records or from sharing the information with the defendant without express judicial approval. Finally, the defense counsel was restricted from offering into evidence any portion of the treatment records without prior approval of the court. In order to introduce any part of the record, counsel had to show that the information was not available from any other source. The model order was essentially the same for both civil and criminal proceedings.

Similarly, Judge John Cratsley wrote a disclosure order that required the defense counsel to make an *in camera* showing "that the information to be offered is relevant, that its relevance outweighs any prejudice to the victim, and that the evidence is not available from any other source" prior to trial.[190] Still another superior court judge, R. Malcolm Graham, specifically protected sexual assault counselor records from disclosure obtained without prior written consent of the survivor.

6. *Commonwealth v. Bishop* Reestablishes the *In Camera* Inspection for Conditionally Privileged Records

In 1993, *Commonwealth v. Bishop*[191] attempted to strike the delicate balance between an overly broad privilege, which would allow defense counsel access to virtually all of the survivor's records, and a narrow result, which would bar the defense from access to obviously exculpatory material.[192] Recognizing that "victims of rape or sexual abuse would likely shy away from forth-

---

the information in the records for use at trial. Further, they had to issue appropriate orders to protect the survivor's privacy. *Id.*

[189] *Id.*

[190] *Stockhammer* Order issued May 11, 1993 by Judge John Cratsley in *Commonwealth v. Doherty*, Indictment # 93414-16.

[191] 617 N.E.2d 990 (Mass. 1993).

[192] *Id.* at 996.

right therapeutic sessions with their counselors if their words were later lent to the perpetrator in aid of his or her defense,"[193] the SJC established the *Two Juveniles* "legitimate need" requirement for conditionally privileged records. According to *Bishop*, a defendant must demonstrate a likelihood that a conditionally privileged record contains relevant evidence in order to justify initial judicial review.[194]

In an effort to clarify the confusion caused by the *Stockhammer* and *Figueroa* decisions, the SJC outlined a five-stage procedure for trial judges to follow in dealing with privileged psychological records of sexual assault survivors. In the first stage, the defendant must move to compel production of the survivor's counseling records, and the judge must decide whether the records are indeed privileged.[195] During the second stage, the defense counsel must submit theories under which the records are likely to be relevant.[196] If the judge decides that the defendant's theories have merit, the judge will conduct an *in camera* inspection to determine which portions of the record are *relevant.*[197]

Again the SJC was again imprecise about what submissions would be sufficient to prove relevancy. The court simply stated that the defense counsel must demonstrate "at least some factual basis which indicates how the privileged records are likely to be relevant to an issue in the case . . . ."[198] The defendant need not show that the record contains information with potential for casting doubt on the criminal charge, but the request shall not be "an unrestrained foray into confidential records in the hope that the unearthing of some unspecified information would enable [the defendant] to impeach the witness."[199]

The judge shall determine relevancy in an *in camera* inspection[200] by considering "the nature of the privilege claimed, the date the target records were produced relative to the date or dates of the alleged incident, and the nature of the crimes charged."[201] The SJC acknowledged the capability of trial judges to identify materials that might be *relevant* but distinguished this ability

---

[193] *Id.* at 994.
[194] *Id.* at 995.
[195] *Id.* at 997.
[196] *Id.*
[197] *Id.* at 998.
[198] *Id.*
[199] *Id.* at 997 (quoting People v. Gissendanner, 399 N.E.2d 924, 928 (1979)).
[200] *Id.*
[201] *Id.* at 997.

from the capability to identify materials *necessary to the defense.*[202]

In the third stage, both the defense counsel and the prosecutor are allowed access to the relevant portions of the records to determine whether disclosure is required for a fair trial.[203] If it is necessary to disclose any part of the counseling records, the trial judge will issue a protective order similar to the model protective order presented in the Appendix to the *Bishop* opinion.[204] The protective order ensures that the information will be disclosed only to the extent "absolutely and unavoidably necessary."[205]

According to the order, counsel may view records only in their capacity as officers of the court,[206] may not photocopy the records without prior permission from the court,[207] and may not introduce the records in evidence without a prior order from the court.[208] Finally, any copies of the records must be returned to the clerk at the conclusion of the trial.[209]

Once the defendant demonstrates that the relevant portions of the counseling records are required for a fair trial, the trial judge will allow disclosure in the fourth stage.[210] The final stage requires the trial judge to be actively involved through the *voir dire* examination, in determining the admissibility of the records in evidence.[211]

7. Problems with *Bishop*

Although the *Bishop* decision outlined a five-stage process, it did not go far enough in defining a workable standard for trial courts to implement. For example, the standard defined in the second stage, which requires the defense counsel to submit theories under which the records are likely to be relevant, is just as

---

[202] The latter standard, which was outlined in *Two Juveniles*, was rejected for the limited privilege in *Stockhammer*, *supra* part IV.A.2, notes 156–170 and accompanying text.

[203] *Bishop*, 617 N.E.2d at 998.

[204] *Id.* at 1001–02. The protective order is similar to the one outlined by Judge Lauriat following the *Stockhammer* decision, *supra* part IV.A.5, notes 188–190 and accompanying text.

[205] *Bishop*, 617 N.E.2d at 998.

[206] *Id.* at 1001.

[207] *Id.* at 1001–02.

[208] *Id.* at 1002.

[209] *Id.* at 1002.

[210] *Id.* at 998.

[211] *Id.*

vague as the *Two Juveniles* "legitimate need" standard. Again, the SJC was unclear about how the defense attorney, without a prior inspection, could show a likelihood that the records contained exculpatory evidence. At best, the showing of relevancy is likely to be a blanket inquiry with boiler-plate language, much like the requests that were at issue in *Commonwealth v. Rape Crisis Services of Greater Lowell, Inc.*[212] and *Commonwealth v. Rape Crisis Program of Worcester, Inc.*[213]

The SJC may have also inadvertently cast a wider net for materials made available to the defense attorney when it required the trial court, in an *in camera* inspection, to search for *relevant* materials rather than materials necessary to the defense. If "relevancy" were interpreted as any statement having to do with the sexual assault or any expression of the survivor's emotional reaction to the assault, virtually the entire file must be exposed to the defense attorney.

B. *Legislation*

Stymied in their judicial endeavors to guarantee absolute confidentiality for communication between a sexual assault survivor and her counselor, prosecutors and advocates have turned their efforts towards legislation.[214] On February 24, 1993, Massachusetts State Representative Peter Forman (R-Plymouth), the House Minority Leader, introduced a bill to protect the privacy rights of sexual assault and domestic violence survivors.[215] The 1993 bill changed the psychotherapist-patient privilege[216] from a limited to an absolute privilege, basing the absolute privilege on the sexual assault counselor-client privilege.[217] The 1993 bill stated in relevant part:

---

[212] 617 N.E.2d 635 (Mass. 1993). *See infra* part IV.C, notes 240–249 and accompanying text.

[213] 617 N.E.2d 637 (Mass. 1993). *See infra* part IV.C., notes 240–249 and accompanying text.

[214] McCafferty, *supra* note 36, at 139 ("[I]f a statute which wholly protects the victim from possible exposure of her personal thoughts is non-existent, there is no way for the rape crisis counselor to ensure the victim that they may privately discuss problems which have occurred due to the rape. This fact could result in instigating a chilling effect on rape prosecutions.").

[215] 1993 Mass. House Bill No. 3500, Massachusetts 178th General Court, 1993 Regular Session [hereinafter Forman Bill 1].

[216] MASS. GEN. L. ch. 233, § 20B (1992).

[217] *Id.* § 20J.

> (g) Notwithstanding the provisions of this section, if the patient is a victim of sexual assault or domestic violence, and the communications relate to said sexual assault or domestic violence, such communications shall be protected to the same extent as confidential communications under section 20J [sexual assault counselor privilege] or 20K of this chapter, as the case may be.[218]

Among the many safeguards suggested to protect the survivor's privacy, the 1993 bill included a requirement for notification of the survivor if the counseling records were expected to be subpoenaed. This requirement provided the survivor with an opportunity to object to the subpoena or to drop her charges.[219] In addition, the 1993 bill suggested similar amendments to other limited privileges that may apply to sexual assault survivors.[220] Further, the bill introduced section 20L, which explicitly defined the guidelines for defense attorneys to follow when examining confidential records.[221] Finally, the bill added a hefty fine, ranging from $25,000 to $50,000, to ensure compliance.[222]

The 1993 bill did not pass, most likely because it eradicated all of the carefully drafted exceptions to the limited privilege regulating psychologist and psychotherapist communications.[223] Even if the bill had passed, it would not have guaranteed absolute confidentiality to sexual assault survivors. The *Two Juve-*

---

[218] Forman Bill 1, *supra* note 215, at 2–3.

[219] *Id.* at 2.

> Section 19A. Prior to the issuance of a subpoena . . . , a notice shall be provided by the Commonwealth to:
> (1) any person whose privacy rights may be affected by such disclosure including the victim of the alleged crime and any third party data subjects reasonably ascertainable . . . .

[220] For example, the 1993 bill suggests the following amendment to Mass. Gen. L. ch. 112, § 135A [limited privilege for social worker-client communications]:

> Notwithstanding the provisions of this section, if the patient is a victim of sexual assault or domestic violence, and the communications relate to said sexual assault or domestic violence, such communications shall be protected to the same extent as confidential communications under section 20J or 20K of chapter 233, as the case may be.

[221] This is a codification of the protective order outlined in *Stockhammer*, 570 N.E.2d at 1002.

[222] Forman Bill 1, *supra* note 215, at 4.

[223] Numerous current exceptions to the limited privilege would be eradicated if the psychotherapist-patient privilege changed to an absolute privilege. Currently, there are exceptions to the privilege if the psychotherapist determines that the patient is a threat to himself or to another person, (MASS. GEN. L. ch. 233, § 20B(a)), if the mental or emotional condition of the patient is introduced as an element of his claim, (*Id.* § 20(B)(c)) or if the patient sues the psychotherapist for malpractice. (*Id.* § 20(B)(f)).

niles ruling has already allowed the section 20J privilege to be breached in certain circumstances.[224]

On January 5, 1994, Rep. Forman introduced another bill, which would essentially codify the *Bishop* decision. The pending 1994 Forman bill does not grant absolute privacy rights to the counseling records of sexual assault survivors but provides certain procedural safeguards.[225] For example, the bill requires a hearing to determine the scope of a subpoena for confidential materials before it can be issued. Both the survivor and any other individual with a substantial interest in the disclosure of the material must be notified of the hearing. However, the requirement for a hearing does not extend to records for which the defendant holds the privilege.[226] Following the hearing, the court must issue an order for disclosure that carefully restricts the use of the confidential materials.[227] The restrictions, however, may be modified upon "a showing by a preponderance of the evidence that such deviation is required by the United States or Massachusetts Constitutions."[228]

The most striking aspect of the 1994 bill is the fact that it compares the scope of protection granted to the section 20J sexual assault counselor privilege to that given to the section 20A clergy privilege.[229] Like section 20J, section 20A absolutely forbids a spiritual counselor from disclosing confessions or from testifying without prior consent of the communicant.[230] Unfortunately, as with the sexual assault counselor privilege, the SJC has never directly addressed the constitutionality of the clergy privilege.[231] Until the SJC upholds the constitutionality of the

---

[224] *Supra* part IV.A.1, notes 142–157 and accompanying text.

[225] 1994 Mass. House Bill No. 1527, Massachusetts 179th General Court, 1994 Regular Session [hereinafter Forman Bill 2].

[226] The proposed Mass. Gen. L. ch. 233, § 19A would read: "[i]n a criminal proceeding, prior to the issuance of a subpoena or court order for any confidential or privileged material, the court of records shall schedule a hearing to determine the extent of such subpoena or court order. Provided, however, that this section shall not apply to any materials where the privilege at issue belongs to the defendant."

