

1-2003

Daubert & Danger: The "Fit" of Expert Predictions in Civil Commitments

Alexander Scherr

Follow this and additional works at: https://repository.uchastings.edu/hastings_law_journal



Part of the [Law Commons](#)

Recommended Citation

Alexander Scherr, *Daubert & Danger: The "Fit" of Expert Predictions in Civil Commitments*, 55 HASTINGS L.J. 1 (2003).

Available at: https://repository.uchastings.edu/hastings_law_journal/vol55/iss1/1

This Article is brought to you for free and open access by the Law Journals at UC Hastings Scholarship Repository. It has been accepted for inclusion in Hastings Law Journal by an authorized editor of UC Hastings Scholarship Repository. For more information, please contact wangangela@uchastings.edu.

Articles

Daubert & Danger: The “Fit” of Expert Predictions in Civil Commitments

ALEXANDER SCHERR*

INTRODUCTION

“Never make predictions, especially about the future.”¹ But in civil commitments, courts do just that, predicting the future behavior of a mentally ill individual. Judicial prediction can have severe results for that person: deprivation of liberty for an indeterminate time; potentially unwanted medication and physical restraint; and the stigma of mental illness. Judicial inaction also poses risks: Erroneous release can lead to harm for the person or for others. Resolving these risks requires determining whether the person poses a danger to self or to others because of the person’s mental illness.² Such illnesses force fact-finders to come to terms with disturbing, chaotic, and unfamiliar mental conditions. To understand these conditions requires knowledge, insight, discipline, and professional balance. In short, it requires expertise.

* Assistant Professor, University of Georgia School of Law. I thank Dan Coenen, John Monahan, Ron Carlson, and Randy Beck for their generous and very useful comments on a draft of this piece. Thanks also go to Rebecca Aubin, Sebastian Suma, and Lisa Taylor for diligent research and thoughtful reflection at various phases of this project. Finally, I express my appreciation to Dean David Shipley both for his direct support of my research, and for his support of the clinical programs under my care.

1. Various attributed to Yogi Berra, Samuel Goldwyn, Casey Stengel, and others. See Gary John Previats, Comment, *Global Multi-Disciplinary Practice: A Word on “The Future”*, 52 CASE W. RES. L. REV. 947, 959 (2002) (attribution to Goldwyn); Eugene Volokh, *Technology and the Future of Law*, 47 STAN. L. REV. 1375, 1403 n.156 (1995) (reviewing M. ETHAN KATSH, *LAW IN A DIGITAL WORLD* (1995)) (citing attributions to Berra, Goldwyn, Stengel, Will Rogers, Leo Durocher, Mark Twain, and Niels Bohr); James A. George, *Wilander—Light at the End of the Labyrinth*, 16 TUL. MAR. L.J. 131, 164 n.208 (1991) (attribution to Stengel).

2. *Addington v. Texas*, 441 U.S. 418, 432–33 (1979) (requirement of ‘clear and convincing’ evidence); *O’Connor v. Donaldson*, 422 U.S. 563, 575–76 (1975) (requirement of danger); see *infra* Part II.C.5 (discussing legal definitions of danger in civil commitment actions).

The opinions of experts in prediction should help the courts in this task, but over thirty years of commentary, judicial opinion, and scientific review argue that predictions of danger lack scientific rigor. The United States Supreme Court has commented regularly on the uncertainty of predictive science.³ The American Psychiatric Association has argued to the Court that “[t]he professional literature uniformly establishes that such predictions are fundamentally of very low reliability.”⁴ Scientific studies indicate that some predictions do little better than chance or lay speculation, and even the best predictions leave substantial room for error about individual cases.⁵ The sharpest critique finds that mental health professionals perform no better than chance at predicting violence, and perhaps perform even worse.⁶

Given these critiques, we should expect the rules of evidence, and specifically the reliability standards of *Daubert v. Merrell Dow Pharmaceuticals, Inc.*,⁷ to require the exclusion of predictive expertise from the civil commitment process. *Daubert* overturned the seventy-year-old *Frye v. United States*⁸ standard for admissibility of expert opinion, replacing it with a standard focused (at least in part) on the scientific reliability of expert opinion.⁹ Civil commitment is a matter of state law, governed by state rules of evidence, and not all states have followed the Supreme Court’s lead.¹⁰ But this should not matter: Given their notorious unreliability and deep division in the professional community, we might predict that no court would admit predictive opinions either under *Daubert* or under *Frye*.

Yet no appellate court has ever ordered exclusion of expert psychiatric testimony about danger in a civil commitment case, either before or after *Daubert*. Courts have shown an extraordinary receptiveness to such opinions, admitting and relying on them in their commitment decision-making. Moreover, courts have had their eyes open: Judicial opinions regularly refer to, and explicitly accept, the imperfections of predictive

3. See *infra* Parts I.B.2, I.C.1, and I.C.3 (discussing various aspects of Supreme Court jurisprudence dealing with psychiatric predictions of danger).

4. Brief Amicus Curiae for the American Psychiatric Association at 8, *Estelle v. Smith*, 451 U.S. 454 (1981) (No. 79-1127).

5. See *infra* Part I.B.1 (describing clinical and actuarial forms of predictive expertise, as well as the scientific assessments of the relative validity of these techniques).

6. Bruce J. Ennis & Thomas R. Litwack, *Psychiatry and the Presumption of Expertise: Flipping Coins in the Courtroom*, 62 CAL. L. REV. 693, 734-35 (1974).

7. 509 U.S. 579, 589-92 (1993).

8. 293 F. 1013, 1014 (D.C. Cir. 1923) (“sufficiently established to have gained general acceptance in the particular field in which it belongs”).

9. See *Daubert*, 509 U.S. at 592-93.

10. See *infra* Part II.A (discussing admissibility of expert predictive testimony in states using *Daubert*, *Frye* and other standards).

testimony.¹¹ What is going on? How can such unreliable opinions survive *Daubert's* stress on scientifically valid and reliable expertise?¹²

The answer is simple: *Daubert* requires more than assessing scientific reliability. Trial courts must also assess how the expert testimony fits the demands of the case.¹³ Even without scientific validation, a court may still use a given opinion if it has a sufficiently strong fit to the demands of the case.¹⁴ This Article uses the example of predictive expertise both to develop a methodology for assessing *Daubert's* notion of "fit," and to find factors for applying that standard. The methodology appraises civil commitment for its characteristic features: its substantive standards; its burdens of proof; its patterns of proof for danger, including expert testimony; and the legal definition of danger. This methodology helps to identify four factors for assessing the fit of expertise in a given case,¹⁵ and to ratify the extraordinary fit between predictive opinions and civil commitments.

The argument proceeds in two phases. Part I discusses how *Daubert* and its progeny have articulated a dual standard of reliability and fit. The first section concludes that, under *Daubert*, the "reliability" of an opinion means not just its scientific validity, but also its practical utility for resolving disputes. It also concludes that *Daubert's* test for "fit" asks not just about bare relevance; it also suggests a comparison of the expert's specialized inferences to the inferences required for fact-finding.¹⁶ The second section then reviews the available modern standards for validating predictive expertise, and concludes that the relevant scientific community has reached a state of guarded optimism about its reliability for judicial

11. See *infra* Part I.C (discussing the role that predictive uncertainty plays in civil commitments).

12. The related question asks, how can such vilified expertise survive *Frye's* mandate of general acceptance? See *infra* Parts II.A.2–3, and notes 370, 386 (discussing the admissibility of predictive testimony under *Frye*).

13. Other explanations exist, most prominently the idea that the imperatives of civil commitment justify a waiver of evidentiary rules. See I DAVID L. FAIGMAN ET AL., MODERN SCIENTIFIC EVIDENCE: THE LAW AND SCIENCE OF EXPERT TESTIMONY § 9-1.2 (2002) (arguing that either explicit statutory suspension or tacit waiver of the evidentiary rules best explains the universal admissibility of predictive opinions). "[C]ourts ignore evidence rules that would otherwise exclude testimony on future violence because of a sense of imperative created by the substantive law in these areas." *Id.* This Article advances the notion that the rules of evidence and specifically the standards of *Daubert* provide coherent justification for admitting these opinions.

14. A complete explanation of all different combinations of reliability and fit, while productive, goes well beyond the scope of this Article.

15. See *infra* Part I.C. The four factors include: how far courts have absorbed concerns over reliability in shaping the structure and process of a case; the inherent difficulty of fact-finding on the central issue, and the extent to which the expertise eases that difficulty; the prevalence of the particular expertise in litigating the case; and the similarity of the inferential process embodied in the opinion to the fact-finding required for resolution.

16. See *infra* Part I.A.

use.¹⁷ Finally, the third section examines the substance and process of civil commitment, assessing how predictive expertise fits the features of commitment, so as to articulate with specificity elements of *Daubert's* fit criterion.¹⁸

Part II turns to evidentiary doctrine, analyzing those cases that have admitted expert predictive opinion in civil commitment proceedings, whether the jurisdiction uses tests from *Frye*, *Daubert* or no test at all. The Part confirms the widespread receptivity of courts to expert opinion evidence about future danger. It also critiques the existing rationales for those results, and proposes a rationale that fully explains the case law and accommodates developments in predictive science.¹⁹ This Article concludes that *Daubert* should permit the admission of expert predictions, but not because of their scientific reliability. Instead, the Article uses the example of civil commitments to develop the notion that *Daubert's* concern with evidentiary fit better explains the courts' receptivity to this form of predictive testimony.

A need exists for this analysis. Most states have yet to rule on the admissibility of expert predictive opinions in commitment cases.²⁰ More broadly, *Daubert* and its successor cases are in danger of being viewed

17. See *infra* Part I.B.

18. See *infra* Part I.C.

19. See *infra* Part II.C.

20. General civil commitment remains a fundamental concern of most states. After the Supreme Court's decision in *Kansas v. Hendricks*, 521 U.S. 346, 370–71 (1997) (upholding a Kansas "sexual predator" statute against constitutional challenge), many states have renewed their focus on special commitment statutes focused on particular classes of individuals with mental disorders. As will be seen, the most recent cases to have ruled on the admissibility of predictive testimony have done so in sexually violent offender commitments. See *infra* Part II.B (discussing cases to have ruled on the admissibility of these opinions). Moreover, the most recent statistical methods for assessing risk actuarially apply similar theories and methods to the assessment of criminals, the mentally ill, and sexually violent offenders. See also Mara Lynn Krongard, Comment, *A Population at Risk: Civil Commitment of Substance Abusers After Kansas v. Hendricks*, 90 CAL. L. REV. 111, 131 (2002) (arguing that the decision in *Kansas v. Hendricks* may permit the expansion of types of special commitment statutes to include substance abusers).

solely as cases on scientific reliability.²¹ This Article counteracts such a view, suggesting that *Daubert's* reliability requirement rests on more than science, that its fit requirement requires exploring how specialized inferences support judicial fact-finding, and that reliability and fit exist in a dynamic relationship. Put another way, there is need for a refined model of *Daubert*, and the cases on prediction of danger help point the way towards construction of that model.

I. EXPERT PREDICTIONS IN CIVIL COMMITMENTS

This Part explores the evidentiary, scientific, and legal dimensions of the use of expert predictions of danger in civil commitments. It begins with a brief assessment of *Daubert* and *Kumho Tire Co. v. Carmichael*,²² continues with an assessment of the science of predicting danger, and ends with an exploration of the use of expert predictions in civil commitment case law. This Part thus lays the groundwork for discussing the “fit” requirement and leads to an assessment of the evidentiary case law in Part II.

A. EXPERT TESTIMONY

1. *Expert Testimony Before Daubert*

The admissibility of expert testimony represents a special application of customary rules differentiating fact from opinion. Opinions consist of collections of inferences conveyed to the fact-finder as “facts,” and include two mutually exclusive sub-categories: lay opinion and expert opinion.²³ In functional terms, expert witnesses can testify in ways that lay witnesses cannot. Lay witnesses must describe the factual bases for their

21. See, e.g., Erica Beecher-Monas & Edgar Garcia-Rill, *Danger at the Edge of Chaos: Predicting Violent Behavior in a Post-Daubert World*, 24 CARDOZO L. REV. 1845, 1856 (2003) (addressing the use of predictive opinion in death penalty cases, and arguing that *Daubert's* requirement of scientific validity should inform the constitutional admissibility of predictive expertise). “[R]elevance and reliability in the context of scientific testimony now requires scientific validity.” *Id.* at 1853. Beecher-Monas and Garcia-Rill argue that *Daubert* would lead to the exclusion from evidence of clinical predictions of violence in death penalty cases, but should permit admission of actuarial assessments in those cases, at least for limited purposes. *Id.* at 1897. They assume that *Daubert* requires no more than “scientific validity,” and that *Daubert's* “fit” component only a minimal connection between an expert’s methodology and conclusion. *Id.* at 1857–58. “Unless an expert can demonstrate sound methodology and scientific reasoning, no opinion testimony is admissible.” *Id.* at 1858.

The present Article reaches an opposite conclusion, at least with respect to civil commitments. In so doing, it focuses more on the notion of “fit” as a test rooted in the courts’ role as pragmatic decision-makers, rather than in a presumed role for courts as validators of scientific reliability.

22. 526 U.S. 137 (1999).

23. See FED. R. EVID. 701 & 702. Under the current version of the Federal Rules, lay opinion refers generally to opinions “not based on scientific, technical, or other specialized knowledge.” FED. R. EVID. 701(c). The quoted language dovetails with that found in Rule 702 governing the admissibility of expert testimony.

opinions; expert witnesses need not (and may not be permitted) to do so.²⁴ The lay witness may only testify to opinions that are “rationally based” on their own perception; the expert witness is not necessarily so limited.²⁵ Finally, the information on which a lay witness relies (the basis for the opinion) must itself be admissible evidence; by contrast, the expert may use inadmissible information, provided it is “of a type reasonably relied upon by experts in the particular field in forming [comparable] opinions.”²⁶

Taken together, these functional differences shift the burden of drawing inferences about critical facts away from the fact-finder and towards the expert. A fact-finder listening to a lay witness hears all of the factual predicates for the opinion, which rest upon the witness’s firsthand knowledge and bear a “rational” relationship to the opinion. By contrast, a fact-finder listening to an expert hears an opinion without necessarily hearing all of the factual predicates, some of which may be inadmissible. For experts, no familiar connection need exist between predicate facts and their opinions; expert inferences are by definition specialized, and potentially unfamiliar to the fact-finder. Finally, by reason of professional or intellectual status or credentials, an expert’s opinions may carry greater weight. These functional differences (coupled with the expert’s status) have led courts to perceive that fact-finders might attribute an “aura of certainty” to expert testimony, justifying special rules of admissibility.²⁷

Frye represented the past century’s most prevalent standard of special admissibility.²⁸ Indeed, *Frye* itself dealt with a form of behavioral or psychic science, the detection of lies based on purported changes in blood pressure resulting from specific psychic states. The proponent had argued standard evidentiary doctrine: The interpretation of blood pressure changes “does not lie within the range of common experience or common knowledge, but requires special experience or special knowl-

24. FED. R. EVID. 703. “[A]n expert is permitted wide latitude to offer opinions, including those that are not based on firsthand knowledge or observation.” *Daubert v. Merrell Dow Pharm., Inc.*, 509 U.S. 579, 592 (1993).

25. FED. R. EVID. 701(a) & 702.

26. FED. R. EVID. 703 (“Facts or data that are otherwise inadmissible shall not be disclosed to the jury by the proponent of the opinion or inference unless the court determines that their probative value . . . substantially outweighs their prejudicial effect.”).