[227] The disclosure order is similar to the *Bishop* model order, *supra* part IV.A.6, notes 191–211 and accompanying text.

[228] Forman Bill 2, *supra* note 225, § 1.

[229] *Id.* § 3.

[230] MASS. GEN. L. ch. 233, § 20A (1992).

[231] In Commonwealth v. Zezima, 310 N.E.2d 590, 592 (Mass. 1974), the SJC acknowledged that § 20A implied that "all conversation . . . incident[al] to a quest for religious or spiritual advice or comfort was inadmissible in the absence of the defendant's consent." The court, however, never questioned the constitutionality of the statute.

clergy privilege, the reference to the clergy privilege may not create any greater protection for the survivor's privacy interests.

The 1994 Forman bill extends the section 20J absolute privilege to licensed psychotherapists, social workers, and nurses, who serve similar functions as sexual assault counselors.[232] By broadening the scope of the definition of the sexual assault counselor to include these professionals within section 20J, the 1994 bill avoids the problems faced by the 1993 Forman bill.

The 1993 bill eradicated the carefully enacted exceptions to the psychotherapist-patient privilege by changing the psychotherapist-patient privilege from a limited privilege to an absolute privilege.[233] The 1994 bill, in contrast, changes section 20J by expanding it to include communications made to psychotherapists, social workers, and nurses if such communications relate to sexual assault. The wording of the 1994 bill may be clarified further by explicitly stating that communications to psychotherapists, social workers, and nurses are only protected by absolute privilege if made for the *primary purpose* of assisting survivors of sexual assault.[234]

An alternate bill introduced by State Representative Bruce E. Tarr (R-Gloucester) on February 1, 1994, proposes amending section 20B, the psychotherapist-patient privilege, rather than expanding section 20J absolute privilege accorded to sexual assault counselors.[235] The Tarr bill would add the following section to section 20B:

> (g) Notwithstanding the provisions of this section, if the patient is a VICTIM of SEXUAL ASSAULT or domestic violence, and the communications relate to said SEXUAL ASSAULT or domestic violence, such communications shall be protected to the same extent as confidential communications under section 20J or 20K of this chapter, as the case may be.[236]

---

[232] Forman Bill 2, *supra* note 225, § 2.

[233] *See supra* notes 216–223 and accompanying text.

[234] The amended definition would read: "'Sexual assault counselor', a person who is employed by or is a volunteer in a rape crisis center . . . and whose primary purpose is the rendering of advice, counseling, or assistance to victims of sexual assault—or a licensed psychotherapist, social worker, or nurse to whom communication is made for the primary purpose of obtaining advice, counseling, or assistance for recovery of sexual assault."

[235] 1994 Mass. House Bill No. 2638, Massachusetts 179th General Court, 1994 Regular Session [hereinafter Tarr Bill].

[236] *Id.* § 2. The following privileges would be similarly amended: Mass. Gen. L. ch. 112, § 135A, the social worker privilege; Mass. Gen. L. ch. 111, § 70E, the privilege pertaining to hospital records; and Mass. Gen. L. ch. 123B, the privilege pertaining to records maintained by the Department of Mental Retardation.

Both the Forman and the Tarr bills present possible paths for granting further protection to sexual assault survivors. While the Tarr bill bases the guarantee of absolute privilege on section 20J, the sexual assault counselor privilege, the 1994 Forman bill looks to section 20A, the clergy privilege. The SJC has never directly addressed the constitutionality of either privilege. However, it has recognized (in the context of the sexual assault counselor privilege) that "in *certain* circumstances a defendant must have access to privileged records so as not to undermine confidence in the outcome of trial."[237] Given the fact that the SJC has already allowed an *in camera* inspection of a sexual assault counselor's records, which are protected by absolute privilege,[238] it seems unlikely that it will guarantee absolute confidentiality to similar records obtained from communications with a psychotherapist, which are protected by a limited privilege.[239]

Therefore, the 1994 Forman bill, which relies upon the clergy privilege, may be a more viable path to pursuing broader rights of privacy for the sexual assault survivor.

C. *Absolute Privilege Applied to Sexual Assault Counselors*

The next battle between the defendant's right to confrontation and the survivor's privacy interests is likely to take place in the courts over the constitutionality of the legislatively created, absolute privilege for sexual assault counselors. The U.S. Supreme Court has never addressed the issue of whether an absolute privilege, such as the Massachusetts sexual assault counselor privilege, satisfies the Confrontation Clause of the U.S. Constitution.[240] Likewise, the question of whether granting an absolute privilege to communications made to sexual assault counselors satisfies the confrontation clause of the Massachusetts Declaration of Rights has yet to be addressed by the highest court in Massachusetts.[241]

---

[237] *Bishop*, 617 N.E.2d at 994–95 (citing *Two Juveniles*, 491 N.E.2d at 238).

[238] *See supra* part IV.A.1, notes 142–157 and accompanying text, for a discussion of *Two Juveniles*.

[239] For further discussion, *see infra* part IV.C, notes 240–249 and accompanying text.

[240] *Cf. Ritchie*, 480 U.S. at 57 (holding that an *in camera* review of the conditionally privileged counseling records struck the correct balance between the defendant's right to a fair trial and the privacy rights of the survivor, but remaining silent on the issue of whether an absolute privilege is constitutional).

[241] In *Two Juveniles*, 491 N.E.2d at 239, the SJC cited the *Advisory Opinion to the*

Recently, the SJC had two opportunities to address the constitutionality of the legislatively created, absolute privilege granted to sexual assault counselors. The court, however, declined the invitation to address the issue in both cases. *Commonwealth v. Rape Crisis Services of Greater Lowell, Inc.*[242] involved a trial court's subpoena for a sexual assault survivor's treatment records maintained at the Rape Crisis Services of Greater Lowell, Inc. (RCS of Lowell). The boiler-plate language in the affidavit in support of the subpoena failed to specifically assert a need for compelling disclosure, but rather made a blanket inquiry. The relevant part stated,

> The documents which are requested flow from the alleged incident. As such they clearly contain statements of the alleged victim which relate to the incident. These statements were presumably made shortly after the incident. For these reasons [the] defense counsel is entitled to review these privileged records to search for evidence of the complainant's bias, prejudice or motive to lie.[243]

RCS of Lowell refused to disclose the records and was held in civil contempt and fined daily until compliance.[244] The defendant subsequently pled guilty and was sentenced, and the judge vacated the contempt order.[245]

On appeal, the SJC dismissed the case and refused to consider the issue of whether the records of sexual assault counselors were protected by absolute privilege. The court stated, "[Defendant's] plea and the judge's order vacating the civil contempt order render this appeal moot."[246] The court rejected the rape crisis center's argument that the stigmatization associated with the civil contempt citation is of such magnitude as to warrant appellate review.[247]

The SJC again passed on the opportunity to rule on the constitutionality of the section 20J privilege in *Commonwealth v. Rape Crisis Program of Worcester, Inc.*[248] The *Worcester* court,

---

House of Representatives, 469 A.2d 1161, 1166 (R.I. 1983), which indicated that a bill substantially similar to § 20J would be ruled unconstitutional if enacted.

[242] 617 N.E.2d 635 (Mass. 1993).

[243] *Id.* at 636.

[244] *Id.*

[245] *Id.*

[246] *Id.*

[247] *Id.* at 636–37.

[248] 617 N.E.2d 637 (Mass. 1993) (*Worcester* also involved a rape counseling center's refusal to disclose confidential records, and like the RCS of Lowell, the RCP of Worcester was held in civil contempt and fined daily until it complied.).

however, directed the trial court to consider the request for disclosure of the absolutely privileged, confidential records in accordance with the procedures set out in *Bishop*. The SJC made this ruling even though the procedures outlined in *Bishop* addressed only the scope of some conditionally privileged psychological records.[249] While it is difficult to predict how the SJC will eventually rule on the constitutionality of the absolute privilege, an examination of the Pennsylvania Supreme Court's rulings on this issue may shed some light.

D. *Pennsylvania's Absolute Privilege for Sexual Assault Counselors*

The Pennsylvania Supreme Court has upheld the constitutionality of an absolute privilege granted to communications between a sexual assault counselor and a survivor. According to the highest court in Pennsylvania, the narrowly tailored privilege can be upheld under the Pennsylvania state constitution because the government has a compelling interest in assisting the survivor in the recovery process, which outweighs even the defendant's right to confrontation.[250] Pennsylvania case law, however, did not always reflect such a belief.

The impetus for change in the law to favor the survivor's privacy interests came from the state legislature rather than from the courts. The Pennsylvania legislature enacted an absolute privilege for confidential communications made to sexual assault counselors in 1981[251] in direct response to an earlier decision by the Pennsylvania Supreme Court, *In re Pittsburgh Action Against Rape* (*PARR*).[252]

In *PAAR*, the court refused to expand the common law in order to create an absolute privilege that would protect communica-

---

[249] *Bishop*, 617 N.E.2d at 993, n.2.

[250] Commonwealth v. Wilson, 602 A.2d 1290, 1296 (Pa. 1992), *cert. denied sub nom.* Aultman v. Pa., 112 S.Ct. 2952 (1992).

[251] 42 PA. CONS. STAT. ANN. § 5945.1(b) (1982) ("[a] sexual assault counselor has a privilege not to be examined as a witness in any civil or criminal proceeding without the prior written consent of the victim being counseled by the counselor as to any confidential communication made by the victim to the counselor or as to any advice, report or working paper given or made in the course of the consultation."). This statutory language is similar to the sexual assault counselor privilege in Massachusetts as codified in MASS. GEN. L. ch. 233, § 20J (1992).

[252] 428 A.2d 126 (Pa. 1981).

tions between a rape counselor and her client.[253] Instead, the court created a limited privilege, allowing the defense counsel, after an *in camera* inspection, to examine confidential statements pertaining directly to the offense.[254] The absolute privilege, created by statute in 1981, abrogated the effects of *PARR*. Pennsylvania now guarantees absolute confidentiality[255] to both oral and written communications between the sexual assault survivor and her counselor in the course of their relationship.[256] The privilege even extends beyond the survivor to "those persons who have a significant relationship with a victim . . . ."[257]

The 1992 Pennsylvania case, *Commonwealth v. Wilson*, upheld the constitutionality of the absolute privilege granted to rape counselor records.[258] The *Wilson* court, recognizing the benefits of counseling,[259] found that the absolute privilege did not violate the confrontation or the compulsory process clauses of the Pennsylvania state constitution.[260] The confrontation clause, which guarantees the defendant's right to question adverse witnesses, was satisfied by the fact that the defense counsel was given wide latitude to question witnesses at trial.[261] The court reasoned that the constitutional right to conduct cross examination "does not include the power to require the pretrial disclosure of any and all information that might be useful in contradicting unfavorable testimony."[262]

The *Wilson* court also ruled that the absolute privilege does not violate the defendant's due process rights, because the state has a compelling interest in protecting the confidentiality of the

---

[253] *Id.* at 127–28.

[254] *Id.* at 132.

[255] The privilege may be breached in certain situations where an attorney-client privilege may be breached, *see infra* notes 264–268.

[256] 42 PA. CONS. STAT. ANN. § 5945.1(a) (1982). *See also* Commonwealth v. Eck, 605 A.2d 1248 (Pa. Super. 1992) (holding that absolute privilege applied to both the testimony of the sexual assault counselor as well as the records created in the course of the relationship); *Wilson*, 602 A.2d at 1295 ("the statutory privilege considered here must extend to the subpoena of records and other documents developed throughout the counseling relationship, any other interpretation of the statute would render the entire privilege meaningless.").

[257] 42 PA. CONST. STAT. ANN. § 5945.1(a) (1982).

[258] 602 A.2d at 1294.

[259] *Id.* at 1298 (Larsen concurring, "the need for and benefits of counseling for a rape victim are extraordinary, and the most accessible counseling of a rape victim is the rape crisis counselor.").

[260] *Id.* at 1296.

[261] *Id.*

[262] *Id.*

records, and the interest is narrowly tailored to achieve the goal of promoting the survivor's recovery.[263] The *Wilson* court may have been influenced by the fact that Pennsylvania had already established an absolute privilege for psychiatrist/psychologist-patient communications in 1978.[264] In order to deny an absolute privilege to the sexual assault counselor-patient relationship, the court would have to justify the denial of equal protection to women who could not afford the more expensive services of a licensed professional.

Pennsylvania constructed an absolute privilege for communications to psychiatrists and psychologists by pegging the privilege to the attorney-client privilege.[265] Like the attorney-client privilege, the privilege for the psychiatrist/psychologist is absolute except in certain limited circumstances.[266] Further, the privilege is waived where a litigant's mental condition is placed directly at issue,[267] and where adequate notice was given during the psychological examination that disclosure may take place.[268]

In *Commonwealth v. Kyle*,[269] a Pennsylvania superior court construed the psychologist-patient privilege as absolute and not

---

[263] *Id.* at 1297. The state's compelling interest is indicated by the enactment of an absolute privilege.

[264] 42 PA. CONS. STAT. ANN. § 5944 (1990) ("No psychiatrist . . . shall be, without the written consent of his client, examined in any civil or criminal matter as to any information acquired in the course of his professional services in behalf of such client. The confidential relations and communications between a psychologist or psychiatrist and his client shall be on the same basis as those provided or prescribed by law between an attorney and client.")

[265] 42 PA. CONS. STAT. ANN. § 5916 (1982) ("In a criminal proceeding counsel shall not be competent or permitted to testify to confidential communications made to him by his client, nor shall the client be compelled to disclose the same, unless in either case this privilege is waived upon the trial by the client."). *See* Samuel J. Knapp, Leon VandeCreek, & Perry A. Zirkel, *Privileged Communications for Psychotherapists in Pennsylvania: A Time for Statutory Reform*, 60 TEMP. L.Q. 267 (1987) (criticizing the comparison to the attorney-client model for failing to protect psychotherapeutic communications made in the presence of third parties).

[266] Such circumstances include when the communication involves a future criminal act likely to cause substantial harm to another party (42 PA. CONS. STAT. ANN. RULES OF PROFESSIONAL CONDUCT, Rule 1.6(c)(1)), when the patient uses the counselor's services to commit a criminal act (*id.*, Rule 1.6(c)(2)), or when the psychiatrist or the psychologist has to defend against the patient in a civil or criminal suit (*id.*, Rule 1.6(c)(3)).

[267] *See* Loftus v. Consolidated Rail Corp., 12 D. & C.4th 357 (1991) (holding that the privilege is waived where plaintiff seeks damages for psychological injury); Russell v. Commercial Union Ins. Co., 9 D. & C.4th 632 (1991) (allowing the insurer access to the insured's psychiatric records where the insured's mental condition is at issue in a claim for payment).

[268] *See* In Interest of Bender, 531 A.2d 504 (Pa. Super. 1987); Matter of Adoption of Embick, 506 A.2d 455 (Pa. Super. 1986), *appeal denied*, 520 A.2d 1385 (Pa. Super. 1987).

[269] 533 A.2d 120 (Pa. Super. 1987), *appeal denied*, 541 A.2d 744 (1988).

outweighed by the defendant's due process rights.[270] The court stated that "a compelling public interest would justify the total non-disclosure of information."[271] The court further noted that the privilege, which serves the public interest by promoting the general well-being of the citizenry[272] and protects the privacy interest of the patient,[273] should be upheld as an absolute privilege.[274]

The *Kyle* court considered the effective treatment of the sexual assault survivor as a paramount concern and refused to subject the confidential file to even an *in camera* review.[275] According to the court, the privilege did not infringe upon the defendant's right to confrontation because the defense counsel could still subject the survivor to a full array of questions on cross examination.[276] The court argued that the defense counsel will not be at a significant disadvantage relative to the prosecution because neither the defense nor the prosecution will have access to these confidential files.[277]

Pennsylvania case history illustrates two points. First, it is possible to argue for the constitutionality of an absolute privilege granted to a sexual assault survivor's counseling records. To uphold the constitutionality of such a privilege, courts must simply recognize the fact that confidential counseling relationships will add definite value to society, while the defendant's access to confidential files may add the possible value of leading to more accurate verdicts. The second point illustrated by Pennsylvania is that even an absolute privilege does not guarantee privacy in all circumstances. Like the attorney-client privilege,

---

[270] *Id.* at 129. *See also* Commonwealth v. Kennedy, 604 A.2d 1036 (Pa. Super. 1992), *appeal denied*, 611 A.2d 711 (1992) ("It should be readily apparent that the general powers of courts do not include the power to order disclosure of materials that the legislature has explicitly directed be kept confidential.") (quoting Commonwealth v. Moore, 584 A.2d 936, 940 (Pa. 1991)). *But see* Commonwealth v. Lloyd, 567 A.2d 1357 (Pa. 1989) (allowing the defendant direct access to psychotherapy records kept by a hospital because the records were not statutorily protected, state-maintained records).

[271] 533 A.2d at 125.

[272] *Id.* at 126.

[273] *Id.* at 127.

[274] *Id.* at 129.

[275] *Id.* at 131.

[276] *Id.* at 130 ("[T]he privilege only limits access to statements made during the course of treatment by the psychologist. It does not foreclose all lines of defense questioning.").

[277] *Id.*

absolute privilege for counseling relationships must yield in certain enumerated circumstances.

V. Conclusion

Courts and legislatures of each state must define a clear scope of counseling privilege for communications between a sexual assault survivor and her counselor. A limited privilege accorded to the sexual assault survivor's counseling communications will not promote the utilization of counseling services. On the other hand, an absolute privilege in which a counselor is forbidden from breaching the privilege will hamper other significant public interests, such as warning third parties when the patient poses an immediate physical threat.

A statutory, semi-absolute privilege that guarantees absolute confidentiality except in certain well-defined, extraordinary circumstances is one possible solution. The semi-absolute privilege will prevent the defense counsel from violating the sanctity of a sexual assault survivor's counseling communications made in a trusting atmosphere. The privilege may be breached in some extraordinary circumstances, but the survivor will receive adequate notice of such situations before entering into the counseling relationship.

Pennsylvania serves as an example of a state that has made an affirmative decision to protect the survivors of sexual assault from further harassment. Massachusetts has also taken one preliminary step towards encouraging the sexual assault survivor to seek counseling. It has required the defense counsel to make a showing of legitimate need before she can obtain an *in camera* inspection of records protected by either an absolute or a limited privilege. In order to provide stronger protection for sexual assault survivors, however, the Massachusetts legislature should enact a semi-absolute privilege for counseling communications.

The semi-absolute privilege should apply to all communications made to counselors for the purpose of assisting the survivor in dealing with the trauma caused by sexual assault. The privilege should be in effect regardless of whether the service is provided by a rape crisis center counselor, a psychologist, psychotherapist, a psychiatrist, a social worker, or a nurse. Two promising bills are currently pending in the legislature. Even if the bills pass, however, the SJC must ultimately affirm the con-

stitutionality of such statutes for survivors to feel completely secure in the confidentiality of their counseling relationships. The value of that comfort for sexual assault survivors is overwhelming and warrants such action.

BOOK REVIEWS

THE TYRANNY OF THE MAJORITY: FUNDAMENTAL FAIRNESS IN REPRESENTATIVE DEMOCRACY. By *Lani Guinier.* New York: The Free Press, 1994. Pp. xx, 324, foreword, notes, acknowledgments, list of cases, subject index. $24.95 cloth.

> It is of great importance in a republic not only to guard the society against the oppression of its rulers, but to guard one part of the society against the injustice of the other part.
>
> —James Madison

Perhaps most easily identified, albeit critically, with the scathing label "Quota Queen,"[1] Lani Guinier was never provided a forum to respond to the harsh indictments waged against her scholarship following her 1993 nomination to be Assistant Attorney General for Civil Rights. President Bill Clinton, responding to media criticism that categorized Guinier's ideas and writings as "anti-democratic," chose to withdraw her nomination within a matter of months. Atypical of recent embattled nominees, Guinier was deprived of the opportunity to respond to criticism that had previously been extended to the likes of Justice Clarence Thomas, Justice Robert Bork, and Zoe Baird. In fact, Guinier's nomination process never even proceeded beyond nomination.[2]

In an attempt to redeem both the substance of her scholarship and the basis for debating the legitimacy of voting rights reforms, Guinier has compiled her previous scholarship into *The Tyranny of the Majority.*[3] Primarily an experimental analysis of reforms for the American democratic process, Guinier's new book is also intended to respond to those critics who either summarily dismissed her reform proposals or hastily designated her "Clinton's Quota Queen." Indeed, *The Tyranny of the Majority* finally enables Guinier to advance her legal and political ideas, which were considered antithetical to American democ-

---

[1] The term "Quota Queen" was originally coined in an editorial column by Clinton Bolick, a lawyer at the Institute for Justice. Clinton Bolick, *Clinton's Quota Queens*, WALL ST. J., Apr. 30, 1993, at A12.

[2] For a full account of the events surrounding Guinier's nomination, *see, e.g.*, Gwen Ifill, *The Guinier Battle; Anatomy of the Failure To Confirm a Nominee*, N.Y. TIMES, June 5, 1993, at A9.

[3] It should be noted that the essays in *The Tyranny of the Majority* have been edited (some sections and footnotes have been omitted) as well as updated (developments since the essays were originally published have been included).

racy and thus undeserving of candid appraisal during the nominations process.

Stephen L. Carter's[4] Foreword in *The Tyranny of the Majority* chronicles the withdrawal of Lani Guinier's nomination to the position of Assistant Attorney General of the Justice Department's Civil Rights Division. Carter recalls the sociopolitical factors that eventually led to Guinier being branded "Quota Queen" and the media's claim that her writings advocated a "racial spoils system" in American politics. In her defense, Carter asserts that Guinier's reform proposals are neither radical nor revolutionary but are, in fact, democratic (p. xv). Carter also writes that Guinier's scholarship was misunderstood largely because her arguments are complicated and therefore too perplexing for a lay audience, particularly for those reporters who so doggedly criticized her proposals (p. xiv). For Carter, the virtue of *The Tyranny of the Majority* is that it will finally permit the public to draw its own conclusions about Guinier's scholarship while avoiding the fatuousness of the media (p. xx). Accordingly, it is the record which should speak for itself.

The book's first essay, "The Tyranny of the Majority," introduces Guinier's justifications for her legal and political agenda and details briefly the salient themes that run throughout her work. For Guinier, genuine democracy cannot be realized where tyranny by the majority exists (p. 6). Essentially, Guinier declares that her political views are "Madisonian" in character, recalling James Madison's fears of majority tyranny. According to Guinier, Madison believed that majority tyranny represented "the great danger" to constitutional democracy because in a heterogeneous community the majority is likely to be self-interested or indifferent to the interests of the minority (p. 3). The ideal democracy, reflected in Madison's writings, always protects the minority from the otherwise overbearing and insurmountable will of the majority.[5] Alluding to Madison's conception of constitutional democracy, Guinier asserts that decisionmaking rules in a multiracial democracy must be structured so as to afford majority rule and thwart majority tyranny (p. 5).