27. *People v. Kelly*, 549 P.2d 1240, 1245 (Cal. 1976) (stating that *Frye* ensures that the “aura of certainty” which can attach to expert scientific testimony does not attach to untested, experimental evidence).

28. *Frye v. United States*, 293 F. 1013, 1014 (D.C. Cir. 1923) (excluding testimony based on a “systolic blood pressure deception test”).

edge."²⁹ Rejecting this argument, the court articulated what came to be the standard form of the rule: "[W]hile 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."³⁰

Frye left many questions unanswered.³¹ The most salient include: which kind of expertise falls within *Frye*'s ambit; which part of the professional community must a trial court consult (or, what is the "particular field" to which the expertise belongs); and exactly what kind of consensus must exist to support a finding of "general acceptance." Many states still retain *Frye*; and courts in at least four of them have used it to review the admissibility of expert predictions in civil commitment cases.³²

2. Daubert & Kumho

Daubert and its two companion cases, *General Electric Co. v. Joiner*³³ and *Kumho Tire*,³⁴ changed *Frye*'s "general acceptance" standard to one more attuned to the merits of the particular expertise itself. This subsection articulates the parameters of *Daubert* and *Kumho Tire*'s new approach.³⁵ The next subsection addresses questions left open by these cases, especially those related to the "fit" requirement.

a. Daubert

Daubert arose out of a dispute over the admissibility of scientific expertise. Relying on *Frye*, the trial court had excluded affidavits from certain experts offered in summary judgment proceedings.³⁶ The Supreme

29. *Id.* Defendant sought to admit the expertise in evidence, and the trial court rejected the proffer; the appellate court affirmed. *Id.*

30. *Id.*

31. See Paul C. Giannelli, *The Admissibility of Novel Scientific Evidence: Frye v. United States, a Half-Century Later*, 80 COLUM. L. REV. 1197, 1250 (1980) (reviewing the history and progress of the *Frye* standard). See also PETER W. HUBER, GALILEO'S REVENGE: JUNK SCIENCE IN THE COURTROOM 204 (1991) (critiquing judicial handling of novel scientific evidence prior to *Daubert*, and suggesting a renewed emphasis on *Frye*'s approach).

32. See *infra* Part II.B.1 and II.B.3 (discussing direct application of *Frye* to predictive testimony, and the exemptions from *Frye* of the same).

33. 522 U.S. 136 (1997).

34. *Kumho Tire Co. v. Carmichael*, 526 U.S. 137 (1999).

35. In *General Electric*, the Court applied an "abuse of discretion" standard to review trial court decisions. 522 U.S. at 146-47. The Court rejected plaintiff's argument that "outcome-determinative" opinions should receive stricter review. *Id.* at 142-43. In applying the long-standing abuse of discretion standard, the Court placed the burden of assessing the inferences implicit in opinion testimony squarely in the hands of trial judges. *Id.* at 141-43, 146-47. The case accepts and asserts the primacy of the trial judge's discretion in ruling on the admissibility of expert testimony. *Id.*

36. *Daubert v. Merrell Dow Pharm., Inc.*, 509 U.S. 579, 583-84 (1993). The factual issue before trial centered on proof of past causation, specifically proof that ingestion of Bendectin by women caused birth defects in their children. *Id.* at 582. In support of their motion for summary judgment,

Court reversed and remanded, holding that *Frye* had not survived the enactment of the Federal Rules of Evidence.³⁷ The Court then articulated a new evidentiary standard for the admissibility of expert testimony, a test focused on reliability and fit.³⁸

In rejecting *Frye*, Justice Blackmun wrote that the language and premises of the rules required trial judges to “ensure that any and all scientific testimony or evidence admitted is not only relevant, but reliable.”³⁹ The opinion discussed science,⁴⁰ but did not restrict “evidentiary reliability” to scientific validity. Rather,

in order to qualify as “scientific knowledge,” an inference or assertion must be derived by the scientific method. Proposed testimony must be supported by *appropriate validation*—i.e., “good grounds,” based on what is known. In short, the requirement that an expert’s testimony pertain to “scientific knowledge” establishes a standard of *evidentiary reliability*.⁴¹

defendants had offered an affidavit from an expert, who concluded that “maternal use of Bendectin during the first trimester of pregnancy [had] not been shown to be a risk factor for human birth defects.” *Id.* In opposition, plaintiffs offered affidavits from eight experts, all of whom concluded that “Bendectin can cause birth defects.” *Id.* at 583. Plaintiffs’ experts based their conclusions on review of studies on animal fetuses and on the chemical similarities between Bendectin and other substances “known to cause birth defects,” as well as on a reanalysis of prior epidemiological studies. *Id.* The trial court used the *Frye* standard to conclude that “expert opinion which is not based on epidemiological evidence is not admissible to establish causation.” *Id.* at 583–84. See *Daubert v. Merrell Dow Pharm., Inc.*, 727 F. Supp. 570, 573–76 (S.D. Cal. 1989). The Ninth Circuit affirmed. See *Daubert v. Merrell Dow Pharm., Inc.*, 951 F.2d 1128, 1130–31 (9th Cir. 1991).

37. *Daubert*, 509 U.S. at 589. Justice Blackmun compared *Frye*, which predated enactment of the Federal Rules of Evidence, and relied in part on other cases which had held that common law doctrines had not survived enactment. See *United States v. Abel*, 469 U.S. 45, 51 (1984); *Bourjaily v. United States*, 483 U.S. 171, 177 (1987). It found no reference to the “general acceptance” standard in the text of Federal Rule of Evidence 702, nor in the drafting history. *Daubert*, 509 U.S. at 588. Finally, it described *Frye* as a “rigid” test, inconsistent “with the ‘liberal thrust’ of the Federal Rules and their ‘general approach of relaxing the traditional barriers to ‘opinion’ testimony.’” *Id.* (quoting *Beech Aircraft Corp. v. Rainey*, 488 U.S. 153, 169 (1988)).

38. On remand, the Ninth Circuit applied the new standard, articulating factors additional to those identified by the Supreme Court. *Daubert v. Merrell Dow Pharm., Inc.*, 43 F.3d 1311, 1317–22 (9th Cir. 1995). But instead of remanding the case, it affirmed the exclusion of the plaintiff’s proffered expert testimony because of an inadequate fit between fact-finding and expertise. *Id.*

39. *Daubert*, 509 U.S. at 589.

40. The parties had litigated the case at least in part on the premise that *Frye* no longer had authority as a test for admissibility of “scientific” evidence. The Court was not presented with an issue about non-scientific expertise, an issue to which it would turn later. See *infra* Part I.A.2.b (discussing *Kumho Tire*).

41. *Daubert*, 509 U.S. at 590 (emphasis added). The quoted language arguably leaves ambiguous whether “scientific method” overtakes and becomes the measure for reliability of “scientific” expertise. On the one hand, the language states that the opinion “must be derived [from] scientific method.” *Id.* On the other hand, the language requires that it be supported by “appropriate validation”; and it notes that the rules establish “a standard of evidentiary reliability.” *Id.* Whatever ambiguity the *Daubert* opinion left, however, the Court later resolved in *Kumho Tire*. See *infra* Part I.A.2.b.

The Court saw reliability as an “evidentiary” standard, rooted in the pragmatic demands of trial, and not tied solely to scientific validity. The Court expanded on *Frye* and articulated four factors to use in assessing reliability, including “general acceptance.” Moreover, the Court presented these as “general observations” or “pertinent consideration[s],” not as exclusive points of reference for every case.⁴²

The Court’s use of “relevance” also rests on evidentiary and not scientific grounds.⁴³ Not restricted to simple logical relevance, the Court’s opinion focuses more on a notion of “fit” or “appropriateness.”⁴⁴ It cites an important pre-*Daubert* case for the proposition that expert testimony must be “sufficiently tied to the facts of the case that it will aid the jury in resolving a factual dispute.”⁴⁵ According to the Court, the notion of fit was not absolute, but functional and relativistic: “‘Fit’ is not always obvious, and scientific validity for one purpose is not necessarily scientific validity for other, unrelated purposes.”⁴⁶

The *Daubert* Court thus fashioned an overarching standard with the potential to go beyond science per se. It used the language of “scientific method,” but its articulated factors neither excluded other methods nor applied to all kinds of expertise. Its standard focused primarily on the relationship between expert opinion and the fact-finding it would influence. It sought to assure that such opinions would cast a spell on the fact-finder only *after* the trial court had found a sufficient connection between the expert inferences and the facts at issue.

42. *Daubert*, 509 U.S. at 593. The four considerations it discussed were: whether the theory or technique has been or can be tested; whether it has been subjected to peer review; what the known or potential error rate of inferences drawn from the technique might be; and whether it has in fact received “general acceptance.” *Id.* at 593–94. These four “*Daubert* factors” threatened to take on a life of their own; references to the “four part test” of *Daubert* were occasionally seen. This restrictive understanding of *Daubert* proved not to survive close examination by the Court. *See infra* Part I.A.2.b (discussing *Kumho Tire*).

43. *See* FED. R. EVID. 401 (“‘Relevant evidence’ means 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.”). This definition denotes logical relevance, concerned solely with the existence of an inference to be drawn from the proposed evidence and a ‘fact of consequence’ in the litigation.

44. The Court derives this concept from the stated requirement in Federal Rule of Evidence 702 that the expert’s knowledge “assist the trier of fact,” a notion of “helpfulness.” *See* discussion *infra* Part II.C.5 (discussing the fit of expert predictions to civil commitments) and II.B (assessing doctrinal approaches to the fit of clinical and actuarial opinions).

45. *Daubert*, 509 U.S. at 591 (quoting *United States v. Downing*, 753 F.2d 1224, 1242 (3d Cir. 1985)).

46. *Id.* (citing James E. Starrs, *Frye v. United States Restructured and Revitalized: A Proposal To Amend Federal Evidence Rule 702*, 26 JURIMETRICS J. 249, 258 (1986)).

b. Kumho Tire

In *Kumho Tire*,⁴⁷ the U.S. Supreme Court rejected the views that *Daubert* dealt only with scientific evidence and that the four *Daubert* factors provide necessary tests of the admissibility of scientific evidence.⁴⁸ In taking this approach, the Court constructed a highly flexible framework within which trial courts should evaluate expertise.

Writing for the Court, Justice Breyer expanded *Daubert* to cover any kind of specialized expertise, even if it does not rest specifically on scientific knowledge. *Daubert* had involved purportedly scientific studies of the health effects of a particular drug; by contrast, *Kumho Tire* involved opinion evidence on the observation of tire wear. Justice Breyer noted that the language of Federal Rule 702 created “no relevant distinction”⁴⁹ among “scientific, technical, or other specialized knowledge.”⁵⁰ The very heart of expertise in evidentiary terms rests on specialization, not science: “[W]hether the specific expert testimony focuses upon specialized observations, the specialized translation of those observations into theory, a specialized theory itself, or the application of such a theory in a particular case, the expert’s testimony often will rest ‘upon an experience confessedly foreign in kind to [the jury’s] own.’”⁵¹

Rejecting slavish reliance on the *Daubert* factors,⁵² the Court asserted reliability and relevancy as the principal goals. For example, the testimony facing the court below consisted of opinions by an expert on

47. 526 U.S. 137 (1999).

48. *Id.* at 141–42. The trial court had excluded evidence by a plaintiff’s expert in “tire failure analysis,” offered in a products liability action to prove causation. *Id.* at 145. The trial court agreed that, even though the proposed expertise was more “technical” than “scientific,” it could still review it using the *Daubert* standard. *Carmichael v. Samyang Tires, Inc.* 923 F. Supp. 1514, 1521–22 (S.D. Ala. 1996). In so doing, the trial court at first used solely the four *Daubert* factors; but upon request by plaintiff for reconsideration, it agreed that these factors were illustrative, and applied a more flexible standard. *Kumho Tire*, 526 U.S. at 145–46. Nonetheless, it again excluded the proposed testimony, focusing especially on the expert’s methodology of analysing data obtained from a visual inspection. *Id.* at 146–47.

49. *Kumho Tire*, 526 U.S. at 147.

50. *Id.* (quoting FED. R. EVID. 702). Justice Breyer noted that each of the adjectives in the quoted phrase modify the word “knowledge,” and quoted *Daubert* for the proposition that the word “establishes a standard of evidentiary reliability.” *Id.* (quoting *Daubert*, 509 U.S. at 589–90).

51. *Id.* at 149 (alteration in original) (quoting Learned Hand, *Historical and Practical Considerations Regarding Expert Testimony*, 15 HARV. L. REV. 40, 54 (1901)).

52. *Kumho Tire*, 526 U.S. at 149–52. Specifically, the Court indicated that trial courts “may consider” the four factors articulated in *Daubert*, but that they were not required to use them. *Id.* at 149–50. Relying on arguments presented by the Solicitor General, the Court noted that “[t]he factors identified in *Daubert* may or may not be pertinent in assessing reliability, depending on the nature of the issue, the expert’s particular expertise, and the subject of his testimony.” *Id.* at 150 (alteration in original). It quoted *Daubert* specifically for the proposition that the four factors “do not constitute a ‘definitive checklist or test,’” *id.*, but that they retained vitality, even when evaluating non-scientific “experience-based testimony.” *Id.* at 151.

tire wear, using a hotly contested methodology of determining tire wear in a tort claim based on tire failure. While the Court reviewed the four considerations posed in *Daubert*, its opinion focused more on how the inferences drawn by the expert did or did not support the fact-finding required by the trial court.

The opinion described the goal of a reliability assessment as an effort “to make certain that an expert, whether basing testimony upon professional studies or personal experience, employs in the courtroom the same level of intellectual rigor that characterizes the practice of an expert in the relevant field.”⁵³ The Court encouraged trial judges to inquire how experts in the given field determine the reliability of opinions, but did not restrict a trial judge to these definitions. Instead, the Court left to trial judges the discretion to determine which standard of reliability to use for a particular form of expertise.⁵⁴

But reliability answers only part of the question. The Court’s opinion stresses that the fit of the opinion to the case should constitute the primary influence on the trial court’s discretion. Indeed, the fit of the opinion to the case constituted the primary influence on the trial court’s discretion. The *Kumho Tire* opinion stressed that trial judges should not evaluate admissibility generally and abstractly. Instead, they should do so specifically, with respect to the expert’s method of drawing “a conclusion regarding *the particular matter to which the expert testimony was directly relevant*.”⁵⁵ The Court took pains to assess the methodology of the particular expertise in the case, not in general, but rather as it applied to the particular facts of the case: “The relevant issue was whether the expert could reliably determine the cause of *this* tire’s [failure].”⁵⁶ It thus articulated a notion of fit that serves the demands of resolving particular disputed issues: “district courts must ‘scrutinize’ whether the ‘principles and methods’ employed by an expert ‘have been properly applied to the facts of the case.’”⁵⁷

53. *Id.* at 152.

54. In so doing, the Court built on the groundwork it had created in *General Electric Co. v. Joiner*, where it applied an abuse of discretion standard of review. 522 U.S. 136, 139 (1997). In the Court’s approach, this discretion extended not only to the decision on admissibility, but also to the selection of measures for reliability, and to the process through which decisions on admissibility would occur. *Kumho Tire*, 526 U.S. at 152. For example, trial judges have discretion to avoid unnecessary hearings on the admissibility of expert testimony, “where the reliability of an expert’s methods is properly taken for granted,” reserving more extensive review for testimony “in the less usual or more complex cases where cause for questioning the expert’s reliability arises.” *Id.*

55. *Id.* at 154.