---

[4] Stephen L. Carter is the William Nelson Cromwell Professor of Law at Yale University.

[5] For a concise treatment of James Madison's political philosophy, *see* BRUCE E. CAIN & W.T. JONES, *Madison's Theory of Representation*, *in* THE FEDERALIST PAPERS AND THE NEW INSTITUTIONALISM 11–30 (Bernard Grofman & Donald Wittman eds., 1989). *See also* THE FEDERALIST No. 51 (James Madison).

In this essay, Guinier introduces her fundamental "principle of taking turns," which is a recurring theme throughout her writings. Here, Guinier again draws on Madisonian tradition, maintaining that the principle of taking turns originates with the political writings of James Madison. This principle is best defined as a "positive-sum, taking turns solution" to ensure that the rules of democracy reflect the values of fairness, compromise, and consensus, which constitute the heart of the democratic ideal (p. 5).

In this introductory essay, Guinier summarizes her basic strategies and proposals for achieving Madisonian democracy in light of the principle of taking turns. These proposals include cumulative voting and supermajority voting in addition to a host of voting reforms that would, in Guinier's view, enhance political participation and government accountability. Guinier expects the debate that *The Tyranny of the Majority* will inevitably engender will garner the political consensus necessary in order to fulfill Madison's goal of safeguarding "one part of the society against the injustice of the other part." (p. 3).

Guinier's next essay in *The Tyranny of the Majority* recounts the rank hostility of the Reagan Administration to traditional civil rights policies in general and race-conscious policies in particular.[6] Entitled "Keeping the Faith—Black Voters in the Post-Reagan Era,"[7] the essay provides an in-depth account of the Reagan Administration's record in the area of voting rights and electoral politics (p. 22). In this review, Guinier indicts the Reagan Administration for seeking to "drive civil rights laws out of the marketplace," thereby creating a divisive political environment that has contributed to an increasing sense of isolation among African Americans (pp. 22–23).[8]

---

[6] Guinier is not alone in her critical assessment of the Reagan Administration's fervent opposition to the traditional civil rights agenda and race-based remedial measures. *See generally*, STEVEN A. SHULL, A KINDER, GENTLER RACISM?: THE REAGAN-BUSH CIVIL RIGHTS LEGACY (1993) (concluding that the Reagan Administration was responsible for the complete overhaul of several major civil rights protections); ROBERT R. DETLEFSEN, CIVIL RIGHTS UNDER REAGAN (1991) (emphasizing that the Reagan Administration aggressively challenged race-based preferences, particularly in the areas of education and employment).

[7] This essay was originally published in 1989 as a law review article. *See* Lani Guinier, *Keeping the Faith: Black Voters in the Post-Reagan Era*, 24 HARV. C.R.-C.L. L. REV. 393 (1989).

[8] *See* Julie Johnson, *Deciding What to Do Next About Civil Rights*, N.Y. TIMES, Mar. 12, 1989, at E5 ("For many blacks, the [Reagan] legacy was a heightened racial tension arising from the rhetoric to the effect that it was wrong to remedy discrimination

In "Keeping the Faith," Guinier characterizes the Reagan Administration's stance as "anti-civil rights" based upon her examination of a range of policies, practices, and procedures that demonstrated the Administration's staunch and unwavering opposition to traditional civil rights policies.[9] Specifically, Guinier emphasizes the Administration's concerted efforts to oppose women and minorities in employment, education, and voting rights cases. She further criticizes the Administration for narrowly interpreting the mandate of the 1965 Voting Rights Act[10] and issuing the first veto against a civil rights bill, the Civil Rights Restoration Act of 1987, in 121 years (p. 23).[11] Guinier also points out that the Administration adopted a neoconservative philosophy that included repealing affirmative action, abrogating many class-based remedies, and abandoning racial discrimination cases except those filed by "identifiable" victims of racism (p. 23).[12] In

---

against blacks with measures that might discriminate against whites."). *See also* HOWARD SCHUMAN et al., RACIAL ATTITUDES IN AMERICA 37–38 (1985) ("The apprehensions of civil rights activists seemed confirmed by a number of Reagan initiatives. The Department of Justice began to enter school desegregation cases on the side of school districts facing desegregation orders. Virtually all pressure to use busing as a means of achieving desegregation was suspended. There was even an attempt to restore the tax-exempt status of the Bob Jones University, a private school that forbade interracial dating. Reagan also was silent on extension of the Voting Rights Act, though once it became clear that congressional support for it was solid, he agreed to sign the bill. Similarly, Reagan initially opposed a national holiday honoring the birthday of Martin Luther King, Jr., but relented when faced with a united Congress.").

[9] *See generally* Drew S. Days III, *Turning Back the Clock: The Reagan Administration and Civil Rights*, 19 HARV. C.R.-C.L. L. REV. 309, 346–47 (1984) ("[The Reagan Administration] has consistently shown an inclination in matters of civil rights to move in precisely the opposite direction from former administrations. It has sought to undermine the achievements of preceding administrations, Republican and Democratic. Mechanically repeating stock phrases about 'busing being bad' and 'quotas being unfair' as articles of faith, its officials have demonstrated an extreme ideological rigidity, refusing to yield in the face of even the most compelling facts and reasoning to the contrary."). *See also* Joel L. Selig, *The Reagan Justice Department and Civil Rights: What Went Wrong*, 1985 U. ILL. L. REV. 785.

[10] Voting Rights Act of 1965, Pub. L. No. 89-110, 79 Stat. 445 (codified as amended at 42 U.S.C. §§ 1971, 1973–1973p (1988)). *See generally* Days, *supra* note 9.

[11] Despite President Reagan's 1988 veto, the Civil Rights Restoration Act was nonetheless enacted by congressional override. Civil Rights Restoration Act of 1987, Pub. L. No. 100-259, 102 Stat. 28 (codified at 20 U.S.C. § 1681) (1988).

[12] Kimberlé Crenshaw, *Race, Reform, and Retrenchment: Transformation and Legitimation in Antidiscrimination Law*, 101 HARV. L. REV. 1331, 1336–37 (1988) ("The Reagan Administration arrived in Washington in 1981 with an agenda that was profoundly hostile to the civil rights policies of the previous two decades. The principal basis of its hostility was a formalistic, color-blind view of civil rights that had developed in the neoconservative "think tanks" during the 1970s. Neoconservative doctrine singles out race-specific civil rights policies as one of the most significant threats to the democratic political system. Emphasizing the need for strictly color-blind policies, this view calls for the repeal of affirmative action and other race-specific remedial policies, urges an end to class-based remedies, and calls for the Administra-

sum, Guinier concludes that the policies and practices of the Reagan Administration served only to thwart the efforts of civil rights proponents to produce a political environment in which blacks could depend upon the rigorous enforcement of voting rights in order to affect legislative policies and elect representatives of their choice (p. 40).

"Keeping the Faith" is perhaps most critical of the Reagan Administration's failure to enforce the 1965 Voting Rights Act. First, Guinier faults the Administration for not filing any new lawsuits upon the enactment of the amended Section 2 of the Voting Rights Act (pp. 26–27).[13] Additionally, Guinier suggests the Administration deliberately intended to limit the interpretation and application of Section 2 altogether (p. 27). For example, Guinier refers to the conduct of the Reagan Administration in the case of *Major v. Treen*[14] to illustrate the Administration's wanton abandonment of the federal government's traditional role in protecting minority voters.[15] According to Guinier, the Administration's failure to enforce the Voting Rights Act properly not only undermined traditional protections for voting rights but also left minority victims to their own devices in an environment overwhelmingly antagonistic to claims of voting rights violations (p. 29).[16]

---

tion to limit remedies to what it calls 'actual victims' of discrimination.") (footnotes omitted). *See also*, Shull, *supra* note 6, at 101–28.

[13] Under Section 2, an electoral system violates the Voting Rights Act if it results in racial or language minorities having "less opportunity than other members of the electorate to participate in the political process and to elect representatives of their choice." Voting Rights Act Amendments of 1982, Pub. L. No. 97-205, 96 Stat. 134 (codified as amended at 42 U.S.C. § 1973(a) (1988)).

[14] 574 F. Supp. 325 (E.D. La. 1983). Guinier claims that in Treen, a voting rights case involving black plaintiffs, the Assistant Attorney General William Bradford Reynolds "failed as a prosecutor to investigate fully the facts of the Treen submission" and "refused to challenge [the] discriminatory redistricting change," notwithstanding his role as the primary federal enforcer of the 1965 Voting Rights Act. (p. 28).

[15] *See generally*, MALCOLM M. FEELEY & SAMUEL KRISLOV, CONSTITUTIONAL LAW 655 (1985) ("[T]he Reagan administration's stance on civil rights . . . is at significant odds with the civil rights positions of the five previous administrations, all of which actively used their discretionary powers to pursue civil rights claims through the administrative process and to press civil rights claims in the courts. In sharp contrast, the Reagan administration has slowed to a near standstill or has actively opposed civil rights actions." *See also*, Shull, *supra* note 6, at 184 (President Reagan was "the first president to retreat on civil rights dramatically and on a broad scale.").

[16] To buttress this claim, Guinier emphasizes that Reagan appointments to the bench (366 in total) currently comprise over half the federal judiciary. Guinier also notes that prior to making nominations for federal judgeships, the Reagan Administration routinely screened prospective nominees on civil rights issues such as school desegregation, affirmative action, and race-based remedies. (p. 23).

In view of the Reagan Administration's record on voting rights enforcement, Guinier recommends immediate action to counteract the ill effects of the "Reagan Legacy." Since "Keeping the Faith" hails the Voting Rights Act as the most effective civil rights legislation to date, Guinier's first recommendation is for subsequent presidential administrations to adopt a national voting rights agenda. This agenda should include a thorough and comprehensive strategy to enforce existing voting rights laws in good faith (p. 38). As a secondary goal, Guinier is certain that the federal government, in the interests of developing and enforcing a national voting rights agenda, must appoint a diverse group of black federal officials while also supporting black candidates for federal, state, and local offices (p. 39). Guinier's suggestion that Congress and the Executive adopt a policy of "Affirmative Recruitment" of minorities will ensure that legislative and executive initiatives in the area of voting rights are both supported and enforced by a "representative" group of federal officials.

"Keeping the Faith" also calls upon the Executive to explore new legal remedies to redress the discrepancy between American democratic ideology and current practice. Guinier suggests the creation of a new executive branch task force to develop and promote a new federal agenda for voting rights reform. In appointing the task force's members and staff, she once again urges a policy of "Affirmative Recruitment" to ensure that a substantial number of nonwhite policy analysts, lawyers, and elected officials are a part of the decisionmaking process for any future federal voting rights reform (pp. 39–40).

In "Triumph of Tokenism—The Voting Rights Act and the Theory of Black Electoral Success,"[17] Guinier addresses the complex issues surrounding black representatives. Her conclusions challenge the predominant viewpoint within American politics that black political success hinges on the election of black politicians.[18] This "Theory of Black Electoral Success," according to Guinier, has not only found sanction in the courts but has be-

[17] This essay, originally published in 1991, was also a law review article. *See* Lani Guinier, *The Triumph of Tokenism: The Voting Rights Act and the Theory of Black Electoral Success*, 89 MICH. L. REV. 1077 (1991).

[18] *See* CAROL M. SWAIN, BLACK FACES, BLACK INTERESTS: THE REPRESENTATION OF AFRICAN AMERICANS IN CONGRESS 207 (1993) ("The most common strategy for electing black representatives has centered on claiming and creating districts with black majorities.").

come the hallmark of contemporary litigation strategy under the Voting Rights Act (p. 42).