56. *Id.*

57. *Id.* at 157 (citing *Advisory Committee’s Note on Proposed Fed. R. Evid. 702, Preliminary Draft of Proposed Amendments to the Federal Rules of Civil Procedure and Evidence: Request for Comment* 126 (1998)).

3. *Reliability and Fit After Kumho Tire*

Despite its reaffirmation of *Daubert's* core concern about pseudoscience, *Kumho Tire* created a remarkably open-ended analytical structure. The Court avoided limiting its reliability inquiry to the *Daubert* factors or other factors and emphasized that the fit inquiry relates to the factual contentions of particular cases. The Court presented this open-endedness as desirable, and expressed its trust of trial courts to work through the tough line-drawing problems in the crucible of litigation. Yet surprisingly basic questions about judicial assessments of expertise remain after *Kumho Tire*, questions that this Article will explore in the context of predictive testimony in civil commitments.

a. *Relative Reliability*

The opinion in *Daubert* speaks of reliability in terms of scientific method: “[I]n order to qualify as ‘scientific knowledge,’ an inference or assertion must be derived by the scientific method.”⁵⁸ Despite this, the Court made clear that “[p]roposed testimony must be supported by appropriate validation,” and described its standard as one of “*evidentiary* reliability.”⁵⁹ Later, the Court confirmed the more expansive reading: in some cases scientific validation may matter; “[i]n other cases, the relevant reliability concerns may focus upon personal knowledge or experience.”⁶⁰ This flexible notion of reliability raises four questions. First, what standards of reliability apply in a given context and what role should the professional community play? Second, will reliability permit less than perfect expertise? Finally, do some forms of expertise so lack reliability that *Daubert* would require per se exclusion, regardless of context?

As to sources, the Court clearly expressed comfort with the practice (well-established under *Frye*) of consultation with the relevant expert community. Both *Daubert* and *Kumho Tire* permit the trial judge to ask whether the particular method (or its application to the issues at hand) has received “general acceptance” in the expert community.⁶¹ Both *Daubert* and *Kumho Tire* advise the trial judge to look directly at the inferential process underlying the expert opinion, assessing, where appropriate, its verifiability, its error rate, and its history of publication and

58. *Daubert v. Merrell Dow Pharm., Inc.*, 509 U.S. 579, 590 (1993).

59. *Id.* (emphasis added).

60. *Kumho Tire Co. v. Carmichael*, 526 U.S. 137, 150–51 (1999) (rejecting the four *Daubert* factors as the sole tests of reliability, but noting their utility even in assessing “experience-based testimony”).

61. *Id.* at 151. “In certain cases, it will be appropriate for the trial judge to ask . . . whether such a method is generally accepted in the relevant . . . community.” *Id.*

critique. Other courts have added additional factors.⁶² Yet the Court clearly does not consider any of these factors conclusive. Instead, it leaves trial courts free to reject any particular form of inference, even when recommended by the relevant community, and to select between methodologies or inferences about which no common consensus exists.⁶³

This flexibility strongly suggests admissibility even if the expertise is less than perfect; it also suggests that, in the right case, highly unreliable expertise might also satisfy the new standard. The Court's opinions consistently speak of "appropriate validation"⁶⁴ or "reasonable measures of reliability,"⁶⁵ suggesting that reliability itself exists on a continuum of strength. Perfect reliability would clearly satisfy its test, while other kinds of expertise, even if well-structured and well-accepted within the field, may completely lack reliability. (Intriguingly, the Court names two disciplines, astrology and necromancy, which both involve prediction to some degree.⁶⁶) How can a trial court determine where on this continuum of strength a given expert opinion rests? The Court has answered this question as well, albeit in general terms: sufficiently strong to help the factfinder.⁶⁷

b. *Fit*

Both *Daubert* and *Kumho Tire* describe the second concern of the new standard as one of "relevance." This notion of relevance goes beyond logical relevance under Rule 401, and beyond the balancing of pro-

62. See, e.g., *Daubert v. Merrell Dow Pharm., Inc.*, 43 F.3d 1311, 1317–21 (9th Cir. 1995). On remand from the Supreme Court, additional factors included: whether the testifying expert had conducted supporting research solely for the litigation; whether the expert used some recognized methodology in reaching their conclusions. *Id.* The Ninth Circuit indicated that, if the sole question were the reliability of the expertise at issue, it might remand the case to allow plaintiffs to supplement affidavits under the new standard. *Id.* at 1320. However, the Court of Appeals refused to remand, given what it saw as an inappropriate "fit" between the inferences offered by the experts and the demands of proving causation in the case. *Id.* at 1320–22. See *infra* Part I.A.3.b. See also *E.I. du Pont de Nemours & Co. v. Robinson*, 923 S.W.2d 549, 557 (Tex. 1995) (additional factors include "the non-judicial uses which have been made of the theory or technique" and "the extent to which the technique relies upon the subjective interpretation of the expert").

63. See *Kumho Tire*, 526 U.S. at 153 (approving the trial court's determination that the expert testimony "fell outside the range where experts might reasonably differ"). See, e.g., *infra* Part I.B (discussing differing methodologies for predicting dangerousness).

64. *Daubert v. Merrell Dow Pharm., Inc.*, 509 U.S. 579, 590 (1993).

65. *Kumho Tire*, 526 U.S. at 153.

66. *Id.* at 151. ("Nor, on the other hand, does the presence of . . . general acceptance . . . help show that an expert's testimony is reliable where the discipline itself lacks reliability, as, for example, do theories grounded in any so-called generally accepted principles of astrology or necromancy.")

67. See FED. R. EVID. 702 ("If scientific, technical, or other specialized knowledge will assist the trier of fact to understand the evidence or to determine a fact in issue . . ."); *Daubert*, 509 U.S. at 591 ("Rule 702 further requires that the evidence or testimony 'assist the trier of fact to understand the evidence or to determine a fact in issue.' This condition goes primarily to relevance. 'Expert testimony which does not relate to any issue in the case is not relevant and, ergo, non-helpful.'").

bative value and prejudicial effect under Rule 403.⁶⁸ While the *Daubert* opinion approaches a logical version of relevance,⁶⁹ it typically uses more expansive language, describing the concern as one of “helpfulness” or “fit.” *Kumho Tire* takes the concern for helpfulness further, specifically assessing how the expertise in a case relates to the specific factual contentions it means to address.⁷⁰ The Court approved a standard that assesses “the reasonableness of using [a given expertise] to draw a conclusion regarding the particular matter to which the expert testimony was directly relevant.”⁷¹ Trial judges have “the discretionary authority . . . to determine reliability in light of the particular facts and circumstances of the particular case.”⁷²

At the extreme, this approach would prevent a generalized model of fit; not until trial, when the proponent of expertise had developed its probative purpose, might a trial judge assess the fit of the expertise to the case. Yet the Court’s opinions do suggest some general parameters: They articulate a dynamic relationship between reliability and fit. One version of this relationship would see reliability as the dominant factor, with helpfulness as solely a relevance check.⁷³ But it is more plausible to see fit as more than just relevance, operating in a dynamic relationship with reliability. A trial court might exclude demonstrably reliable expertise if it had an insufficient fit to the facts of the case. By the same token, a trial court might admit clearly unreliable evidence if it bore a sufficiently strong connection to the facts at issue.

Kumho Tire excluded questionable expertise because it did not help the fact-finder in that case, for the proffered purpose. In a sense, it affirmed the exclusion of weak (but not valueless) expertise because the particular inferences it expressed did not fit the necessary fact-finding. The Court’s cases do *not* address and thus do not answer whether a trial court might accept questionable expertise which has an especially strong fit to a particular case. The existence of this possibility suggests yet another range of questions: What factors (in addition to logical relevance)

68. See FED. R. EVID. 401 (defining relevant evidence) & 403 (requiring a balance of probative value and prejudicial impact).

69. See *Daubert*, 509 U.S. at 591–92 (“Expert testimony which does not relate to any issue in the case is not relevant and, ergo, non-helpful.”) (internal quotations omitted).

70. *Kumho Tire*, 526 U.S. at 153–57.

71. *Id.* at 153–54. “The trial court had to decide whether this particular expert had sufficient specialized knowledge to assist the jurors ‘in deciding the particular issues in the case.’” *Id.* at 156 (internal quotations omitted).

72. *Id.* at 158.

73. The Court’s reference to the phases of the moon in *Daubert* appears to approve this dynamic; here, the science of tracking moon phases appears relatively reliable, but the irrelevance of a particular moon phase to prove a given individual’s state of mind (“irrationality”) fails to satisfy our baseline notion of “relevance.” See *Daubert*, 509 U.S. at 591–92.

might a trial court use in assessing the strength of fit between expert opinions and the evidentiary demands of a given case?

This Article answers these questions by exploring predictive testimony in civil commitments. As the next subsection indicates, predictive testimony is not only "questionable," it has been roundly and thoroughly denigrated for forensic purposes, often by the very community of experts from which it derives.⁷⁴ Yet, predictive expertise bears a special relationship to civil commitment determinations, as explained further below.⁷⁵

B. DANGER

Mental health professionals use two primary approaches in predicting danger: clinical predictions and actuarial assessments.⁷⁶ The relevant literature reveals intense criticism of each of these techniques, criticism rooted both in science and in policy. As a result, some experts have persistently argued for the exclusion of psychiatric predictions from judicial proceedings, including civil commitments, while others have typically made guarded, and carefully limited, claims for admissibility.

I. *Methods of Inferring Future Violence*

Clinical predictions and actuarial assessments use different inferential processes. The first entails review of disparate information about a given individual, followed by the exercise of judgment about the risk of danger that individual might pose. The second entails the identification of specific characteristics of the individual that have been statistically correlated with a specified risk of violence. Clinical predictions make statements about the individual; actuarial assessments make statements about a group with which the individual shares characteristics, and from which one might draw inferences about the individual.⁷⁷

a. *Clinical Opinions*

Clinical opinion represents the earliest, and still the most widespread, form of predictive expertise. As we will see below,⁷⁸ psychiatrists and psychologists play a significant role in the initiation, continuation

74. See *infra* Part I.B (discussing the science of prediction, and the literature critical of that science).

75. See *infra* Part I.C (discussing the integration of predictive expertise into civil commitments).

76. William M. Grove & Paul E. Meehl, *Comparative Efficiency of Informal (Subjective, Impressionistic) and Formal (Mechanical, Algorithmic) Prediction Procedures: The Clinical Statistical Controversy*, 2 PSYCHOL., PUB. POL'Y & L. 293, 294 (1996) ("[There are] two ways of forecasting behavior. One, a formal method, uses an equation, a formula, a graph, or an actuarial table to arrive at a probability, or expected value, of some outcome; the other method relies on an informal, 'in the head,' impressionistic subjective conclusion, reached . . . by a human clinical judge.")

77. See Christopher Slobogin, *Dangerousness and Expertise*, 133 U. PA. L. REV. 97, 109-10 (1984).

78. See *infra* Part I.C (discussing the roles that psychiatrists play in initiating, advocating, reviewing and terminating civil commitment orders).

and termination of commitment orders; a majority of the reported decisions involve clinical predictions. Moreover, clinical opinions play a regular role in the routine practice of hospitals, including those designated to treat committed patients, as well as in other forensic contexts.⁷⁹

Clinical opinions share a common methodology. The clinical assessor interviews the patient directly, obtains the patient's medical history, and formulates a diagnosis of the patient's psychiatric condition. Data may include conversations with people familiar with the patient, written medical records of that patient, or conversations with other clinicians. The clinician may also compare the collected information to other similar cases with which the clinician is familiar, and may consult literature that provides further bases for comparison. This literature review may also include review of actuarial methods as they might apply to the patient.⁸⁰ The clinician then applies his or her judgment to this collected data in order to reach an opinion about whether the patient is likely to act dangerously.

Descriptions of clinical opinions stress their subjective and intuitive aspects: "[T]he clinical method relies on human judgment that is based on informal contemplation and, sometimes, discussion with others (e.g. case conferences)."⁸¹ During the late 1970s and into the 1980s, a second (or perhaps a third) wave of clinical approaches emerged, combining traditional clinical judgment with structured integration of actuarial conclusions into an overall opinion.⁸² Even when they incorporate actuarial information, however, clinical judgments apply intuitive and subjective assessments to patient-specific data.⁸³

79. See Kirk Heilbrun & Gretchen Witte, *The Macarthur Risk Assessment Study: Implications for Practice, Research and Policy*, 82 MARQ. L. REV. 733, 734-38 (1999) (listing the following uses of risk assessments, including clinical predictions: civil commitments; child custody and parental fitness; malpractice cases involving breach of a duty to warn; criminal sentencing; criminal commitment; correctional transfers; and sexual offender commitment).

80. See *infra* Part I.B.1.b. (discussing the uses of actuarial data in clinical assessments).

81. Grove & Meehl, *supra* note 76, at 293 (comparing different studies of the effectiveness of clinical as opposed to actuarial assessments, and concluding that "the mechanical method is almost invariably equal to or superior to the clinical method"). See also Caroline M. Mee & Harold V. Hall, *Risky Business: Assessing Dangerousness in Hawai'i*, 24 U. HAW. L. REV. 63, 90 (2001) ("A psychiatrist or psychologist reviewed medical records and conducted a personal interview, which formed the basis for a professional judgment as the subject's potential dangerousness. Unstructured clinical opinions were, by definition, subjective . . .").

82. Mee & Hall, *supra* note 81, at 90 (identifying "structured clinical opinions" as the second stage of violence prediction methods, and "empirically guided evaluations" as the third phase).

83. See R. Karl Hanson, *What Do We Know About Sex Offender Risk Assessment?*, 4 PSYCHOL., PUB. POL'Y & L. 50, 52-53 (1998) (distinguishing still further between "guided clinical" approaches and "adjusted actuarial" approaches).

In the *guided clinical approach*, expert evaluators consider a wide range of empirically validated risk factors and then form an overall opinion concerning the offender's recidivism

Clinical opinions have never received high marks for reliability. Early literature and studies almost completely discounted them, finding that clinicians did little better than chance.⁸⁴ A 1981 study by John Monahan, an early critic of predictive accuracy, summarized these studies, and critiqued their methodological shortcomings,⁸⁵ resulting in a “second generation” of research into the accuracy of clinical methods.⁸⁶ Over the past decade, these second generation research methods have led to a conclusion that clinical methods perform somewhat better than random, but are still deeply imperfect.⁸⁷ Assessments that incorporate actuarial data appear to have performed somewhat better than unguided and particularly unstructured assessments,⁸⁸ increasing the rate of reli-

risk. In the guided clinical approach, the method for translating the identified risk factors into recidivism rates is not explicitly determined. . . . The *adjusted actuarial approach* begins with an actuarial prediction, but expert evaluators can then adjust (or not) the actuarial prediction after considering potentially important factors that were not included in the actuarial measure.

Id. at 52–53 (internal quotations omitted). The key distinction remains the explicitness (and presumably the statistical soundness) with which the clinician translates the data into predictions of danger.

84. See JOHN MONAHAN, *THE CLINICAL PREDICTION OF VIOLENT BEHAVIOR* 44–56 (1981) (reviewing HENRY J. STEADMAN & JOSEPH J. COCOZZA, *CAREERS OF THE CRIMINALLY INSANE: EXCESSIVE SOCIAL CONTROL OF DEVIANCE* (1974); TERENCE P. THORNBERRY & JOSEPH E. JACOBY, *THE CRIMINALLY INSANE: A COMMUNITY FOLLOW-UP OF MENTALLY ILL OFFENDERS 200–201* (1979); Joseph J. Cocozza & Henry J. Steadman, *The Failure of Psychiatric Predictions of Dangerousness: Clear and Convincing Evidence*, 29 RUTGERS L. REV. 1084 (1976) [hereinafter Cocozza & Steadman, *The Failure of Psychiatric Predictions*]; Harry L. Kozol et al., *The Diagnosis and Treatment of Dangerousness*, 18 CRIME & DELINQ. 371 (1972); Henry J. Steadman, *A New Look at Recidivism Among Patuxent Inmates*, 5 BULL. AM. ACAD. PSYCHIATRY & L. 200, 200 (1977)).

85. *Id.*

86. See Randy K. Otto, *On the Ability of Mental Health Professionals to “Predict Dangerousness”*: *A Commentary on Interpretations of the “Dangerousness” Literature*, 18 LAW & PSYCHOL. REV. 43, 63 (1994). The changes in methodology implemented in this second generation include: greater use of clinicians’ predictions and assessments of risk (rather than on those of non-clinicians, including courts through commitment orders); expansion in the categories of behavior considered violent (such as the inclusion of verbal threats in addition to acts); expanding the range of sources from which violence after a prediction may be detected; and greater exploration of the role of situational and environmental factors.

87. See John Monahan, *Violence Risk Assessment: Scientific Validity and Evidentiary Admissibility*, 57 WASH. & LEE L. REV. 901, 903–05 (2000) (reviewing two studies in the 1990s which attribute greater reliability to clinical predictions than mere chance; Charles W. Lidz et al., *The Accuracy of Predictions of Violence to Others*, 269 JAMA 1007 (1993); Dale E. McNeil & Renee L. Binder, *Clinical Assessment of the Risk of Violence Among Psychiatric Inpatients*, 148 AM. J. PSYCHIATRY 1317 (1991)). See also Joseph T. McCann, *Risk Assessment and the Prediction of Violent Behavior*, 44. FED. LAW. 18, 18 (1997) (“Research over the past several years has shown that mental health professionals’ predictions of violence are better than chance . . .”).

88. See Hanson, *supra* note 83, at 60–67 (distinguishing between and assessing clinical, guided clinical, actuarial and adjusted actuarial methods of prediction).

Although most clinical assessments would be expected to have little predictive accuracy, there is some evidence that empirically guided risk assessments can provide estimates worthy of consideration. . . . Guided clinical assessments include a range of empirically vali-

ability from 1 in 3 to 1 in 2.⁸⁹ Overall, Monahan concluded that “the sober conclusion that clinicians are ‘modestly better than chance’ at predicting violence appears to be becoming the consensus view.”⁹⁰

The community of prediction researchers remains divided between experts who advocate use of clinical methods and those who stress the use of actuarial methods. Indeed, some of the sharpest critiques of clinical prediction come from those advocating actuarial approaches.⁹¹ However, clinical assessments remain dominant in the judicial process. Indeed, to the extent that clinical predictions reflect statements about a particular individual, rather than a group with which the individual shares some characteristics, they may receive warmer welcome at trial, which necessarily focuses on the individual.⁹²

b. Actuarial / Statistical

While statistical methods for assessing risk have existed since the early 1950s, only over the last ten to fifteen years have they become formally and widely used.⁹³ During this period, an extraordinarily diverse range of methodologies has emerged, both for the assessment of violent

dated risk indicators and then make recidivism estimates on the basis of the offenders’ rankings on these factors and the expected base rates for similar offenders.

Id. at 61–62 (internal citations omitted). *See also* Otto, *supra* note 86, at 63 (“[T]his body of research generally suggests that mental health professionals have some ability to assess risk and make predictions of violence (i.e. they do better than chance.)” (citing Randy K. Otto, *Prediction of Dangerous Behavior: A Review and Analysis of “Second Generation” Research*, 5 FORENSIC REP. 103 (1992))).

89. Otto, *supra* note 86, at 63 (“[W]hereas first generation research suggested that perhaps one out of three people predicted to engage in some kind of violent behavior will actually go on to do so, more recent studies suggest that one out of every two people predicted to be violent would go on to engage in some kind of legally relevant, violent behavior.”).

90. John Monahan, *Predictions of Violence: The Scientific Status of Research and Actuarial Predictions of Violence* § 9-2, in MODERN SCIENTIFIC EVIDENCE: THE LAW AND SCIENCE OF EXPERT TESTIMONY (David L. Faigman et al. eds., 2002) (assessing modern status of clinical predictions).

91. For example, Paul Meehl has proven a persistent and articulate critic of clinical approaches and advocate for stastically based methods. *See* Grove & Meehl, *supra* note 76, at 293; *see also* Mee & Hall, *supra* note 81, at 293.

92. *See* Eric S. Janus & Paul E. Meehl, *Assessing the Legal Standard for Predictions of Dangerousness in Sex Offender Commitment Proceedings*, 3 PSYCHOL., PUB. POL’Y & L. 33, 60–61 (1997) (arguing that clinical predictions also rest on assessments of groups, but acknowledging a “contemporary American value” for assessing the unique features of individuals).

Many courts, however, appear chary of actuarial methods of prediction. These methods treat the subject of commitment as a member of a group rather than as an individual, which appears to be contrary to contemporary American values. . . .

. . . [T]he use of actuarial methods may appear too mechanical and too general, undercutting the ideal that decisions about personal liberty should be individualized, taking into account the unique constellation of facts that make up this person.

Id.

93. Mee & Hall, *supra* note 81, at 91–92 (identifying the mid-1990s as the date for first development of pure actuarial measures, followed by two additional stages of refinement (combined actuarial, and classification tree approaches)).

behavior generally and for assessment of more specific populations, including the mentally ill and sex offenders.⁹⁴ The sources, methods, and purposes of the tools vary substantially, and this Article neither catalogues nor critiques them in detail. However, they share common assumptions and have similar effects. Literature (and case law) discusses them together under the rubric of “actuarial approaches.”

Actuarial researchers collect and analyze data on the characteristics and behaviors of a given group of individuals. The researchers identify features in that data that statistically correlate to acts of violence (generally or of a certain type) by members of that group. The assessments make statements about statistical correlations: for the given group, one

94. *Id.* at 97–113. The authors catalog seventeen different methods for violence risk assessment, including: a seriousness scoring system for quantifying the harmful consequences of past violence; a violence meta-analysis which compared predictors of recidivism for mentally disordered with those of non-disordered offenders; the Psychopathy Checklist-Revised (PCL-R), which “measures behaviours and personality traits that are considered fundamental to the clinical construct of psychopathy,” *id.* at 101 (quoting Ivan Zinger & Adelle E. Forth, *Psychopathology and Canadian Criminal Proceedings: The Potential for Human Rights Abuses*, 40 REVUE CANADIENNE DE CRIMINOLOGIE 237, 248 (1998)); the Violence Risk Appraisal Guide (VRAG), which produces a weighted score based on 11 different predictors of violence, including the PCL-R; the Sex Offender Risk Appraisal Guide (SORAG), which “predicts the risk that an adult male sex offender will commit another sex offense within a seven to ten year period,” *id.* at 104; a sexual recidivism meta-analysis, which measured a base-rate for reoffending for 28,972 sex offenders, as explored in eighty-seven articles; the Rapid Risk Assessment for Sexual Offense Recidivism (RRASOR), which “predicts the risk that an individual will commit a new sex offense within the next five to ten years,” by using a 4 factor weighted scoring system, *id.* at 106; the Minnesota Sex Offender Screening Test-Revised (MnSOST-R), which “predicts the risk that violent offender will commit another act of violence within the next six years,” using a sixteen factor weighted scoring system, *id.* at 107; static and dynamic risk assessment tools, which “yield low, medium and high risk predictors for sexual offenders,” *id.* at 108; a Dangerousness Prediction Decision Tree, which integrated both static violence risk factors and dynamic factors to predict risks of violence within the next three months; the Spousal Assault Risk Assessment Guide (SARA), which assesses risks of violence towards an intimate partner; the Sexual Violence Risk-20 (SVR-20), which act as a “set of professional guidelines for evaluating the risk of sexual violence,” *id.* at 110, but do not predict the risk of future sexual violence; the Historical Clinical Risk Management -20 (HCR-20), which act as a “set of professional guidelines to assess the risk of violence,” *id.* at 110; the California Actuarial Risk Assessment Table (CARAT) which present the base rates for reoffending of both child molesters and rapists; the Level of Service Inventory-Revised (LSI-R), which includes both static and dynamic factors to help assess the appropriateness of various professional interventions; and the Workplace Violence Risk Assessment Checklist (WVRAC), which serves as a screening tool to help those who manage personnel in a workplace make appropriate referrals for staff. The authors also take note of other methods about which they had insufficient data, including the Static-99, “which attempts to predict sexual recidivism by using a stepwise regression approach to classify offenders.” *Id.* at 113. *See also* Hanson, *supra* note 83, at 62–65 (describing and comparing the MnSOST, the VRAG, the RRASOR); Monahan, *supra* note 87, at 905–11 (assessing actuarial methods generally, including: the VRAG; a version of the PCL-R (Hare Psychopathy Checklist-Revised); the HCR-20; and the MacArthur Violence Risk Assessment Study, which uses an Iterative Classification Tree (ICT)). It is beyond the scope of this Article to offer an exhaustive review of all forms of statistically based violence risk assessment.

or more factors do or do not correlate to violence. In other words, the assessments identify “risk factors.”⁹⁵

To some extent, assessors can use these risk factors individually. The assessor notices which of these risk factors a particular individual exhibits, takes note of the probabilities associated with those various factors, considers them in light of overall base rates for violence for someone like the individual, and produces an individualized assessment of risk for that individual.⁹⁶ For example, in 1998, Hanson and Bussière examined a number of different risk factors for sexual recidivism, using a dataset that covered 28,972 sexual offenders.⁹⁷ As summarized later by Karl Hanson, ten factors had statistically significant correlations with later sexual reoffense.⁹⁸ An assessor who identifies any of these factors in a particular individual (and accepts the validity of the underlying study) can validly assert that the given risk factor in one individual correlates with violence in a group of other individuals with the same risk factor.

Risk factors also separate into categories. The differences influence the weighting of multiple risk factors. Static (or fixed) factors permit different inferences from dynamic (or changeable) risk factors.⁹⁹ Static factors include relatively long-term characteristics of a given individual, “such as a history of childhood maladjustment or prior offenses.”¹⁰⁰ Dynamic factors include characteristics that are either short-term or more susceptible to change. Literature recognizes a further subdivision of dynamic factors into stable and acute factors.¹⁰¹ Stable dynamic factors

95. Monahan, *supra* note 87, at 905 n.27 (citing Helena Chmura Kraemer et al., *Coming to Terms With the Terms of Risk*, 54 ARCHIVES GEN. PSYCHIATRY 337 (1997)). To say that a particular characteristic is a risk factor means that: “(1) the variable correlates with the outcome (in this case, violence), and (2) the variable precedes the outcome. To call a variable a ‘risk factor’ does not imply that its relationship to the outcome is in any sense ‘causal.’” *Id.*

96. See Hanson, *supra* note 83, at 56–58 (describing the use of individual risk factors in clinical assessments).

97. R. Karl Hanson & Monique T. Bussière, *Predicting Relapse: A Meta-Analysis of Sexual Offender Recidivism Studies*, 66 J. CONSULTING & CLINICAL PSYCHOL. 348, 351 (1998).

98. Hanson, *supra* note 83, at 57 tbl.1. The factors included: sexual preference for children; any deviant sexual preference; prior sexual offenses; failure to complete treatment; antisocial personality disorder; any prior offenses; age (youth); never married; any unrelated victims; any male child victims. *Id.*

99. *Id.* at 51–52 (describing the differences between different kinds of risk factors). The discussion which follows draws heavily on Hanson’s overview of different risk factors.

100. *Id.*

101. *Id.* (“Stable factors have the potential of changing but typically endure for long periods of time (e.g., deviant sexual preferences or alcoholism). Acute factors, in contrast, are rapidly changing states (e.g., sexual arousal or drunkenness) that immediately precede sexual crimes.”). Knowledge of static factors for a given individual permits inferences about that person’s long-term tendency towards particular kinds of behavior, but does not help to predict when violence might occur, or whether the individual has a reduced risk of offending (or reoffending). Dynamic factors “are those that predict

permit inferences about likely fluctuations in a given person's risk profile; by contrast, acute risk factors "help to determine the timing of reoffense."¹⁰²

The categorization of risk factors informs a central effort of modern actuarial research: the creation of overall risk assessment scales. An assessor using a "guided clinical" approach to prediction may attend to the statistical research underlying the identification of individual risk factors, but has no explicitly defined method for combining them; the combination rests on the clinician's judgment.¹⁰³ By contrast, modern risk assessment scales select a limited group of salient risk factors, assign values to these risk factors (often weighting some more heavily than others), and create a formula for producing an overall probability of risk (a "risk score") for the person who exhibits the relevant risk factors. The combination rests on tested pre-weighting of factors, not on clinical discretion.

For example, Hanson describes the creation of the Rapid Risk Assessment for Sex Offense Recidivism (RRASOR) as an "objective risk assessment procedure" designed to predict sexual recidivism.¹⁰⁴ Hanson reviewed and reanalysed data from eight different data sets, then reassessed the data using variables that had been identified as significant in a previous analysis. Using various statistical methods,¹⁰⁵ Hanson produced "[f]our variables that independently predicted recidivism": the number of prior sexual offenses; age at the time of release (over or under 25); earlier targeting of male victims; and targeting of unrelated victims by the offender.¹⁰⁶ He weighted each of these factors separately, then added them together.¹⁰⁷ The assessor uses the score so obtained to assess the

recidivism, have the potential of changing, and, when changed, are associated with corresponding increases or decreases in recidivism." *Id.*

The scientific literature appears to divide these factors in different ways; the discussion here seeks only to describe the utility of categorizing risk factors, especially for purposes of creating overall risk assessment scales, without offering any view on which categorization is most appropriate. *See, e.g.*, Mee & Hall, *supra* note 81, at 90–91 (contrasting historical factors, opportunity factors, and triggering stimuli).

102. Hanson, *supra* note 83, at 51.

103. *See supra* note 83 and accompanying text; Hanson, *supra* note 83, at 63–64.

104. *See* Hanson, *supra* note 83, at 63–64.

105. *Id.* ("[T]he average intercorrelations between the factors were calculated using standard meta-analytic procedures. The averaged correlation matrix was then analyzed using stepwise regression to identify the best subset of nonredundant predictors.").

106. *Id.* at 64.