Nevertheless, Guinier has three reasons for viewing the black electoral success theory as the main impediment to achieving the overriding objective of a political system that is more responsive to black political interests. First, the theory "romanticizes" black political figures as black empowerment role models while ignoring the seemingly apparent problem of tokenism (p. 42). Second, the theory fails either to broaden the base of black political participation or to reform the substance of political decisions appropriately. Finally, black officials elected from majority-black districts[19] with geographically and socially isolated constituencies cannot substantively influence the policies of their white counterparts because the entire political environment is racially polarized (p. 43). Thus, in "The Triumph of Tokenism," Guinier primarily seeks to examine more carefully the key assumption of the theory of black electoral success that the election of black representatives best fulfills political self-confidence and legislative influence for black voters (p. 42).

Guinier first attacks the four basic assumptions that provide the foundation for the theory. The first assumption is what Guinier refers to as the "Authenticity Assumption," whereby representation provides some psychological value to those voters who have similarly situated representatives (p. 55). This assumption presumes that race is the prime indicator in determining who best represents the interests of black voters. The Authenticity Assumption, therefore, demonstrates what Guinier calls an "essentialist impulse" among black voters: because black officials are black, they are representative (p. 55).

While recognizing that several claims of the Authenticity Assumption are legitimate, Guinier contends that the assumption is, at best, a limited empowerment tool. She argues that the Authenticity Assumption's main shortcoming is its failure to inquire sufficiently into the credentials possessed by black candidates and the extent to which they can serve as responsive and accountable representatives (p. 58). Moreover, the assumption falls short of explaining how black voters can hold black representatives accountable when the political support of white voters

[19] The term "majority-black districting" refers to the current practice of gerrymandering the voting districts of a state so that certain districts are created where black voters constitute the majority.

is involved. In Guinier's view, notwithstanding the electoral ratification by a black majority, a black representative also elected by whites may be incapable of discerning the respective mandates of black and white voters (p. 58).

The second assumption of black electoral success theory is the "Mobilization Assumption," suggesting that black electoral successes directly affect black political participation at the grassroots level. This assumption claims that black voters are encouraged to participate in electoral politics if and when there is an opportunity to elect black candidates (p. 59). Though correct in recognizing the increase in black voter turn out during elections involving black candidates, the Mobilization Assumption, Guinier maintains, ignores the critical importance of post-election participation.[20] For Guinier, post-election participation cannot be neglected because it encourages a healthy political dialogue between representatives and constituents that helps achieve a more responsive government (p. 59).

The third assumption, which Guinier terms the "Polarization Assumption," is founded on empirical evidence supporting the proposition that blacks and whites both tend to vote for candidates of their own race.[21] In view of racial bloc voting, the Polarization Assumption provides that black electoral success is possible to the extent that protective voting rights measures are in force. Thus, in the absence of protective measures, the persistence of racial bloc voting perpetuates white majority rule because black candidates cannot build an effective multiracial majority (p. 60).

Guinier's critique of the Polarization Assumption is that a more adequate response is needed to address problems associated with racial bloc voting. Since the assumption witlessly deals with consensus-building and the need to encourage political bargaining, Guinier dismisses the Polarization Assumption because it summarily presumes that black representatives can participate substantively in the decisionmaking process and will not be im-

---

[20] Guinier uses the term "post-election participation" to refer to the ideal of a politically active constituency that carefully monitors the official actions of elected representatives.

[21] *See, e.g.*, Linda Williams, *Black Political Progress in the 1980s: The Electoral Arena*, *in* THE NEW BLACK POLITICS: THE SEARCH FOR POLITICAL POWER 126 (Michael B. Preston et al. eds., 1987) ("[R]acial voting patterns are still closely tied to the election of black officials. Most black elected officials represent majority black constituencies.").

peded in a racially polarized climate where racial bloc voting is prevalent (p. 61).

The fourth and final assumption of black electoral success theory is the "Responsiveness Assumption," which maintains that black representatives support the traditional civil rights agenda and most other issues related to black political interests. Implicit in this assumption is the view that black representatives possess a unique understanding of critical issues affecting the black community (p. 66).

According to Guinier, however, the Responsiveness Assumption does not provide an adequate check on the elected representative. The assumption neither explains how responsiveness is achieved merely by virtue of electing a black representative nor does it account for the often competing interests of poor or working-class blacks and middle-class blacks (p. 69).

In contrast to the implications of the black electoral success theory, Guinier offers two comprehensive proposals. First, she calls on voting rights proponents to challenge the legitimacy of a political system in which the interests of a racially homogeneous majority consistently prevail. Her point is that voting rights proponents must redefine conceptions of fundamental fairness throughout the American political process (p. 70). Second, voting rights proponents should cast doubt upon the prevailing notion that physical representation sufficiently realizes voters' political preferences (p. 70). Altogether, "Triumph of Tokenism" demands that the American voting system be restructured to ensure a fair chance for minority voters to achieve equal recognition within the legislative decisionmaking process.

In her next essay, entitled "No Two Seats—The Elusive Quest for Political Equality,"[22] Guinier considers remedial alternatives to the traditional method of creating majority-black districts, such as cumulative voting and supermajority rules,[23] and their implications for collective decisionmaking within municipal and county legislative bodies. In the interests of pursuing a more

---

[22] This essay was originally published as a law review article in 1991. *See* Lani Guinier, *No Two Seats: The Elusive Quest for Political Equality*, 77 VA. L. REV. 1413 (1991).

[23] Although little treatment is given to supermajority rules and their would-be effect in *The Tyranny of the Majority*, Guinier suggests that a supermajority rule may be appropriate where egregious diminution of black political influence is evident. This rule would require a supermajority vote on issues of importance to the majority and a minority veto on issues critical to the minority, which would effectively promote cross-racial coalition building. (p. 108).

inclusive voting system, Guinier suggests that these remedial alternatives may enable politically cohesive minorities to participate more fully in the political process, particularly in cases where efforts to eliminate illegal voting practices have proved unsuccessful. For Guinier, implementing cumulative voting and supermajority rules includes the additional benefit of promoting the development of cross-racial coalitions.

Before examining remedial alternatives to majority-black districting, Guinier once again addresses the practice's implications for both black political influence and participation. While Guinier's critique of black electoral success theory is similar to the critique previously offered in "The Triumph of Tokenism," here she analyzes the theory in light of the specific objectives of the Voting Rights Act.

According to Guinier, the Voting Rights Act was intended to make state and local government more responsive to minority voters (p. 72). Thus, Guinier maintains that the Act was designed to promote both political equality and political empowerment for black voters (p. 72). The Act's political equality objective is accomplished when each voter exercises his or her right to vote pursuant to the constitutional mandate of "one person, one vote."[24] The political empowerment objective, on the other hand, is less technical and relies on a proposition implicit in the "one person, one vote" mandate: equal voting weight and equal voting power (p. 72). Guinier concedes that the political empowerment objective is closer to being a moral proposition than a technical rule. Nevertheless, she points out that accomplishing this objective would be impossible if the technical mandate of "one person, one vote" was neglected. Protecting the right to vote is the first step toward achieving political empowerment, which she defines as using the vote to mobilize support for black political interests. Guinier expects blacks to employ such power to ameliorate their depressed and isolated social and economic status (p. 73).

Guinier concludes in "No Two Seats" that the current practice of majority-black districting, as applied on the county and mu-

---

[24] According to Guinier, the Voting Rights Act is the statutory embodiment of the "one person, one vote" principle already enunciated and recognized in Supreme Court jurisprudence. The "one person, one vote" principle was first introduced by the Supreme Court in Gray v. Sanders, 372 U.S. 368 (1963). "Once the geographical unit for which a representative is to be chosen is designated, all who participate in the election are to have an equal vote." *Id.* at 379. The Supreme Court first applied the "one person, one vote" rule in the case Wesberry v. Sanders, 376 U.S. 1 (1964). "One man's vote in a congressional election is to be worth as much as another's." *Id.* at 8.

nicipal levels, fails to achieve the political equality and political empowerment goals of the Voting Rights Act (p. 74). The practice does not account for prejudice and racism among white incumbent legislators, a factor that hampers the ability of elected minority representatives to work effectively in the legislature (pp. 75–77).[25] Also, as Guinier explained earlier in "Triumph of Tokenism," majority-black districting fails to ensure the responsiveness and accountability of representatives to their constituents (pp. 80–82). Finally, Guinier criticizes pursuing majority-black districting on the county and municipal level because it tends to depress the level of political competition while also discouraging the type of grassroots political organization necessary to mobilize voters to participate vigorously in the political process (pp. 82–86).

Yet, Guinier is somewhat less critical of plans to create minority influence districts. This kind of districting disperses the black voting population throughout several districts, thereby creating pockets of black electoral influence. Since minority influence districts guarantee that there is no majority monopoly of all political power in an "influenced" district, minority voters may potentially exercise a fair and reasonable share of procedural resources (p. 88). More importantly, these districts may provide minority voters with opportunities to influence elections and, indirectly, decisionmaking. Such involvement enables minorities to exchange information with other groups of voters concerning their policy preferences, thus stimulating an environment for political bargaining and consensus-building (p. 88).

Nonetheless, even after considering the seemingly apparent benefits that minority voters would derive from the creation of minority influence districts, Guinier objects, though mildly, to their application, particularly in cases where whites function as a politically cohesive district majority. In such a case, representatives elected by the racial majority wield disproportionate power since they can forgo accommodating the interests of minority voters at little or no political expense (p. 89). Conse-

---

[25] Guinier cites several examples of blatant discriminatory acts on the part of local governments that deprived minorities of the opportunity to participate equally in local politics. *See* Rojas v. Victoria Independent School District, 1988 WL 920531 (S.D. Tex. 1988), *aff'd*, 490 U.S. 1001 (1989) (school board altered its procedural rules, effectively preventing newly elected Mexican American woman from placing items on school board agenda); Major v. Treen, 574 F. Supp. 325 (E.D. La. 1983) (black legislators excluded from secret meeting on congressional redistricting).

quently, Guinier contends that minorities would not realize the purported benefits from implementing minority influence districts.

In rejecting majority-black districts and minority influence districts, Guinier instead calls for a "reconceptualization" of local minority political exclusion and the adoption of "Interest Representation." By definition, Interest Representation identifies voters that comprise a politically cohesive group on account of shared interests (p. 94). Essentially, Guinier's theory of Interest Representation assesses the extent to which politically cohesive minorities have their voting interests undermined. Under this theory, winner-take-all voting systems that dilute the power of politically cohesive minority voting groups violate the right to a meaningful vote (p. 94).

Interest Representation is best achieved, according to Guinier, through a system of cumulative voting that disaggregates the voting majority to ensure that fifty percent plus one of the voters will not control one hundred percent of the electoral or legislative power (p. 96). Voters would have the same number of votes as open seats and may, depending upon their particular voting interests, strategically concentrate their votes for a particular candidate (p. 95). In cases where there is a politically cohesive minority voting group, a cumulative voting system increases the odds that minority voters could elect a representative of their choice.

Guinier justifies cumulative voting on several grounds. First, it would create a more accountable political system by mobilizing voter interest and participation in electoral politics. Guinier suggests that implementing cumulative voting would increase voter turnout because voters will more actively participate in a political system that is no longer defined by winner-take-all results (p. 99). Such a voting regime would force incumbents to develop more substantive programs and proposals to ensure their re-election (p. 99). Moreover, during the course of election campaigns, new candidates would also develop better proposals to distinguish themselves from incumbents (p. 99). Thus, Guinier contends that a cumulative voting system would offer voters more substantive programs and proposals from which to choose while also providing an unprecedented degree of political accountability.