107. *Id.* For example, "[p]rior sexual offenses were coded as follows: 0 = none; 1 = one prior conviction or one to two prior arrests; 2 = two or three prior convictions or three to five prior arrests; and 3 = four or more prior convictions or six or more prior arrests." *Id.* The assessor can adjust the score further for age, targeting of male victims, and offending against an unrelated victim. *Id.* "In practice, this scoring procedure gives the lowest scores to incest offenders with no prior records and the highest scores to young, boy-object child molesters with multiple prior sexual convictions." *Id.*

risk of reoffense for any member of the group who has the same configuration of factors. In effect, “actuarial approaches use explicit, formal procedures for translating ratings on a limited number of risk factors into an overall risk score.”¹⁰⁸

Hanson and others acknowledge a range of weaknesses and imperfections in such models, including:

- (1) The lack of comprehensiveness in the included factors—the underlying risk factors in the RRASOR, as in other devices, are “static and [thus] cannot be used to measure change.”¹⁰⁹ Other devices may include greater reliance on dynamic factors, thus improving the responsiveness of the device to changeable conditions; but to date, none appear to have an authoritative inventory of both static and dynamic factors.¹¹⁰
- (2) The imprecision in the categories of risk factors—for example, under the RRASOR, offending against a family member is weighted lower than offending against a stranger. Hanson asks, “what about an offender who selected a 14-year-old foster child who had been living in the home for 2 months as compared with a foster child who had been living in the home for 5 years?”¹¹¹ Even the RRASOR’s age cut-off of twenty-five, which appears to offer a bright line, will still leave “borderline cases.”¹¹²
- (3) Errors in identifying individual characteristics, such as frequent reliance on the individual’s self-reporting, the lack of relevant records, and contradictions in otherwise com-

108. *Id.*

109. *Id.* at 65.

110. *Id.* Hanson refers to the MnSOST as an example of a scale that includes at least a potential “dynamic risk factor.” *Id.* Moreover, he reports on his own efforts to develop a more comprehensive cataloging and assessment of dynamic factors. *Id.* at 58–60. Harold Hall identifies another such device, the Dangerousness Prediction Decision Tree, which seeks to incorporate a range of dynamic and static factors, together with both internal and external triggers. Mee & Hall, *supra* note 81, at 108–09 (describing the Decision Tree as another scoring instrument). The Decision Tree draws on a research conclusion that “[t]he best predictor of short-term dangerousness [was the presence of] multiple stimulus triggers (at least two), short-term in duration (less than one month), high in impact, superimposed on past violence.” *Id.* at 109 (second alteration in original) (internal quotations omitted). The purpose of the Decision Tree is thus “to predict the risk that an individual will commit another violent offense within the next three months.” *Id.*

This Article offers these as illustrations of possible methods of combining risk factors. A more subtle and nuanced cataloging of all methods is beyond the scope of this Article.

111. Hanson, *supra* note 83, at 61.

112. *Id.*

plete records, which can lower confidence in the overall score.¹¹³

How effectively do actuarial scales measure the probability of danger? To answer this question, one must distinguish between accuracy about groups and accuracy about individuals. As to groups, the “accuracy” of the scale means the number of correct predictions (both positive and negative) divided into the total number of predictions (including both correct and incorrect predictions).¹¹⁴ When a given scale has an accuracy of seventy-five percent, that percentage of the individuals scored as high risk did in fact commit the relevant violent act, as determined by actuarial studies.

To individualize these statements (i.e., to use the data to make predictions about individuals) requires knowledge of the “base rate” of violence (i.e., the frequency with which the behavior defined as violent occurs in the overall group). Even assuming a high accuracy for the overall group, the probability that a “positive” result is true varies with the base rate of the group on which a test is applied. If a condition is relatively rare in a group, even accurate tests identify lots of “false positives.”¹¹⁵

The group accuracy rate of most pure actuarial instruments is relatively high. Janus and Meehl posit accuracy rates hovering around seventy or seventy-five percent.¹¹⁶ Group rates thus compare favorably to those achieved by clinical predictions for individuals. As a result, many actuarial proponents argue the superiority of actuarial over clinical

113. *Id.* See also Donna Cropp Bechman, *Sex Offender Civil Commitments: Scientists or Psychics?*, 16 CRIM. JUST., Summer 2001, at 24, 26–30 (identifying the “five fatal flaws” of actuarial instruments, including underuse of dynamic factors, lack of peer review, uncertainty about the margin of error, lack of appropriate cross-validation, and disparities in results between different users of the instruments (“inter-rater reliability”)).

114. See Janus & Meehl, *supra* note 92, at 47–48.

115. *Id.* at 48. Janus and Meehl offer the following example of the interaction between test accuracy and base rates in computing individualized probabilities. Assume a scientist has developed a blood test that can detect cancer with 99% accuracy; 99 out of 100 times, the test reaches the correct result. Assume also that we test a screening group of 1,000. The example offers two different scenarios, involving different base rates:

Example 1: 0.1% base rate: that is, one person out of the 1000 will have cancer. With 99% accuracy, the test will almost certainly identify that one person correctly. But a 99% accuracy rate *also* means a 1% error rate; applied to the remaining 999 cancer-free patients means that the test will produce around ten (9.9) “false positives,” bringing the total identified with cancer to eleven people. “[T]he probability that any of the 11 ‘positives’ actually has cancer is only 1/11, or about 9%.” *Id.*

Example 2: 50% base rate: that is, 500 people out of the 1000 will have cancer. With 99% accuracy, the test will almost certainly identify 495 people correctly; with a 1% error rate, it will identify five people incorrectly, for a total of 500 people identified with cancer. “The probability that any one of those 500 actually has cancer is 495 / 500 or 99%.” *Id.*

116. *Id.* at 49.

methods.¹¹⁷ The most moderate advocates of actuarial approaches accept a combination of clinical and actuarial methods, with increasing reliance on actuarial methods as research techniques advance.¹¹⁸ More stringent advocates reject any form of clinical judgment, including an “adjusted actuarial” approach,¹¹⁹ or the cumulation and combination of disparate actuarial methods.¹²⁰ A few actuarial advocates claim that predictions based on actuarial methods can “reliably identify . . . offenders with an enduring propensity” to act violently, translating statistical data about risk factors for a group into a statement about an internal psychic state of a single member of that group.¹²¹

While general consensus does exist about the statistical validity of actuarial tools as statements about group risks, no comparable consensus appears to exist about the respective merits and appropriate combinations of clinical and actuarial methods to predict violence, especially for a single individual. Even strong actuarial advocates acknowledge that “the proactuarial position is apparently held by only a minority of practitioners.”¹²² As we will see, clinical approaches have also dominated the pre-

117. *Id.* at 48–49 (citing Grove & Meehl, *supra* note 76, at 298 (using meta-analytical techniques to compare the outcomes of studies of clinical and actuarial methods, and concluding, *inter alia*, “of the 136 studies, 64 favored the actuary . . . , 64 showed approximately equivalent accuracy, and 8 favored the clinician”). “Meehl’s thesis, confirmed in many contexts, is that clinical judgment is at best as good as, but often worse than, actuarial methods.” *Id.*

118. Hanson, *supra* note 83, at 67.

I have argued that there are three approaches that can provide credible risk assessments: the empirically guided clinical approach, the pure actuarial approach, and the adjusted actuarial approach. . . . [E]ach of these approaches can be expected to reliably identify a small subgroup of offenders with an enduring propensity to sexually reoffend.

Id. Hanson estimates the accuracy rate for these three methods as ranging between fifty percent and eighty percent, *id.* at 67–68, but separately cautions that “evaluators should, nevertheless, be exceedingly cautious about how much they adjust actuarial estimates.” *Id.* at 66.

119. *Id.* at 65. “The adjusted actuarial approach begins with actuarial predictions and then adjusts these assessments on the basis of other compelling evidence.” *Id.* Hanson asserts that “[e]ven the strongest advocates of actuarial prediction have always believed that adjustments to the statistical predictions can be justified in certain circumstances.” *Id.* (citing Grove & Meehl, *supra* note 76). Hanson also notes that “[t]here has been ongoing controversy concerning the likelihood that clinical adjustments will dilute rather than enhance actuarial predictions.” *Id.*

120. See Janus & Meehl, *supra* note 92, at 50.

It might be argued that the cumulative accuracy of seriatim review by a number of different evaluators will tend to increase the accuracy of the entire process. But this would be true only if the clinicians regularly knew when to overrule the actuarial studies or their colleagues’ predictions. . . . Further, a weak link in the chain could actually decrease accuracy.

Id.

121. Hanson, *supra* note 83, at 67.

122. Grove & Meehl, *supra* note 76, at 317. “How is it possible that thousands of MDs, PhDs, and MSWs . . . could be so wrong, as we allege?” *Id.* Grove and Meehl speculate about a range of reasons: fear of technological unemployment; self-concept; attachment to theory; misperception of the actuarial method as dehumanizing; general dislike of computers in competition with the human mind; and poor education. *Id.* at 317–18.

dictive expertise used in and reviewed by courts.¹²³ Yet whatever the opinion in the professional community, until recently, courts have taken a negative view of *any* prediction, clinical or otherwise. The next subsection offers a brief history of this skeptical view.

2. *Attitudes Towards Predictions*

Until the early 1970s, psychiatric testimony in civil commitments went largely unchallenged, both in courts and in the literature. However, for nearly fifteen years after 1970, a series of articles assessing psychiatric prediction appeared, much of it sharply critical.¹²⁴ John Monahan's influential study of both legal and psychiatric research literature concluded in 1981 that "the 'best' clinical research currently in existence indicates that psychiatrists and psychologists are accurate in no more than one out of three predictions of violent behavior."¹²⁵ In other words, "mental health professionals . . . are more likely to be wrong than right when they predict legally relevant behavior. When predicting violence, dangerousness, and suicide, they are far more likely to be wrong than right."¹²⁶ Indeed, according to some critics, not only were mental health professionals no better than lay people (or chance) at predicting violence, they might even be worse.¹²⁷

The intensity and frequency of the critique had at least two notable consequences. First, largely as a result of Monahan's seminal critique of the methodology of earlier psychological studies, the scientific commu-

123. See *infra* Part II.C. (reviewing predictive expertise in civil commitments).

124. Coccozza & Steadman, *The Failure of Psychiatric Predictions*, *supra* note 84, at 1098-1100; Alan M. Dershowitz, *The Law of Dangerousness: Some Fictions About Predictions*, 23 J. LEGAL EDUC. 24, 46-47 (1970); Bernard L. Diamond, *The Psychiatric Prediction of Dangerousness*, 123 U. PA. L. REV. 439, 440 (1974); see also George E. Dix, *The Death Penalty, "Dangerousness," Psychiatric Testimony, and Professional Ethics*, 5 AM. J. CRIM. L. 151, 212 (1977); George E. Dix, *Clinical Evaluation of the "Dangerousness" of "Normal" Criminal Defendants*, 66 VA. L. REV. 523, 546-47 (1980); George E. Dix, *Expert Prediction Testimony in Capital Sentencing: Evidentiary and Constitutional Considerations*, 19 AM. CRIM. L. REV. 1, 16 (1981); Ennis & Litwack, *supra* note 6, at 734-35; Kozol, *supra* note 84, at 392. (concluding that dangerousness can reliably be diagnosed); MONAHAN, *supra* note 84, at 44-56; Stephen J. Morse, *Crazy Behavior, Morals, and Science: An Analysis of Mental Health Law*, 51 S. CAL. L. REV. 527, 600 (1978); Bernard Rubin, *Prediction of Dangerousness in Mentally Ill Criminals*, 27 ARCHIVES GEN. PSYCHIATRY 397, 398 (1972); Slobogin, *supra* note 77, at 111; Henry J. Steadman, *Some Evidence on the Inadequacy of the Concept and Determination of Dangerousness in Law and Psychiatry*, 1 J. PSYCHIATRY & L. 409, 410 (1973); Henry J. Steadman & Joseph P. Morrissey, *The Statistical Prediction of Violent Behavior*, 5 L. & HUM. BEHAV. 263, 272-73 (1981); Andrew von Hirsch, *Prediction of Criminal Conduct and Preventive Confinement of Convicted Persons*, 21 BUFF. L. REV. 717, 736 (1971); AM. PSYCHIATRIC ASS'N, TASK FORCE REPORT 8, CLINICAL ASPECTS OF THE VIOLENT INDIVIDUAL 28 (1974) (concluding that "[n]either psychiatrists nor anyone else have demonstrated an ability to predict future violence or dangerousness"); Comment, *The Psychologist as Expert Witness: Science in the Courtroom?*, 38 MD. L. REV. 539, 577-81 (1979).

125. MONAHAN, *supra* note 84, at 47-49 (emphasis omitted).

126. Morse, *supra* note 124, at 600.

127. Ennis & Litwack, *supra* note 6, at 734-35.

nity entered into a second generation of assessments of the accuracy of clinical predictions.¹²⁸ That community renewed its efforts to discover more verifiable empirical bases on which to make those predictions, leading to a wave of statistical and actuarial studies and prediction methods in the 1990s.¹²⁹

Second, the attitude that predictive expertise lacked any reliability took root in reported legal opinions. In a series of constitutional cases, the United States Supreme Court assessed the due process implications both of civil commitments and of the use of predictive testimony.¹³⁰ In many of these cases, the Court strongly criticized the reliability of psychiatric predictions, noting a consensus opinion that psychiatric opinions predicting danger lack any reliability.¹³¹ To some extent, the Court developed its attitude from citations to the early literature.¹³² However, its views were also shaped by submissions from the American Psychiatric Association (APA), which filed repeated amicus briefs during this period, expressing its views on the lack of psychiatric expertise to make predictions about future behavior.¹³³ These opinions matched the harshness of the strongest critics of predictive expertise:

The professional literature uniformly establishes that such predictions are fundamentally of very low reliability, and that psychiatric testimony and expertise are irrelevant to such predictions. In view of these findings, psychiatric testimony on the issue of future criminal behavior only distorts the factfinding process. To the extent that there are important facts for a jury to consider on this issue, they can be fully presented by lay witnesses who do not testify with the mantle of professional expertise.¹³⁴

128. MONAHAN, *supra* note 84, at 44–56. *See supra* note 86 and accompanying text (identifying methodological reforms accomplished in the second generation research).

129. *See infra* Part I.B.1.b (discussing modern actuarial methods of assessing risk).

130. *See infra* Part I.C.1 (discussing the constitutional parameters of civil commitments and predictive testimony) and I.C.3 (discussing constitutional requirements for burdens of proof).

131. *See infra* Parts I.C.1 & I.C.3. *See, e.g.*, *Addington v. Texas*, 441 U.S. 418, 429 (1979) (“Given the lack of certainty and the fallibility of psychiatric diagnosis, there is a serious question as to whether a state could ever prove beyond a reasonable doubt that an individual is both mentally ill and likely to be dangerous.”).

132. *See, e.g.*, *Barefoot v. Estelle*, 463 U.S. 880, 920–24, (1983) (Blackmun, J., dissenting).

133. *Addington*, 441 U.S. at 431–33 (holding that due process required at least a “clear and convincing” evidence burden of proof in civil commitment actions); *Estelle v. Smith*, 451 U.S. 454, 467–69 (1981) (requiring the knowing waiver of right of self-incrimination prior to court-ordered psychiatric examination); *Barefoot*, 463 U.S. at 896 (holding that use of psychiatric predictions in death penalty determinations, however unreliable, did not violate due process). *See also* *Foucha v. Louisiana*, 504 U.S. 71, 78–80 (1992) (holding continued confinement of an insanity acquittee unconstitutional in the absence of a finding of mental illness).

134. Brief for Amicus Curiae American Psychiatric Association at 8; *Estelle v. Smith*, 451 U.S. 454 (1981).

The APA carefully limited its argument to long-term predictions, such as those required as part of the penalty phase in capital cases.¹³⁵ Yet its argument cut broadly, casting doubt on assessments in both civil and criminal cases and identifying weaknesses in methodology that applied to all predictive testimony.¹³⁶

As we will see, modern state cases dealing with the admissibility of expert predictions of danger have tended to assume the worst about predictions, yet still have readily used them as evidence.¹³⁷ The developing psychiatric literature has found a slightly improved reliability for clinical predictions and much stronger reliability for actuarial assessments in recent years.¹³⁸ The legal academy has also begun to embrace a more accepting view. Some analysts have continued the critique of psychiatric predictions as evidence, especially after *Daubert*.¹³⁹ John Monahan, however, has applied a slightly modified *Daubert* test to reach the conclusion

135. By contrast, it asserted that “in civil commitment cases . . . psychiatrists often can and do make reliable predictions about short-term prognoses, and such predictions often include potential violence.” *Id.* at n.5. The amicus brief clearly refers to clinical prediction methods with respect to short-term prediction: “the psychiatrist is able to evaluate the patient’s current mental condition and to discern likely behavioral patterns, including potential violent behavior in the near future, if the illness remains untreated.” *Id.* Interestingly, later cases and scholars rarely note this limitation on the A.P.A.’s opinion.

136. Its primary methodological complaint related to clinical reliance on base rates. Clinicians may produce reliable predictions where, for example, “past behavior clearly and repetitively evidence physical abuse of his or her children.” *Id.* at 13. However, this confidence results from knowledge of high base rates for violence in those situations. “Thus, to the extent that a psychiatrist’s prediction of ‘dangerousness’ is based solely on this knowledge of base rates of behavior, his prediction involves no more ‘expertise’—and certainly no more ‘psychiatric expertise’—than does that of the average nonexpert.” *Id.* at 14.

137. See *infra* Part II.B. (analyzing cases which address the admissibility of expert psychiatric predictions under evidentiary standards).

138. See *supra* Part I.B. (discussing current psychiatric literature).

139. Gary Gleb, Comment, *Washington’s Sexually Violent Predator Law: The Need to Bar Unreliable Psychiatric Predictions of Dangerousness from Civil Commitment Proceedings*, 39 UCLA L. REV. 213, 248–49 (1991) (recommending adoption of a modified *Frye* standard, and arguing that, under this standard, psychiatric predictions of long term danger “would be admissible in criminal proceedings but not in civil commitment proceedings”); Clayton Skaggs, *Kansas’ Sexual Predator Act and the Impact of Expert Predictions: Psyched Out by the Daubert Test*, 34 WASHBURN L. J. 320, 342 (1995) (assessing psychiatric predictions under the four identified *Daubert* factors, and concluding “a trial judge applying the *Daubert* Court’s ‘general observations’ would find the expert predictions inadmissible.”); Michael H. Gottesman, *From Barefoot to Daubert to Joiner: Triple Play or Double Error*, 40 ARIZ. L. REV. 753, 755 (1998) (arguing that *Daubert* will require the exclusion of psychiatric predictions and that “*Daubert* cannot be squared with *Barefoot*.”); Erica Beecher-Monas & Edgar Garcia-Rill, *The Law and The Brain: Judging Scientific Evidence of Intent*, 1 J. APP. PRAC. AND PROCESS 243, 273–74 (1999) (critiquing psychiatric predictions, and suggesting that “[i]t is doubtful that testimony about future dangerousness could withstand *Daubert* analysis”); Erica Beecher-Monas & Edgar Garcia-Rill, *Danger at the Edge of Chaos: Predicting Violent Behavior in a Post-Daubert World*, 24 CARDOZO L. REV. 1845, 1895–1901 (2003) (arguing that *Daubert* requires the exclusion of clinical predictions, but that actuarial instruments may improve judgments of juries in death penalty sentencing proceedings).

that predictive testimony should ordinarily be admissible.¹⁴⁰ Finally, many primary treatises and texts on scientific or expert evidence have concluded that *Daubert* will not result in the exclusion of expert predictions.¹⁴¹ As stated in *The Handbook of Forensic Psychology*, “it is highly unlikely that the *Daubert* decision will affect the admissibility of professional assessments of dangerousness in federal courts or in states that follow the *Daubert* decision.”¹⁴²

This Article agrees that *Daubert* should permit the admission of expert predictions, both clinical and actuarial, but not because of their scientific reliability. Instead, the Article uses the example of civil commitments to develop the notion that *Daubert*'s concern with evidentiary fit better explains the courts' receptivity to this form of predictive testimony.

C. CIVIL COMMITMENT

Civil commitment law requires a finding of “danger” as a prerequisite to commitment for an indefinite period. How well does psychiatric expertise fit this requirement? This section develops the proposition that trial judges make decisions about the fit of expertise in the context of standardized elements characteristic of any litigation. These elements include: the nature and purpose of the cause of action; the constitutional limits within which the litigation occurs; the allocation and measure of burdens of proof; the types of evidence offered and found sufficient to meet the burden; and the legal definition of the specific facts towards which the expert testifies. Some of these elements have special significance for civil commitment,¹⁴³ yet each will influence a trial judge in de-

140. Monahan, *supra* note 87, at 910–18. Monahan uses a six-factor test proposed by the Texas Supreme Court in adopting its version of *Daubert*. *Id.* at 911. See also E.I. du Pont de Nemours & Co., Inc. v. Robinson, 923 S.W.2d 549, 557 (Tex. 1995).

141. See 1 DAVID L. FAIGMAN ET AL., *Modern Scientific Evidence: The Law and Science of Expert Testimony* 420–21 (2002) (arguing that *Daubert* is unlikely to change judicial receptivity to predictive expertise, especially as predictive science develops greater demonstrated reliability); GARY B. MELTON ET AL., *PSYCHOLOGICAL EVALUATIONS FOR THE COURTS: A HANDBOOK FOR MENTAL HEALTH PROFESSIONALS & LAWYERS*, 292–93 (2d ed. 1997) (arguing that both empirical and clinical predictions entail non-obvious interpretation of data that require expertise, and thus justify expert testimony).

142. Thomas R. Litwack & Louise B. Schlesinger, *Dangerousness Risk Assessments: Research, Legal, and Clinical Considerations*, in *THE HANDBOOK OF FORENSIC PSYCHOLOGY* 192–93 (Allen K. Hess & Irving B. Weiner eds., 2d ed. 1999). The authors identify six different factors in support of this prediction: the Supreme Court's historical receptivity to predictive testimony; the likelihood that predictions will rest on validated risk factors; the argument that validation of expertise would require release of potentially dangerous individuals; the prevalence of professional predictions of danger in other parts of society; an asserted rise in the “general acceptance” of clinical predictions; and the likelihood that psychiatric witnesses will better communicate and assess risk factors than law people. *Id.*

143. In particular, the constitutional limits on civil commitment have no counterpart in the cases from which *Daubert*, *Joiner* and *Kumho Tire* have emerged.

cluding how particular expertise will affect the disposition of a given litigation. In assessing these elements, this section suggests a more refined model of *Daubert's* concerns for reliability and fit, one rooted less in science and more in the demands of dispute resolution.

I. Purpose and Constitutional Dimensions of Civil Commitment

Every state has enacted a form of civil commitment law.¹⁴⁴ In every jurisdiction, the law serves as the vehicle through which public or even private actors can obtain (or compel) treatment for those who need it. These laws mandate treatment of those with mental illness in the exercise of the state's *parens patriae* and police powers.¹⁴⁵ States' commitment laws usually distinguish among mental illness, developmental disability,¹⁴⁶ and other mental disorders justifying long-term treatment.¹⁴⁷ Civil commitment of any kind embodies and constrains conflict over three important policy goals: assuring public safety against predictable violence (including self-directed violence); alleviating the mental anguish of those suffering from an illness; and maximizing an individual's personal liberty.

Federal constitutional concerns place boundaries on state commitment laws. "Danger" itself has a constitutional dimension. In *O'Connor*

144. See *infra* Part I.C.5 (describing and citing to state commitment statutes' definitions of danger).

145. Mary Lynn Krongard, *A Population at Risk: Civil Commitment of Substance Abusers After Kansas v. Hendricks*, 90 CAL. L. REV. 111, 117-27 (2002). These same justifications underlie the handling of the mentally ill through the criminal justice system, a topic beyond the scope of this Article. Common examples of criminal proceedings in which predictions of danger might affect the outcome include: bail proceedings; sentencing proceedings; and criminal commitment proceedings.

146. See *Heller v. Doe*, 509 U.S. 312, 321-28 (1993) (assessing constitutionally required burden of proof in commitment for the developmentally disabled, and discussing differences). As an element in legal cases brought to compel treatment, "danger" plays a comparable role in the treatment of developmentally disabled patients. The author has found no cases which have assessed the admissibility of predictive testimony in these cases.

147. Recent statutory efforts to impose civil commitment as a remedy for "sexual predators" have received greater attention over the last decade. See *Kansas v. Hendricks*, 521 U.S. 346, 359 (1997) (upholding a Kansas "sexual predator" statute, and according state legislatures broad leeway in formulating the terms it uses to describe mental illnesses and mental disorders). But see Mary Lynn Krongard, *A Population at Risk: Civil Commitment of Substance Abusers After Kansas v. Hendricks*, 90 CAL. L. REV. 111, 131 (2002) ("Justice Thomas transformed the issue from a question of whether the new statute's target population fulfilled Fouca's constitutionally required 'mental illness' standard to a matter of synonym choices . . ."). States enact these special commitment statutes separately from general commitment statutes; while the mental condition required by law varies between special and general commitment statutes, both have been held to require a finding of "danger." *Kansas v. Hendricks*, 521 U.S. at 357. Many of the most recent evidentiary assessments of predictive testimony have occurred in cases dealing with these special commitment statutes.

The current round of controversy over "sexual predators" represents only a recent phase of a longer term use of commitment process to confine sexual offenders. See *Baxstrom v. Herold*, 383 U.S. 107, 110 (1966) (requiring state to provide an equivalent access to jury trials in general commitment and sexual offender statutes); *Humphrey v. Cady*, 405 U.S. 504, 509 (1972) (holding that a prisoner was entitled to prove that the denial of a jury trial under a sexual predator statute violated equal protection).

v. *Donaldson*,¹⁴⁸ the U.S. Supreme Court held that “a State cannot constitutionally confine without more a nondangerous individual who is capable of surviving safely in freedom”¹⁴⁹ The Court found that the U.S. Constitution required a finding of danger, even if the state statute required less.¹⁵⁰ It held that the state must prove dangerousness both at the time of initial commitment, and at the time of any later request for release.¹⁵¹ The state and the patient have multiple occasions to litigate the issue, in multiple procedural contexts, and thus have regular opportunity to assess the patient’s danger at different stages of treatment.¹⁵²

The *O’Connor* decision came down in the midst of the first wave of criticism of predictive expertise.¹⁵³ The strength and pervasiveness of the critique raised a constitutional question: If expert predictions have so little reliability, how can their use satisfy due process? The Court had already adverted to the unreliability of psychiatric predictions in commitment cases,¹⁵⁴ but it resolved the due process concern in a separate context, evaluating psychiatric predictions of danger during the death penalty phase of capital prosecutions.

148. *O’Connor v. Donaldson*, 422 U.S. 563 (1975).

149. *Id.* at 576. The Court’s ruling affirmed both trial and appeals court decisions in favor of a Florida man who had been held for 15 years in the absence of any showing of dangerousness. *Id.* at 568–70. It held that Donaldson’s confinement violated a constitutional guarantee of treatment sufficient to create “a reasonable opportunity to be cured or to improve his mental condition.” *Id.* at 572 (quoting *O’Connor v. Donaldson*, 493 F.2d 507, 520 (1974)). The Court rested its ruling solely on the lack of any evidence that Donaldson was dangerous. *O’Connor*, 422 U.S. at 573. Because no evidence of danger existed, the Court had no occasion to deal with testimony about danger, expert or otherwise.

150. *Id.* at 574 (“The fact that state law may have authorized confinement of the harmless mentally ill does not itself establish a constitutionally adequate purpose for the confinement.”).

151. *Id.* at 574–75 (“Nor is it enough that Donaldson’s original confinement was founded upon a constitutionally adequate basis, if in fact it was, because even if his involuntary confinement was initially permissible, it could not constitutionally continue after that basis no longer existed.”). See also *Kansas v. Hendricks*, 521 U.S. at 358 (reaffirming the danger requirement, while holding that states have latitude to define “mental illness” to include any “volitional impairment rendering them dangerous beyond their control”). But see *Parham v. J.R.*, 442 U.S. 584, 620 (1979) (holding constitutional a statute that permitted commitment of children by the voluntary act of the parent, without consideration of the child’s possible dangerousness).

152. Most statutory schemes contemplate multiple findings of dangerousness for a given patient: upon first entry into compulsory treatment; upon continuation of that treatment; upon transfer from hospitalization to community treatment, and upon return to the hospital; and upon the termination of treatment entirely. The nature of the danger determination may vary from issue to issue: the statutory definition of danger required for initial commitment, for example, may differ from the statutory definition required for subsequent release. Moreover, as we will see below, statutes define danger itself in different ways, and the range of different fact patterns which might justify a finding of danger are similarly diverse. See *infra* Part I.C.5 (discussing statutory definitions of danger).

153. See *supra* notes 124–27 and accompanying text.

154. See *infra* Part I.C.2 (discussing *Addington v. Texas*, 441 U.S. 418 (1979)).

In *Barefoot v. Estelle*,¹⁵⁵ the Court reviewed the adequacy of a sentencing hearing in which psychiatrists testified that the defendant “would probably commit further acts of violence and represented a continuing threat to society.”¹⁵⁶ Barefoot contended that psychiatric predictions lacked sufficient reliability to satisfy federal due process concerns and the APA concurred in an amicus brief.¹⁵⁷ The Supreme Court rejected the argument, however, over a sharply worded dissent by Justice Blackmun.¹⁵⁸ The majority reasoned that, because the Constitution did not prevent a state from requiring a jury to consider future dangerousness,¹⁵⁹

155. *Barefoot v. Estelle*, 463 U.S. 880 (1983).

156. *Id.* at 884. The jury agreed, and sentenced the defendant, Thomas Barefoot, to death. The defendant pursued appeals both through state and federal court systems, by direct appeal and by habeas corpus petition. *Id.* at 884–87. The appeal to the Supreme Court originated in a federal district court’s denial of Barefoot’s habeas corpus petition; the Fifth Circuit reviewed and affirmed that denial. Both lower court’s denied Barefoot’s claims with respect to the use of psychiatric expertise.

157. *Id.* at 899–902. See *supra* notes 133–36 (describing the position of the APA).

158. Justice Blackmun argued that, as a matter of due process, the defects in psychiatric testimony on danger run too deep to justify its use in a death penalty case. *Barefoot*, 463 U.S. at 916–38. He reviews the then current critique of predictive accuracy, noting the lack of professional support for the opinion, and the error rate of two in three. *Id.* at 920.

Interestingly, Justice Blackmun sounded themes that would reappear later in his majority opinion in *Daubert*: the need for trial court review of the impact of expert testimony; a mandate (albeit constitutional, not evidentiary) that expert evidence survive a scan for reliability; and his careful assessment of the scientific basis for the particular evidence in question. *Id.* at 929–36; see *Daubert v. Merrell Dow Pharm., Inc.*, 509 U.S. 579 (1993). He expressed strong reservations that the rules of evidence or the safeguards of cross-examination would adequately point out the deficiencies of such evidence. *Barefoot*, 463 U.S. at 931–32.