Cumulative voting also avoids race-conscious districting remedies that she believes may inadvertently fuel the resentment of

groups not protected under the Voting Rights Act,[26] such as whites and religious minorities (p. 100). These groups often perceive that their interests are submerged in majority-black districts (p. 100). Guinier also believes that cumulative voting promotes the formation of well-informed consensus in group decisionmaking better than the winner-take-all tradition of American electoral politics (p. 101).

Guinier concludes "No Two Seats" by defending the constitutionality of Interest Representation. She first argues that the Voting Rights Act must be interpreted in its broadest sense because the Supreme Court has already deferred to Congress' judgment that group-based affirmative remedies are not only necessary and appropriate but that any risks pursuant to such remedies are either nonexistent or *de minimus* (p. 115).[27] Secondly, Guinier claims that interest representation is consistent with the constitutional mandate of "one person, one vote" because cumulative voting ensures that each voter exercises a similarly "meaningful" vote (p. 116). Since all voters have the same number of votes to cast under a cumulative voting system, Guinier proffers that there is no violation of the "one person, one vote" principle.[28]

Interest Representation, as enunciated in "No Two Seats," might be viewed as redefining principles of political fairness and equality because interests, rather than voters, become accounted for and incorporated into the legislative decisionmaking process. "No Two Seats" is consistent with the thesis Guinier advances throughout *The Tyranny of the Majority*: one must measure the fairness of a voting system by its success in ensuring the deliberative process reflects the will of politically cohesive minorities as well as the will of homogeneous majorities.

---

[26] Section 2 of the Voting Rights Act, for example, protects "racial or language minorities" from standards, practices, or procedures that result in "less opportunity than other members of the electorate to participate in the political process." *See supra* note 13.

[27] Guinier cites the opinion in Gingles v. Edmisten, 590 F. Supp. 345 (1984), *aff'd in part and rev'd in part sub nom.* Thornburg v. Gingles, 478 U.S. 30 (1986), a Section 2 voting rights case, in support of this proposition. "In enacting amended Section 2, Congress made a deliberate political judgment that . . . national policy respecting minority voting rights could no longer await the securing of those rights by normal political processes . . . . In making that political judgment, Congress necessarily took into account and rejected as unfounded, or assumed as outweighed, several risks to fundamental political values that opponents of the amendment urged in committee deliberations and floor debate." *Id.* at 356.

[28] *See supra* note 23.

In "Groups, Representation, and Race Conscious Districting—A Case of the Emperor's Clothes,"[29] Guinier discredits the premises of the "Theory of Virtual Representation"—the idea that a majority of voters within a particular geographic community comprise a politically homogeneous unit.[30] Guinier finds fault with this traditional view because she questions whether a geographically districted group has a coherent political identity (p. 131). This theory belies electoral results that demonstrate that constituencies do not uniformly evince the same or similar political behavior.

Guinier mainly criticizes districting according to the theory of Virtual Representation for wasting votes, not only by denying voters who supported losing candidates with direct political representation but also by packing voters into homogeneous districts. "Packing,"[31] according to Guinier, dilutes the voting strength of the district in that particular jurisdiction since any votes beyond what is required to elect the winning candidate otherwise could have provided the necessary electoral margin for an appealing candidate from another district (p. 134).

Guinier disputes the legitimacy of Virtual Representation because she believes that some geographic districts contain constituents who may, in fact, reside there involuntarily (p. 129). Therefore, geographic districts may not reliably indicate choice but instead may more accurately reflect socioeconomic phenomenon such as race and income level (p. 129). Blacks and other minorities may disproportionately suffer from territorial districting since minorities—more often constrained by socioeconomic forces than whites—may lack flexibility in deciding where to reside in the first place.

"Groups, Representation, and Race Conscious Districting" also examines the degree to which race functions as an arguably inaccurate political proxy. While race is a reliable indicator of group consciousness, Guinier nonetheless criticizes the way it has been used to categorize the political interests of those who share a particular racial identity (p. 142). Guinier argues that race-conscious districting, like territorial districting, only iso-

---

[29] This essay was first published in 1993 as a law review article. *See* Lani Guinier, *Groups, Representation, and Race-Conscious Districting: A Case of the Emperor's Clothes*, 71 TEX. L. REV. 1589 (1993).

[30] *See generally*, ABIGAIL THERNSTROM, WHOSE VOTES COUNT? (1987).

[31] "Packing" is a districting technique that intentionally concentrates minority voters into districts with large minority populations.

lates minority voters, thereby stifling cross-racial legislative coalition-building (p. 135). Furthermore, Guinier views race-conscious districting as detrimental to minority voting interests because any such effort presumes the creation of white majority districts, which tend to dilute the influence of minority votes within those jurisdictions (p. 135).

For Guinier, at-large voting systems such as cumulative voting yield better results than territorial and race-conscious districting. Cumulative voting enables politically cohesive minority groups to vote strategically in an attempt to elect representatives of their own choice. Also, under a cumulative voting regime, coalition building is more feasible since constituencies may be formed on the basis of voters' interests (p. 149).

The final essay in *The Tyranny of the Majority* departs significantly from the previous essays in both form and substance. In "Lines in the Sand: A Review of Charles Fried's Order and Law: Arguing the Reagan Revolution—A Firsthand Account,"[32] Guinier offers yet another indictment of the Reagan Administration's hostility toward traditional civil rights policies. This piece reviews Charles Fried's *Order and Law: Arguing the Reagan Revolution—A Firsthand Account*.[33] For Guinier, *Order and Law* is significant for it details the impact of Fried's color-blind approach to constitutional adjudication and forecasts the future of litigation-based federal law reform (pp. 160–61).

Guinier criticizes Fried's effort to force government to be color-blind in its policies, practices, and procedures (p. 167). Specifically, Guinier asserts that Reagan's legal philosophy led the Executive branch to attempt to halt race-based remedies and racially based preferences unless applied to specific individuals who could legally establish they were harmed by specific acts of racial discrimination (pp. 176–77).[34] Thus, Guinier views Fried, the vanguard of the Reagan legal philosophy, as having been partly successful in aligning the federal courts with the Reagan Administration's conservative view of the law and, in turn, al-

---

[32] This Book Review was first published in 1993. *See* Lani Guinier, *Order and Law: Arguing the Reagan Revolution—A Firsthand Account*, 72 TEX. L. REV. 315 (1993) (book review).

[33] CHARLES FRIED, ORDER AND LAW: ARGUING THE REAGAN REVOLUTION—A FIRSTHAND ACCOUNT (1991). Charles Fried, a self-proclaimed Reagan Revolutionary, was Solicitor General during Reagan's second presidential term between the years 1985 and 1989.

[34] *See supra* notes 8–12 and accompanying text.

tering the course of traditional civil rights adjudication in America.

Since Reagan and Bush left behind an increasingly conservative federal judiciary, Guinier proposes both legal and political strategies as a counterbalance. First and foremost, Guinier suggests that progressive lawyers aggressively expose the ideological preferences that inform the decisionmaking of entrenched conservative judicial activists (p. 185). Revealing the ideological biases of the federal judiciary will, in Guinier's view, highlight the intensely political character of judicial decisionmaking (p. 185) For Guinier, recognition of the now-tenured conservative coalition in the federal courts is the first step toward challenging the claim of a restrained and accountable judiciary (p. 185).

Guinier additionally proposes that progressive activists resort to electoral politics to ensure judicial accountability. Specifically, Guinier suggests exploring term limitations to curtail the reign of the entrenched conservative federal judiciary (p. 185). As a corollary to her emphasis on the elections process, Guinier also proposes that progressives wage a more concerted political campaign within Congress and state legislatures. Since legislative bodies encourage interest group participation and are more accessible than the courts, Guinier believes that they offer the best opportunity for resolving differences and developing a public consensus on policy goals (p. 186).

The Tyranny of the Majority closes with the speech Guinier delivered upon the withdrawal of her nomination for Assistant Attorney General for Civil Rights.[35] Entitled "Lessons," this intensely personal epilogue is a fitting close to the concepts explored throughout the book and provides a brief but thought-provoking insight into Guinier's commitment to civil rights, both as a civil rights attorney and as an activist. Here, Guinier offers perhaps the cardinal admonition of *The Tyranny of the Majority*: "I hope that what has happened to my nomination does not mean that future nominees will not be allowed to explain their views as soon as any controversy arises. I hope that we are not witnessing that dawning of a new intellectual orthodoxy in which thoughtful people can no longer debate provocative ideas without denying the country their talents as public servants."

[35] Guinier's withdrawal speech for the nomination of Assistant Attorney General for Civil Rights was originally delivered on June 4, 1993 (p. 188).

With *The Tyranny of the Majority*, Guinier seeks to generate widespread debate over the efficacy of various voting rights reform proposals in protecting Madisonian democracy. In framing this debate, Guinier poses for the reader the ultimate question: Would her proposed reforms best fulfill Madison's ideal of guarding one part of society against the injustices of another part?

Guinier clearly is most concerned with protecting black voters and their political interests; perhaps, her legal and political reform proposals, while theoretical, are best evaluated in terms of their potential effects on the representation of black political interests. In this light, one can direct several criticisms at *The Tyranny of the Majority* and its proposals to reform the American electoral process.

First, Guinier's critical assessment of black electoral success theory—the idea that black political interests are best served by electing black representatives—underestimates the significant role of institutional power in the legislative decisionmaking process.[36] For example, Guinier admonishes against assuming that black voters are best represented by black representatives. She suggests that black political participation is considerably compromised when black electoral success amounts to nothing more than tokenism. Yet Guinier's fear is overly exaggerated, especially given the current dynamics of black federal institutional power and its current role in legislative decisionmaking.[37] Particularly on the federal level, black representatives have recently garnered unprecedented political success on a wide range of policy issues, which have benefitted both black voters and the general population.[38]

---

[36] *See* JOHN W. KINGDON, CONGRESSMEN'S VOTING DECISIONS 110–11 (2d ed. 1981) ("Among fellow congressmen, those who occupy formal positions in the elected party leadership or on the standing committees might have a special place in voting decisions . . . . Previous studies of legislative voting have discovered that the party leadership tends to vote with the ranking members of the committee which reported a bill under consideration.") (footnote omitted).

[37] *See* Steven V. Roberts, *New Black Power on Capitol Hill*, U.S. NEWS & WORLD REP., May 23, 1994, at 36 ("[T]he decision by House Republicans not to cooperate on many issues means that Democratic leaders rely more than ever on black-caucus votes.").

[38] *See* Kenneth J. Cooper, *For Enlarged Congressional Black Caucus, a New Kind of Impact*, WASH. POST, Sept. 19, 1993, at A4 (noting that the Congressional Black Caucus' refusal to vote the party line led to it playing the decisive role in defeating efforts by conservative and moderate Democrats to cap entitlements to Medicaid and Medicare; the Caucus also gained invaluable political concessions during two rounds of budget negotiations with the Clinton Administration).

For instance, the recent expansion of black institutional power within Congress has afforded blacks with a degree of substantive representation never before witnessed in American politics.[39] In fact, the Congressional Black Caucus, the epitome of black federal power, has delivered critical votes as well as input for a range of high-profile policy matters, including passage of the Violent Crime Control and Law Enforcement Act of 1994.[40] The Congressional Black Caucus has even decisively influenced American foreign policy.[41]

Guinier's reluctance to acknowledge the role of institutional power causes her to undervalue the creation of majority-black districts—the very process that facilitated the increased role of the Congressional Black Caucus in congressional decisionmaking.[42] In view of this contemporary political development, Guinier's

---

[39] Black representatives currently hold key positions within the Congressional power structure. *60 Minutes: The New Black Power*, (CBS television broadcast, Sept. 25, 1994) (transcript on file with the *Harvard Journal on Legislation*) [hereinafter *60 Minutes*]. "The Black [Congressional] Caucus now has 40 members in Congress, up from 26 just two years ago. Some of the seats are from newly created black districts ordered by the federal court, and some are occupied by powerful committee chairmen like Ron Dellums of the House Armed Services Committee, John Conyers of Government Operations and William Clay, chairman of the Post Office and Civil Services Committees. There's Deputy Whip John Lewis, and elder statesmen like Louis Stokes and Charles Rangel. In all, a block of 37 Democratic votes in the House of Representatives." *Id.* at 3.

[40] Pub. L. No. 103-322, 108 Stat. 1796 (1994). The force behind allocating $7 billion in prevention programs into the Crime Prevention and Criminal Justice Reform Act was the Congressional Black Caucus, which had declared this allocation to be non-negotiable. "[W]ithout the [Congressional] Black Caucus, there certainly would not have been a crime bill." *60 Minutes* at 4.

[41] The Congressional Black Caucus became the driving force behind restoring democracy in Haiti. *60 Minutes* at 2.

[42] However, a conservative backlash has recently emerged to curtail the broad interpretation and application of Section 2 of the Voting Rights Act. The case of Shaw v. Reno, 113 S. Ct. 2816 (1993), is the most publicized example of this movement against race-based districting. In *Shaw*, where white plaintiffs challenged the constitutionality of a majority-black district, the Supreme Court held that a cognizable 14th Amendment equal protection claim is stated when a district is "so irrational on its face that it can be understood only as an effort to segregate voters into separate districts because of their race." *Id.* at 2832. *See, e.g.*, Note, *White Lines, Black Districts:* Shaw v. Reno *and the Dilution of the Anti-Dilution Principle*, 29 HARV. C.R.-C.L. L. REV. 231 (1994). In fact, white plaintiffs have brought suits in several states claiming that creating majority-black districts violates 14th Amendment equal protection. *See, e.g.*, Hays v. State of Louisiana, 839 F. Supp. 1188 (1993). *See generally* David G. Savage, *Minority-Based Gerrymandering Facing Backlash*, L.A. TIMES, Oct. 8, 1994, at A1 ("Two years ago, 39 blacks were elected to the House of Representatives, a historic high-water mark that civil rights advocates saw as a culmination of the Voting Rights Act . . . . Now, however, a powerful backlash is sweeping through the federal courts, one that threatens to unseat many of those new minority representatives. In the last two months, judges in Louisiana and Texas and Georgia have struck down majority-minority districts as unconstitutional; districts in North Carolina and Florida remain under legal challenge."). *See also* Elaine R. Jones, *In Peril: Black Lawmakers*, N.Y. TIMES, Sept.

reform proposals unwittingly sacrifice more responsive government, expanded diversity of viewpoints, and substantial influence over federal policy for the sake of restraining the purported ill effects of "tokenism."[43]

Black institutional power furthers several of the goals Guinier seeks to advance in the first place. The goals of voter mobilization and post-election participation, for example, may be successfully achieved through the incidental benefits of institutional power.[44] Irrespective of race and ethnicity, the fundamental baseline for voter mobilization and post-election participation is the voter's belief that his or her vote counts. Even more significant in stimulating voter mobilization and post-election participation is the corresponding notion among voters that their *particular* vote made a difference in a given election's outcome. Arguably, the mere presence of black elected representatives, while neither ensuring substantive representation nor political concessions for black voters, makes blacks more politically aware and imbues in them a sense that their *particular* vote counts.[45]

Absent considerations of political viability,[46] the most provocative of Guinier's voting reform proposals is cumulative voting. By definition, cumulative voting increases the likelihood that minority voters can elect minority candidates, provided minority voters conscientiously choose to cast their votes strategically. At the very least, cumulative voting would promote minority voter mobilization by increasing the perception, as well

---

11, 1994, at E19 (summarizing recent litigation challenging the constitutionality of majority-black districts).

[43] For additional reflections on minority voting and political behavior, *see* Bernard Grofman & Lisa Handley, *Preconditions for Black and Hispanic Congressional Success*, *in* UNITED STATES ELECTORAL SYSTEMS: THEIR IMPACT ON WOMEN AND MINORITIES 32 (Wilma Rule & Joseph F. Zimmerman eds., 1992); *see also* HANES WALTON, JR., INVISIBLE POLITICS: BLACK POLITICAL BEHAVIOR (1985).

[44] *See, e.g.*, Milton D. Morris, *Black Electoral Participation and the Distribution of Public Benefits*, *in* MINORITY VOTE DILUTION 277 (1984) ("BEOs [Black Elected Officials] are the products of black votes, and they are in many respects the focus of black expectations for a more responsive government.").

[45] *See* Kathryn Abrams, *"Raising Politics Up": Minority Political Participation and Section 2 of the Voting Rights Act*, 63 N.Y.U. L. REV. 449, 519 (1988) ("[T]he election of minority representatives has been a crucial first step toward incorporating minority voters into the political process, a step that has brought minority participants political self-esteem.").

[46] Regardless of their benefits, questions regarding the political viability and degree of support for alternative remedies like cumulative voting necessarily involve examining, as well as challenging, fundamental assumptions of political participation and American politics. Articulating the particular ends sought constitutes only a small part of the entire exercise.

as the likelihood, that minority voters could readily elect representatives of their choice.

An additional benefit of cumulative voting, and perhaps even more critical in political terms than minority voter mobilization, is its ability to mitigate the resentment the majority typically feels in response to preferential voting rights policies for minorities. Guinier aptly notes that cumulative voting likely would avoid the incessant political opposition to race-based remedial measures. Cumulative voting may indeed calm the fears of many white voters that the divisive policy of affirmative action, traditionally reserved for employment and higher education, will not "infiltrate" the domain of the elections process. It may, therefore, escape the current political backlash surrounding majority-black districts.[47] Although unstated in her essays, this prediction may account for Guinier's unwillingness to view majority-black districts as a fail-safe resolution. In this regard, cumulative voting is Guinier's most appealing proposal in *The Tyranny of the Majority* since it circumvents the problems commonly associated with race-based remedies while still affording minorities a favorable chance to elect representatives of their choosing.

The Tyranny of the Majority has reinvigorated debate over the political quandary James Madison identified more than two centuries ago. Any well-functioning democracy to fulfill its ideals must guarantee that its governing and decisionmaking processes not only represent majority interests but also seek to reconcile the interests of minorities. As Madison once pointed out, providing the appropriate mechanism to do so is the challenge of the American experiment with democracy. *The Tyranny of the Majority* endeavors to meet that challenge.

—*Anthony Q. Fletcher*

---

[47] The Democratic party has increasingly become viewed as the "party of blacks." *See* Jerelyn Eddings et al., *Republicans whistle Dixie: Democrats battle demographics, economics, and a backlash against President Clinton*, U.S. NEWS & WORLD REP., Oct. 17, 1994, at 32. "The presidential campaigns of Jesse Jackson and the increased number of black Democrats in Congress, elected largely from black-majority districts drawn under voting rights guidelines, have strengthened [this] impression among whites . . . ." *Id.* at 34. *See also*, Joe Klein, *Bubba Is Back: Racial gerrymandering turns the New South ugly*, NEWSWEEK, Sept. 26, 1994, at 46 (noting that the creation of new "minority" congressional districts has further polarized politics in the South). *See also supra* note 42.

"HERE, THE PEOPLE RULE:" A CONSTITUTIONAL POPULIST MANIFESTO. By *Richard D. Parker*. Cambridge, Mass.: Harvard University Press, 1994. Pp. i, 132. $29.95 cloth, $14.95 paper.

America has a long tradition of carefully creating protections for the rights of the individual against the depredations of the many. In *"Here, the People Rule,"* Richard Parker argues that this tradition is an enervating, classist one that produces a culture of alienation among ordinary citizens. He suggests that America ought to shift its focus from protecting minorities to encouraging the majority's participation in politics. Unfortunately, Parker's critique of the liberal and allegedly "anti-populist" state is more interesting than his simplistic alternative—mass decisionmaking processes guarded by the courts. Parker fails to recognize that the intricate balance of power in the American polity not only serves the participation interests of the general public, but also takes into account other compelling social interests, such as protecting minorities and facilitating cost-effective government.

Parker's "manifesto" begins with two interpretations of Thomas Mann's symbolic short story "Mario and the Magician" (pp. 9–49).[1] Parker suggests that Mann's story, set in fascist Italy, evokes "some of our deepest, most problematic attitudes about the nature and peril of popular political energy in our own democracy" (pp. 9–10).

"Mario and the Magician" recounts the visit of the narrator, a bourgeois German, and his family to a crowded Tyrrhenian seaside village. The story culminates in an offensive performance given by a hunchback named Cippiola who entertains his boisterous audience by hypnotizing and then humiliating some of its members. The crowd does not protect these people from embarrassment; rather, it responds to the hypnotist's tricks with awe and amazement. To the narrator's dismay, Cippiola assumes the role of a classic demagogue.[2]

Parker claims that many readers interpret "Mario and the Magician" as a warning against the perils of the mob, signified by the easily led audience. Critics of the enthralled crowd invoke "the nightmare of fascism—mass energy arm-in-arm with abuse

---

[1] Thomas Mann, *Mario and the Magician, in* DEATH IN VENICE AND SEVEN OTHER STORIES 135 (H.T. Lowe-Porte, trans., Vintage Books 1936) (1929).

[2] *Id.* at 148–79.

of power, suppression of disfavored, nonconforming individuals, free-floating pugnacity, all in the name of a group identity" (p. 20). Parker concludes that such an interpretation suggests that "we are safe if a check by reason on power—by law on the febrile political energy of ordinary people—is built into institutions that give power expression" (p. 26). He observes that this interpretation, emphasizing the negative aspects of majoritarian power, has enjoyed popularity among the law students and professors with whom he has discussed the story (p. 67, n.32).

Parker, however, prefers a second interpretation that "look[s] at the narrator *as* the problem" rather than as a sympathetic victim of a crowded vacation and awful show (p. 27, emphasis in the original). Parker notes that the narrator responds to the stimuli of the crowd with a curious mix of haughtiness, deriding the citizenry for its boorishness and unsophistication while reacting passively to their offenses (p. 33). To Parker, this combination is inevitable, for "[t]he capacity—the energy—to assert and resist power can only decay in the absence of some concrete purpose rooted in a concrete connection to other people" (p. 41). Under this second interpretation, the moral of "Mario and the Magician" lies not in the power of the masses, but rather in the detachment of haughty bourgeois like the narrator from the people. "Freedom," Parker explains, "requires political mobilization of 'ordinary people' and for purposes of politics, you should embrace your lot as one 'ordinary' person among many, taking to heart the fear and hope that is the lot of us all" (p. 48).

Parker concludes that the two interpretations he offers for "Mario and the Magician" represent conflicting interpretations of the United States Constitution (pp. 53–115). He defines the dominant interpretation as "anti-populist." It espouses a mistrust and "genteel loathing" of popular political energy (p. 60). Anti-populists fear that mass political participation risks, at best, common, "vulgar," "reckless," and "abusive" politics and, at worst, the suppression of minorities (pp. 58, 69–70). Anti-populists respond to an environment imbued with political energy of this sort either with dignified withdrawal from the political arena, or with an attempt to transform that arena—they "try to contain or to retard, to tame or to manipulate, the forces of ordinary politics," by elevating the process above ordinary people (p. 60).