Were Justice Blackmun’s arguments in *Daubert* and *Barefoot* in some way related? On at least three of the four illustrative *Daubert* factors, Justice Blackmun’s views of psychiatric predictions in *Barefoot* could well have led to their exclusion under *Daubert*. Compare *Barefoot*, 463 U.S. at 932 (“intuitive clinical judgments not susceptible to cross-examination and rebuttal”), with *Daubert*, 509 U.S. at 593 (“whether [the scientific expertise] can be (and has been) tested”); compare *Barefoot*, 463 U.S. at 928 (“Psychiatric predictions of future dangerousness are not accurate; wrong two times out of three, their probative value . . . is virtually non-existent”), with *Daubert*, 509 U.S. at 594 (“the court should ordinarily consider the known or potential rate of error”); compare *Barefoot*, 463 U.S. at 920 (quoting the American Psychiatric Association’s Brief saying “[t]he unreliability of psychiatric predictions of long-term future dangerousness is by now an established fact within the profession”), with *Daubert*, 509 U.S. at 594 (“‘general acceptance’ can yet have a bearing on the inquiry. . . . ‘A known technique which has been able to attract only minimal support within the community’ . . . may properly be viewed with skepticism.”) (citation omitted). One can hear echoes of *Barefoot* in the Justice’s rhetoric in *Daubert*.

On the (evidentiary) merits, some indication exists in *Barefoot* that Justice Blackmun would have found actuarial assessments more reliable than clinical predictions. See *Barefoot*, 463 U.S. at 922 n.4 (contrasting the unreliability of clinical predictions with the greater reliability of statistical predictions: “Statistical prediction is clearly more reliable than clinical prediction.”).

See also *Flores v. Johnson*, 210 F.3d 456, 464 n.11 (5th Cir. 2000) (concurring opinion) (“It bears mentioning that Justice Blackmun, the author of *Daubert*, was also the author of the *Barefoot* dissent which harshly criticized the use of psychiatric evidence of future dangerousness.”).

159. *Barefoot*, 463 U.S. at 896–97. See *Jurek v. Texas*, 428 U.S. 262, 276 (1976) (approving Texas statute permitting consideration of the “likelihood of a defendant committing further crimes” as a fac-

it likewise did not limit the methods by which a state might choose to prove future danger, specifically including psychiatric testimony.¹⁶⁰ The majority worried that an opposite conclusion would call into question the use of psychiatrists in other contexts, especially civil commitment proceedings.¹⁶¹ It saw no reason to assume that psychiatrists were any less reliable than lay persons in finding a likelihood of danger.¹⁶²

At first blush, the *Barefoot* opinion appears not only to cut off due process objections to predictive testimony, but also to override objections to its reliability, including in civil commitment proceedings.¹⁶³ Yet a careful reading indicates that *Barefoot* specifically splits the evidentiary from the constitutional questions, and reserves the former for resolution at trial. In this view, the rules of evidence, and the processes of cross-examination, would adequately expose deficiencies in psychiatric testimony on danger: "Psychiatric testimony predicting dangerousness may be countered not only as erroneous in a particular case but also as generally so unreliable that it should be ignored."¹⁶⁴ *Barefoot* says that if such opinions pass evidentiary muster under sub-constitutional evidentiary standards, they do not violate any constitutional notions of sufficiency to support the death penalty.¹⁶⁵ The opinion leaves open the question as to what such evidentiary standards might require.

tor in imposing the death penalty). See also *Estelle v. Smith*, 451 U.S. 454, 463-66 (1981) (reversing a trial based on testimony by a psychiatrist that included statements made by defendant without advising the defendant of *Miranda* rights and without the benefit of his attorney).

160. *Barefoot*, 463 U.S. at 896-97.

161. *Id.* at 898-99 (citing *O'Connor v. Donaldson* and *Addington v. Texas*). See *supra* notes 149-51 (discussing *O'Connor*) and *infra* notes 169-73 (discussing *Addington*).

162. *Id.* at 899-903. "[I]t makes little sense, if any, to submit that psychiatrists, out of the entire universe of persons who might have an opinion on the issue, would know so little about the subject that they should not be permitted to testify." *Id.* at 897. The Court noted that many psychiatrists disagreed with the position of the A.P.A., and it reviewed with some care the conflict in testimony both at the state trial and in the District Court's habeas proceedings. The Court stressed the dissent within the psychiatric community about the validity of predicting danger, and noted that the only witness to suggest directly that "no reliable psychiatric predictions of dangerousness could ever be made" had conceded that evidence existed of some degree of accuracy in the predictions made by psychiatrists. *Id.* at 899-901 & n.7, (citing JOHN MONAHAN, *THE CLINICAL PREDICTION OF VIOLENT BEHAVIOR* 47-49 (1981)); see *infra* notes 125-28 and accompanying text (discussing further analysis by Professor Monahan).

163. The majority raised the prospect of damage to civil commitment process as a justification for its decision. *Id.* at 898 (citing *O'Connor* and *Addington*: "Acceptance of petitioner's position . . . would immediately call into question those other contexts in which predictions of future behavior are constantly made"). Even Justice Blackmun in dissent noted differences which made such predictive testimony more palatable in civil commitment cases. *Id.* at 936 n.14 (Blackmun, J., dissenting). Justice Blackmun distinguished civil commitments as cases which involve short-term predictions with short-term consequences, and argued both that the predictive accuracy was higher and the constitutional stakes lower in commitments. *Id.*

164. *Id.* at 898.

165. *Id.* at 901 ("[T]he submission is that this category of testimony should be excised entirely from all trials. We are unconvinced . . . that the adversary process cannot be trusted to sort out the reliable

The Court's due process analysis in *Barefoot* adds two elements to our discussion of the evidentiary admissibility of predictive expertise. First, as noted, the majority opinion at once agreed with the widespread skepticism of predictive testimony, yet found that such testimony had some reliability. As the Court itself stated, "[w]e are not persuaded that such testimony is almost entirely unreliable." Second, even with such a minimalist assessment, the majority felt that predictive testimony had sufficient reliability for use in the courtroom, if subjected to fair adversarial testing. Indeed, both majority and dissent saw a special need for such evidence in the context of civil commitment. This balancing of helpfulness and reliability for constitutional purposes anticipates the similar approach of *Daubert*. It also set a pattern that would recur in both pre- and post-*Daubert* evidentiary rulings.¹⁶⁶ In effect, *Barefoot* established an analytical template for admitting unreliable testimony in light of its close fit with the issues raised by a particular case.

2. *Burden of proof*

Part of the difficulty of predictive testimony lies less with the expertise than with the task itself: Prediction inevitably brings with it a risk of error. We can hope to reconstruct past events, but future events have not yet happened. Making "findings" about the future thus carries a greater risk of error. In civil commitments, error has severe consequences either way: loss of liberty for the patient if commitment wrongly succeeds; and violent behavior if commitment wrongly fails. States regulate this zone of risk in civil commitment by allocating burdens of proof, and the legal allocations take account of the unreliability of prediction and of predictive testimony.

In *Addington v. Texas*,¹⁶⁷ the Court held that the proponent of civil commitment must prove all elements of a civil commitment action, including future dangerousness, using the standard of "clear and convincing evidence."¹⁶⁸ The Court found in this standard a sensible middle

from the unreliable evidence and opinion about future dangerousness . . ."). See also *In re Brown*, 493 A.2d 447, 450 (N.H. 1985) (holding that use of psychiatric predictions of danger in civil commitment cases does not violate due process, given the presence of procedural "safeguards" against error, citing *In re Scott L.*, 469 A.2d 1336, 1337 (N.H. 1983) (dicta that a stringent burden of proof and the requirement of recent overt acts constitute adequate safeguards)); *Edmonds v. Commonwealth*, 329 S.E.2d 807, 813 (Va. 1985) (approving use of expert psychiatric predictions of danger in death penalty cases); *Saunders v. Commonwealth*, 406 S.E.2d 39, 43 (Va. 1991) (while constitutionally admissible, juries not bound to accept expert psychiatric predictions of danger).

166. See *infra* Part II.B (assessing cases affirming evidentiary admission of predictive opinion).

167. *Addington v. Texas*, 441 U.S. 418, 433 (1979).

168. *Id.* at 433. See *supra* note 92.

ground,¹⁶⁹ rooted in the special characteristics of civil commitments. Given the deprivation of liberty and profound stigma imposed on patients, something more than a preponderance should be required.¹⁷⁰ But the fallibility of predictions, and especially of psychiatric testimony, rendered the “reasonable doubt” standard unworkable:

[W]hether the individual is mentally ill and dangerous . . . turns on the *meaning* of the facts which must be interpreted by expert psychiatrists and psychologists. Given the lack of certainty and the fallibility of psychiatric diagnosis, there is a serious question as to whether a state could ever prove beyond a reasonable doubt that an individual is both mentally ill and likely to be dangerous.¹⁷¹

The Court stressed the fundamental uncertainties of psychiatric testimony:

The subtleties and nuances of psychiatric diagnosis render certainties virtually beyond reach in most situations. . . . Psychiatric diagnosis . . . is to a large extent based on medical “impressions” drawn from subjective analysis and filtered through the experience of the diagnostician. This process often makes it very difficult for the expert physician to offer definite conclusions about any particular patient. . . . If a trained psychiatrist has difficulty with the categorical “beyond a reasonable doubt” standard, the untrained lay juror—or indeed even a trained judge—who is required to rely upon expert opinion could be forced by the criminal law standard of proof to reject commitment for many patients desperately in need of institutionalized psychiatric care.¹⁷²

The Court established the clear and convincing standard as the minimum required by the U.S. Constitution; states remained free to adopt higher burdens.¹⁷³

The Court revisited these questions in *Heller v. Doe*, when it decided that equal protection did not require the same burden of proof for different kinds of commitments.¹⁷⁴ The statute in question required a “clear

169. *Addington*, 441 U.S. at 423–24 (comparing the “three standards or levels of proof for different types of cases,” including a preponderance of the evidence, clear and convincing evidence and beyond a reasonable doubt). The Court also expressed skepticism that burdens of proof really mattered. *Id.* at 424–25 (“Candor suggests that, to a degree, efforts to analyze what lay jurors understand concerning the differences among these three tests . . . may well be largely an academic exercise . . .”).

170. *Id.* at 426–27 (“Loss of liberty calls for a showing that the individual suffers from something more serious than is demonstrated by idiosyncratic behavior. Increasing the burden of proof is one way to impress the factfinder with the importance of the decision . . .”).

171. *Id.* at 429.

172. *Id.* at 430. Note that the Court’s description of expertise entails clinical rather than actuarial prediction.

173. *Id.* at 431.

174. *Heller v. Doe*, 509 U.S. 312, 326–28 (1993). Plaintiffs challenged a Kentucky statute which set a “clear and convincing standard” for commitment of mentally retarded individuals, and a “reasonable

and convincing” standard for committing the mentally retarded and a “reasonable doubt” standard for committing the mentally ill. The Court found a rational basis for the disparity.¹⁷⁵ It noted three distinctions between mental illness and mental retardation that satisfied this minimal standard; one of these distinctions was the relative difficulty of proving dangerousness.¹⁷⁶ In discussing predictions of danger, Justice Kennedy noted that “mental retardation is a permanent, relatively static condition, so a determination of dangerousness may be made with some accuracy based on previous behavior.”¹⁷⁷ The Court contrasted this with mental illness:

Manifestations of mental illness may be sudden, and past behavior may not be an adequate predictor of future actions. Prediction of future behavior is complicated as well by the difficulties inherent in diagnosis of mental illness. It is thus no surprise that many psychiatric predictions of future violent behavior by the mentally ill are inaccurate.¹⁷⁸

Given the greater uncertainty of predicting the behavior of the mentally ill, Kentucky could rationally allocate a higher burden of proof in this case: “a higher burden of proof for [commitment of the mentally ill] . . . tends to equalize the risks of an erroneous determination”¹⁷⁹

Like Kentucky, some states have imposed the more stringent, “beyond a reasonable doubt” standard,¹⁸⁰ while others use the less stringent

doubt” standard for commitment of mentally ill individuals. *Id.* at 315–17 (discussing Ky. Rev. Stat. § 202B.160(2) (mental retardation) and Ky. Rev. Stat. § 202A.076 (2) (mental illness)).

175. *Id.* at 319–21 (seeking “any reasonably conceivable state of facts that could provide a rational basis” for the difference in standards).

176. *Id.* at 321–28. The other two were the relative ease of diagnosis; and relative intrusiveness of treatment. As to ease of diagnosis, the Court rested its view largely on the proposition that, unlike mental illness, “mental retardation is a developmental disability that becomes apparent before adulthood”; the disparity in ease of diagnosis justified the state’s use of a stricter standard where the risk of error was higher. *Id.* at 321–23. As to intrusiveness of treatment, the Court noted that mentally ill patients often undergo “intrusive inquiries into the patient’s innermost thoughts . . . and use of psychotropic drugs”; the disparity in ease of diagnosis justified a stricter standard where the respondent suffered more severe consequences. *Id.* at 324–25. The Court also noted both the historical and the modern day division of statutes regulating the commitment of the two different categories of condition. *Id.* at 326–28.

177. *Id.* at 323 (citation omitted).

178. *Id.* at 323–24 (citation omitted).

179. *Id.* at 322. (discussing the relationship of more difficult diagnosis of mental illness to Kentucky’s allocation of the burden of proof). The majority opinion on this point drew only 5 votes.

180. *People v. Burnick*, 535 P.2d 352, 369 (Cal. 1975) (requiring “beyond a reasonable doubt” for civil commitments of sexual offenders); *Superintendent of Worcester State Hosp. v. Hagberg*, 372 N.E.2d 242, 245–46 (Mass. 1978) (requiring “beyond a reasonable doubt standard” for mental health commitments); *but cf.* *Guardianship of Doe*, 583 N.E.2d 1263, 1271 (Mass. 1992) (“in cases involving important personal rights, ‘we have refused to apply either the “beyond a reasonable doubt” standard or the “clear and convincing” standard.’ Rather, we have determined that ‘fact-finding is enhanced by requiring that it be done in writing and in meticulous detail.’”); *In re D.D.*, 920 P.2d 973, 975 (Mont. 1996) (noting that MONT. CODE ANN. § 53-21-126(2) requires proof of “any physical fact or evidence”

standard required by *Addington*.¹⁸¹ The rationales in state cases which discuss burdens typically focus on the need for a stringent standard in light of the difficulty of predicting danger and the uncertainty of psychiatric testimony on the point. In *People v. Burnick*,¹⁸² for example, the California Supreme Court reviewed the "first generation" critique of predicting danger, and articulated a severely critical assessment' of the science of prediction:

In the light of recent studies it is no longer heresy to question the reliability of psychiatric predictions. Psychiatrists themselves would be the first to admit that however desirable an infallible crystal ball might be, it is not among the tools of their profession. . . . "The evidence, as well as the consensus of opinion by responsible scientific authorities, is now unequivocal." . . . In the words of spokesmen for the psychiatric profession itself, "Unfortunately, this is the state of the art. Neither psychiatrists nor anyone else have reliably demonstrated an ability to predict future violence or 'dangerousness.' Neither has any special psychiatric 'expertise' in this area been established." (Task Force Report, Clinical Aspects of the Violent Individual (American Psychiatric Assn., 1974) p. 28.)¹⁸³

The court used this harsh assessment to justify a "reasonable doubt" standard, while rejecting calls for an absolute constitutional exclusion.¹⁸⁴

These opinions deal with the most unreliable form of predictive expertise, clinical opinion. They accept and stress the unreliability not just of prediction, but also of predictive testimony. In *Addington* and *Heller*, the Court showed its willingness to use the burden of proof in civil commitments to accommodate the risk of error inherent in both predictive fact-finding and expert testimony on prediction. This accommodation reinforces the fit between predictive opinion and the civil commitment process, and helps to identify the burden of proof as another element of the context within which trial courts can make determinations of fit under *Daubert*.

beyond a reasonable doubt, and proof of "all other matters" by clear and convincing evidence). See also *Proctor v. Butler*, 380 A.2d 673, 677 (N.H. 1977) (establishing "beyond a reasonable doubt" standard for civil commitments), *overruled by In re Sanborn*, 545 A.2d 726, 736 (N.H. 1988) (establishing "clear and convincing evidence" standard for civil commitments).