Parker claims that anti-populists have set many of the parameters of modern constitutional debate.[3] He characterizes their contribution as one that combines empty, somewhat disdainful flat-

tery of the concept of majority rule with a fixation on the "correction of failures allegedly endemic to majority rule" (pp. 66–70). These corrections are made by a judiciary removed from ordinary people. Parker characterizes the judiciary's opinions as both too complex and overly concerned with protecting a "fragile" constitution against the depredations of the many (pp. 74–76).[4] Parker claims that the resulting system has created a governing class that combines "leadership" of the people with a strong distrust of them (pp. 78–93).[5]

Instead, Parker suggests that the country should promote majority rule (pp. 94–115). Majorities—if organized in a grassroots fashion—are "courageous," "vigorous," and "spontaneous" (p. 63).[6] Parker argues that these qualities helpfully inform decisionmaking. Therefore, substantive social decisions about rights and entitlements should be examined only to see if they were made through a process that would foster popular participation in politics, ensuring that society would benefit from majority involvement (p. 105). Toward that end, Parker would create a new set of constitutional claims to protect populism. For example, he would encourage openness by prosecuting such claims against officials who fail adequately to make their decisionmak-

---

[3] Fear of the majority has a rich constitutional history. For example, in 1905 the United States Supreme Court protested rhetorically "[A]re we all at the mercy of legislative majorities?" as it infamously overruled one. *Lochner v. New York*, 198 U.S. 45, 59 (1905).

[4] Parker notes that an increasing number of judges have been recruited from law school faculties (p. 75). He argues that these judges produce complicated rulings that ordinary Americans find difficult to penetrate.

[5] Parker's conclusion that judges are removed from ordinary experience has been seconded by other literary legal analysts. Richard Delgado and Jean Stefancic argue that judges' "inability to identify imaginatively, with the person whose fate is being decided" can be rectified by the study of Law and Literature. Richard Delgado and Jean Stefancic, *Norms and Narratives: Can Judges Avoid Serious Moral Error?* 69 TEX. L. REV. 1929, 1930 (1991). "Members of the movement believe that by reading lawyers and judges may gain empathy through vicarious experience." *Id.* at 1931. *See also*, John Fischer, Note, *Reading Literature/Reading Law: Is There a Literary Jurisprudence?* 72 TEXAS L. REV. 135, 136–37 (1993) ("One of the primary claims often made in that movement is that literature teaches us in a uniquely *ethical* manner because of its power to bind the lawyer to the larger community of which she is a part") (emphasis in the original).

[6] A preference for vigor, spontaneity, and a lack of inhibition is nothing new. *See* FRIEDRICH NIETZSCHE, ON THE GENEALOGY OF MORALS 157–59 (Walter Kaufmann & R.J. Hollingdale trans., 1967) (1887). Parker's view of the dominant anti-populist paradigm of constitutional law, drawing on a tradition opposed to emotion and impulsiveness, also finds parallels in Nietzsche. Nietzsche argued that "from a historical point of view, law represents on earth the struggle *against* the reactive feelings." *Id.* at 75 (emphasis in the original). However, Nietzsche's solution to such an oppressive law embraced vigor and spontaneity in a rather anti-populist sense. He urged the cultivation of a noble—as distinct from a mass—ideal. *See id.* at 27–35.

ing processes accessible to the citizenry (pp. 100–01). Also, his theory would promote equality and participation by exposing hierarchical, undemocratic organizations (including businesses, perhaps) to new levels of judicial scrutiny (pp. 102–04).

Judicial review would therefore protect the populist process. However, society's most important decisions would be made through democratic politics, instead of legal proceedings (p. 106). This majoritarian mechanism would therefore replace the legal system as the arbiter of the constitutional value of our rights and entitlements. Courts, rather than defining our constitutional prerogatives, would become the defenders of populist decisionmaking methods. Parker concludes that constitutional law should not be viewed as some kind of "higher law," rather it is only "a political controversy about democracy" that will evolve over time as the concerns of the constituents of that democracy change (p. 109).

Parker usefully identifies the costs of a constitutional order that places some issues, such as limits on free speech and abortion rights, beyond the control of the majority. This manner of regulating an individual's vote devalues not only democratic participation in a general sense but also the vote itself. Subjecting decisions of the majority to the constant scrutiny of the judiciary surely demoralizes voters. Constitutional regulation also rewards the select groups, such as the legal profession, that define the constitutional order. Unfortunately, however, Parker's populist solution offers little guidance on how we might reconcile our constitutional order with vigorous popular politics. More fundamentally, his analysis ignores both the costs of unrestricted majority rule and the balanced majoritarianism that already exists in the American polity.

Parker neglects to explore his claim that individuals personally benefit from participating in a group. His populists avoid being "other-directed as opposed to inner-directed" and "diffident as opposed to self-confident" (p. 63). But it is not obvious that group membership will in fact boost independence and self-confidence. Groups, of course, may create unhealthy dependencies and a strong pressure to conform.[7] Moreover, they can inhibit creativity and expression. But Parker does not consider these possibilities.

---

[7] Cults or gangs, for example, may cultivate senses of belonging among their members but do not necessarily inspire individualism.

If Parker's conclusions are unsatisfying on a personal level, they may be unworkable on a constitutional level. Parker urges the measurement of all rule-making procedures against one standard—the degree to which the procedure facilitates participation. But maximizing participation could make it impossible for society to accomplish anything. Indeed, the public might be interested in trading a portion of its right to rule to bureaucracies that grease the wheels of governance. For instance, people could conceivably barter participation for efficiency, perhaps supplied by an executive. Parker, however, does not seem to allow for such trade-offs within his social contract.

Instead of considering these possibilities, Parker reifies process-oriented constitutionalism. Any value passes constitutional muster for Parker as long as it was chosen by a vocal and active majority. This substitution of democratic process for the articulation of substantive rights, however, is disingenuous. One can easily imagine majoritarian decisions that effectively curtail vocal and extensive debate or universal political participation. One need only imagine a horrible substantive decision made by a majority (say, racial segregation) to recognize the flaws of a solely procedural constitutional theory. As Laurence Tribe has observed, "the constitutional theme of perfecting the processes of governmental decision is radically indeterminate and fundamentally incomplete."[8] Parker's scheme is certainly flawed in this respect, given that he fails to offer a well-articulated set of values that could inform his majority decision-makers.[9] For him, a democratic procedure alone will suffice.

It is possible that such a procedure would actually inhibit widespread involvement in political life. Strong majority rule, after all, might demoralize minority participation in the political process, considering that minorities could reasonably believe that their concerns would carry no weight. Conversely, in a heterogeneous and diverse culture, we might maximize participation precisely by limiting the possibilities for majority rule,

---

[8] Laurence Tribe, *The Puzzling Persistence of Process-Based Constitutional Theories*, 89 YALE L.J. 1063, 1064 (1980). Parker himself has criticized process-oriented theory for its classism. *See* Richard D. Parker, *The Past of Constitutional Theory—And its Future*, 42 OHIO ST. L.J. 223 (1981).

[9] Tribe argues that process theory needs to be supplemented by a full theory of substantive rights and values. *See* Tribe, *supra* note 8, at 1064. Parker, however, claims only to offer a "sensibility" and not a "set of principles" (p. 94).

so that each group could rest assured of a voice worth exercising in popular politics.

Risking this sense of minority disenfranchisement by further empowering majority rule may well be unnecessary because part of the American government already recognizes majoritarian, participatory values. Congress is undoubtedly the kind of populist entity that Parker glorifies. It acts by majority vote after an extensive process of debate, hearings, and comment.[10] It is recomposed by voters every two years and is institutionally structured to stay in relatively close contact with constituents, as Members' franking privileges[11] and large staffs[12] attest. Congress is therefore reasonably accessible to everyday people and acts openly with an eye toward soliciting the approval of the public. Its popular, participatory framework helps to offset Parker's concern that one of the legislature's counterpart branches, the judiciary, is primarily an elitist institution.

To be sure, Congress cannot serve as an institutional response to all of Parker's worries. The seniority system and rigid procedural rules of Congress limit the internal democratic possibilities of the institution.[13] These possibilities are also externally limited by a judiciary that responds to minority concerns and by an executive that often values the administrability of policies more highly than popular participation in policy-making. Nonetheless, Congress as an institution stands as an important counter-example to Parker's notion that majority rule is only granted "courtly disdain" in American political discourse (p. 68).

While Parker usefully analyzes constitutional debate as a contest between populists and anti-populists and presents a clear and simple solution to the problem allegedly generated by anti-popu-

---

[10] The Senate features especially generous rules encouraging a free, open, and unregulated debate. In addition, while the Senate has been criticized in the past for allegedly placing too much authority in the hands of a few leaders via its seniority system, evidence indicates this is much less of a problem today. Even ten years ago, the *New York Times* reported that "[a] decade of diffusion of authority has steadily eroded the powers of seniority and leadership" in the Senate. Tolchin, *Senators Assail Anarchy in New Chamber of Equals*, N.Y. TIMES, Nov. 25, 1984, at 40.

[11] The statutory franking privilege gives Members of Congress and the Vice President the power to send mail free of charge through the postal service to their constituents "in order to assist and expedite the conduct of the official business, activities, and duties of the Congress of the United States." 39 U.S.C. § 3210(a)(1) (1992).

[12] Over half of the work done by personal staffs may involve constituency affairs. MORRIS P. FIORINA, CONGRESS: KEYSTONE OF THE WASHINGTON ESTABLISHMENT 53–60 (1989).

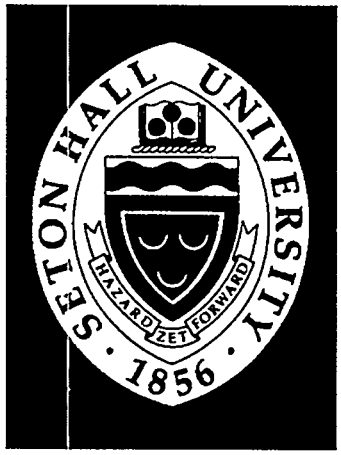
[13] Conservatives also criticize some of the more rigid Congressional structures. *See,*

lism, he goes too far. In reducing America's delicate constitutional balances to a mantra of participatory politics, Parker ignores the complexities of constitutional rule. Parker suggests that the populist and anti-populist paradigms wage a constitutional war of absolutes—that one can view the constitutional landscape only through one lens or the other, but not both. In reality, populist impulses, reflected by such institutions as elections and legislatures, coexist in America with other less democratic notions that similarly find expression in such well-established procedures as judicial review. Viewing American constitutional structure as an amalgam of populism and anti-populism does not make for easy constitutional analysis, but it accurately captures the ideological compromises that Americans make to achieve a functional society. Although populist impulses surely appeal to a democratic culture, that culture can only prosper when society recognizes that the excesses of majority rule are occasionally too expensive. Populism and anti-populism should thus be analyzed as two intellectual poles on the continuum within which political compromises are made. In this regard, Parker's clash between ideological absolutes mischaracterizes the American reality and the advantages it offers. That reality can better be understood as an uneasy coexistence of populist outlets checked by some anti-populist procedures.

—*David T. Zaring*

---

e.g., GEORGE F. WILL, RESTORATION: CONGRESS, TERM LIMITS, AND THE RECOVERY OF DELIBERATIVE DEMOCRACY (1992).



SETON HALL LEGISLATIVE JOURNAL

The *Seton Hall Legislative Journal* is a nationally recognized periodical that focuses on legislative, statutory, and regulatory topics. Authors include state and federal legislators, judges, and members of the academic and practicing communities.

Our upcoming issue, Volume 19, Book 1, includes articles on independent counsel, federal boxing regulation, the New Jersey death penalty, and national health care reform.

The subscription rate is $20.00 per two-issue volume and individual issues are available at $12.00 per copy.

Subscription Manager
Seton Hall Legislative Journal
Seton Hall University School of Law
One Newark Center
Newark, NJ 07102
(201) 642-8261

The *Seton Hall Legislative Journal* welcomes the submission of articles, book reviews, essays, and perspectives.