181. See, e.g., *State v. Ward*, 369 N.W.2d 293, 295 (Minn. 1985) (requiring "clear and convincing evidence" in commitment of sexual offenders); *People v. Taylor*, 618 P.2d 1127, 1135-36 (Colo. 1980) (requiring "clear and convincing evidence" in civil commitments).

182. 535 P.2d 352 (Cal. 1975).

183. *Id.* at 365-66. See also Part II.B.3 *infra* (discussing California's exemption of predictive testimony from the *Frye* standard of "general acceptance.")

184. *Id.* ("Nor do we go so far as to join in the conclusion of certain well-known writers that in civil commitment proceedings no psychiatrists should be permitted to give their opinions as to future dangerousness. . . .").

3. *Other Evidence*

"A brick is not a wall."¹⁸⁵ Experts typically do not testify in an evidentiary vacuum. In commitment proceedings, the proponent of predictive testimony not only has other statutory elements either to establish (or rebut),¹⁸⁶ but also has other means to prove future danger. In assessing the "fit" of expert testimony to the demands of a particular kind of case, then, it makes sense to assess the evidentiary patterns within which the expert's opinion will be offered. Trial court decisions about the fit of expertise to a case will most certainly consider the body of other evidence. In other words, one should know what other bricks are needed for the wall to stand. If one means for a flawed brick to bear weight, one must consider not only its flaws but also its fit with other bricks, and its impact on their arrangement.

While avoiding a comprehensive review of all evidence on danger, this section does attempt to categorize that evidence and to assess how predictive expertise fits as part of a greater whole. First, the section considers the prevalence of expert testimony. Then, it discusses its sufficiency to sustain a finding of danger, in the absence of other evidence. Finally, we assess rules requiring additional proof of danger, including proof of "overt acts."

a. *The Prevalence of Expert Testimony*

Many jurisdictions trump the evidentiary question about expert psychiatric predictions by requiring experts to testify in commitment cases.¹⁸⁷ In some states, statutes require the testimony,¹⁸⁸ while court-made rules

185. 1 MCCORMICK ON EVIDENCE § 185, at 542-43 (E. Cleary ed., 3d ed. 1984).

186. These statutory elements include the constitutionally required finding of mental illness, *see* O'Connor v. Donaldson, 422 U.S. 563, 576 (1975); MICHAEL L. PERLIN, LAW AND MENTAL DISABILITY 10-15 (1994) (discussing the mean of the statutory requirement of mental illness—"there can be no doubt that some finding of 'mental illness' is a prerequisite to an application for involuntary civil commitment . . ."), as well as a finding that commitment is the "least restrictive alternative" form of treatment. PERLIN, *supra*, at 114-20 (discussing the "least restrictive alternative" doctrine—"this principle has been incorporated in many civil commitment statutes, and is routinely invoked at individual commitment hearings on a daily basis . . . [T]he importance of the concept to the fabric of the commitment process cannot be overstated.").

187. 1 DAVID L. FAIGMAN ET AL., MODERN SCIENTIFIC EVIDENCE 411-15 (2002): "Psychiatric predictions are thus admissible on the basis of substantive law, and neither lack of general acceptance nor lack of scientific validity effects this conclusion." *Id.* at 414.

188. *See, e.g.*, 405 ILL. COMP. STAT. ANN. § 5/3-807 (West 1997) (requiring at least one mental health professional). Other states strongly imply the testimony of a physician. *See, e.g.*, R.I. GEN. LAWS § 40.1-5-2 (1997). ("In determining whether there exists a likelihood of serious harm the physician and the court may consider previous acts, diagnosis, words or thoughts of the patient."). As does Rhode Island, some statutes limit the requirement to particular topics. *See, e.g.*, ME. REV. STAT. ANN. tit. 34-B, § 3864 (West 1964 & Supp. 2002) (requirement of expert testimony on the proposed treatment plan for a potential committee). *See also* Matter of Oseing, 296 N.W.2d 797, 799 (Iowa 1980) (noting an Iowa

require it in others.¹⁸⁹ Requirements such as these obviate the decision on admissibility; evidence that must be offered necessarily must be heard.¹⁹⁰ Absent such provisions, however, how prevalent is predictive expertise in commitment cases?

The answer is "very prevalent," not just as a matter of effective proof, but also as a matter of statutory structure.¹⁹¹ Virtually all statutes have provisions requiring an examination by a physician or psychiatrist, which typically includes an assessment of the person's dangerousness.¹⁹² Typically, this assessment results in the petition for commitment. Moreover, some states also require written assessments, or even the presence

statute requiring recommendations by the chief medical officer of the state hospital concerning alternative placements).

189. Many cases note and apply a relevant statutory requirement. *See, e.g., J.W.K. v. State*, 370 N.W.2d 294 (Wis.App. 1985) (unpublished opinion) (applying Wis. STATE ANN. 51.20(4) (West 1997)); *State for Interest of P.W.*, 801 S.W.2d 1, 4-5 (Tex.App. 1990) (noting requirement in TEX. REV. CIV. STAT. ANN. art. 5547, §50(c) (Vernon Supp. 1990)); *People v. Henderson*, 162 Cal. Rptr. 886, 892 (Cal. Ct. App. 1980) (noting requirement of expert testimony on danger in sexually violent predator statute). Other state courts express the requirement without referring to a statutory basis. *See, e.g., In re Cochran*, 487 N.E.2d 389, 391 (Ill. App. Ct. 1985) ("[I]n order to meet its burden of proof, the State must proffer explicit medical testimony asserting that as a direct result of mental illness, the respondent can reasonably be expected to harm himself or another."). Still other states do not impose a requirement of expert predictive testimony, but argue that it is usually necessary. *See cases cited infra* note 194.

190. Again, one should distinguish between rules of sufficiency and those of admissibility. Where the jurisdiction requires expert testimony before a valid commitment order can issue, a rule of sufficiency exists, and only questions over the wisdom of such a provision as a matter of substantive commitment law remain. This Article's analysis governs those states in which no requirement of predictive expertise exists, nor any statutory rule of admissibility.

191. No empirical study exists on the frequency with which experts appear at commitment hearings, nor have I attempted one. Virtually all of the appellate decisions dealing with commitment challenges contain descriptions of psychiatric testimony. To some degree, this reflects structural choices by the relevant jurisdiction to encourage and even require the involvement of psychiatric experts in the commitment process. To some degree, it reflects the high likelihood of admissibility, coupled with the persuasive advantages of having testimony from a skilled witness on critical issues of proof.

192. *See, e.g., MO. REV. STAT. § 632.300(2)* (2000).

If, as the result of personal observation or investigation, the mental health coordinator has reasonable cause to believe that such person is mentally disordered and, as a result, presents a likelihood of serious harm to himself or others, the mental health coordinator may file an application with the court

Id. See also DEL. CODE ANN. tit. 16 § 5003 (1995).

No person shall be involuntarily admitted to the hospital as a patient except pursuant to the written certification of a psychiatrist that based upon the psychiatrist's examination of such person, such person suffers from a disease or condition which requires him to be observed and treated at a mental hospital for his own welfare and which either renders such person unable to make responsible decisions with respect to his hospitalization, or poses a present threat, based upon manifest indications, that such person is likely to commit or suffer serious harm to himself or others or to property, if not given immediate hospital care and treatment.

Id.

of the initial psychiatrist, at trial.¹⁹³ Mental health professionals thus play a critical role in initiating the commitment process.

Courts have also expressed the view that expert predictions are pragmatically necessary in commitment actions: "In civil commitment cases, where the trier of fact is required by statute to determine whether a person is dangerous or likely to be dangerous, expert prediction may be the only evidence available."¹⁹⁴ This necessity rationale acknowledges the limitations of expert testimony on danger, but finds it either better than other available proof, or the only available proof of danger.

Yet strictly speaking, expert testimony which predicts danger is not required. Fact-finders can receive (through independent proof) the same data available to the expert.¹⁹⁵ While experts may add an "aura of certainty," their demonstrated (in)accuracy makes their opinions something less than conclusive on the issue. Moreover, as Justice Blackmun noted in *Barefoot*, if expert opinion is little better than lay opinion, "statistical prediction is clearly more reliable than clinical prediction, . . . and prediction based on statistics alone may be done by anyone."¹⁹⁶ It seems best to characterize expert testimony as a preferred, but not essential, form of proof.

For present purposes, it suffices to note that information from psychiatric sources permeates civil commitment actions, including initial certification, medical histories maintained by psychiatric personnel,¹⁹⁷ and

193. Note that requiring a mental health expert to appear does not require testimony on danger. As noted elsewhere, the elements of civil commitment cases include multiple elements, including many about which a mental health professional will have specialized knowledge. These facts include the diagnosis of mental illness, the appropriateness of different treatments, and the availability of treatments both in restricted and unrestricted settings. It would be plausible to have a commitment proceeding where mental health experts testify only on these issues, leaving the prediction of danger as a task for judicial resolution based on other proof.

194. *People v. Ward*, 83 Cal. Rptr. 2d 828, 832 (Cal. Ct. App. 1999) (admitting expert psychiatric predictions in a sexually violent predator commitment); *accord* *People v. Bennett*, 182 Cal. Rptr. 473, 497 (Cal. Ct. App. 1982) ("In the context of a petition for an extension of commitment . . . a finding on whether the individual is dangerous to others because of mental illness is essential. Testimony by mental health experts in this context will often be the only way to establish whether such dangerousness exists."); *People v. Devers*, No. A095661, 2002 WL 724931, at *3 (Cal. Ct. App. Apr. 25, 2002) ("The question of whether an individual suffers from a mental disease, disorder or defect that renders him or her a danger to others is not a question of law, but one of fact to be resolved with the assistance of expert testimony.").

195. See *infra* text accompanying notes 362–63 (discussing the use of lay opinion testimony).

196. *Barefoot v. Estelle*, 463 U.S. 880, 992 n.4 (1983) (Blackmun, J., dissenting).

197. See, e.g., *Devers*, 2002 WL 724931, at *3 (holding psychiatric testimony sufficient to support a commitment: "Aside from his testimony, [the psychiatric expert] made reference to numerous records and reports, which appear as part of the court record. These reports reflect that appellant suffers from a mental disorder described variously as, Substance Induced Psychotic Disorder with Hallucinations, In Remission, and Antisocial Personality Disorder with Borderline and Narcissistic Traits. These hear-

proof of other elements of a commitment proceeding. Psychiatric witnesses, psychiatric facts, and psychiatric opinion constitute practical and legal commonplaces at trial, and are neither novel nor unfamiliar.¹⁹⁸ This familiarity assures the exposure of trial judges to the opinions and methodologies of mental health experts on matters relating to the commitment. Moreover, if we assume competent advocacy, the prevalence of these experts assures full and fair testing of the weaknesses and inherent uncertainties of predictive testimony. In short, one can more readily accept the risk of failure in a particular brick if one regularly sees similar bricks placed under stress in other parts of the structure.

b. The Sufficiency of Predictive Expertise

An overwhelming majority of cases holds that expert testimony standing alone without other proof cannot sustain a commitment.¹⁹⁹ Some courts reach this result with only a bare assertion of insufficiency.²⁰⁰ Many others have specifically adopted a prediction-plus-more approach. These courts emphasize both the difficulty of prediction, and the unreliability of predictive expertise, in requiring additional proof as a check on these uncertainties.²⁰¹ In some states, this requirement rests on an argument about

say materials provide ample details about appellant's mental condition and his need for further treatment.").

198. *Compare* *Daubert v. Merrell Dow Pharm., Inc.*, 43 F.3d 1311, 1317 (9th Cir. 1995) ("[O]ne very significant fact to be considered [in determining the reliability of expert testimony] is whether the experts are proposing to testify about matters growing naturally and directly out of research they have conducted independent of the litigation, or whether they have developed their opinions expressly for purposes of testifying."). Judge Kozinski noted that some forms of opinion are so closely tied to the judicial system as to have the courtroom as "a principal theatre of operations," a fact which he would not weigh against determining such an opinion reliable. *Id.* at 1317 n.5. Judge Kozinski's concerns related to the reliability component of *Daubert*, not the relevance component.

199. *See infra* Part IV.D.2-3 (noting that virtually all states require some proof of danger in addition to the expert's opinion).

200. *See, e.g., Broussard v. State*, 827 S.W.2d 619, 622 (Tex. App. 1992) ("We find no evidence that Broussard was likely to cause serious harm to others as a result of her mental illness. Though the expert medical witnesses acknowledged this as a 'possibility' and there is a vague reference to some sort of incident involving another person's child, both experts agreed that it was not likely or a clear imminent risk. Bare psychiatric expert opinion of a 'potential danger' to others is insufficient to support a commitment."); *Mezick v. State*, 920 S.W.2d 427, 430 (Tex. App. 1996) (finding adequate factual basis in testimony of lay witness "expert diagnosis alone is not sufficient to confine a patient for compulsory treatment The expert opinions and recommendations must be supported by a showing of the factual bases on which they are grounded"); *In re Interest and Protection of C.O.*, 65 S.W.3d 175, 181 (Tex. App. 2001) ("Expert testimony is essential . . . but expert diagnosis alone is not sufficient to confine a patient for compulsory treatment The State cannot meet its burden of proof without presenting evidence of the behavior of the proposed patient that provides the factual basis for the expert opinion.").

201. *See, e.g., State v. Stanley*, 843 P.2d 1018, 1020 (Or. Ct. App. 1992) ("Apprehensions and speculation alone are not enough to find a person in need of treatment There must be facts and observations that lead to such a conclusion.") (finding no other evidence sufficient to justify commitment).

a court's institutional role: Courts have the institutional duty to predict danger, not doctors:

A person suspected of mental illness ought not to be deprived of liberty on the basis of expert opinion alone. . . . If the conclusion of a psychiatrist were sufficient for a civil commitment, no necessity would exist for a court, with or without a jury, to make a judicial determination of the facts necessary for involuntary hospitalization.²⁰²

In this vein, courts often stress that a finding of danger requires a balancing of societal and individual interests that is primarily a judicial function:

The determination of dangerousness involves a delicate balancing of society's interest in protection from harmful conduct against the individual's interest in personal liberty and autonomy. This decision, while requiring the court to make use of the assistance which medical testimony may provide, is ultimately a legal one, not a medical one.²⁰³

Even those rare cases that accept bare psychiatric testimony as sufficient stress the judicial function in making the ultimate decision on dan-

202. *Moss v. State*, 539 S.W.2d 936, 949 (Tex. Civ. App. 1976) (finding insufficient other proof of danger to justify commitment).

203. *State v. Krol*, 344 A.2d 289, 302 (N.J. 1975) (reviewing and reforming New Jersey's then existing commitment scheme, and requiring proof of mental illness and danger by clear and convincing evidence); followed in "*Matter of*" *Newsome*, 424 A.2d 222, 225 (N.J. Super., 1980); "*Matter of*" *D.C.*, 679 A.2d 634, 648 (N.J. 1996). See also *Hatcher v. Wachtel*, 269 S.E.2d 849, 852 (W. Va., 1980) (finding sufficient evidence to commit, and rejecting an argument that West Virginia law required a finding of "imminent danger"); *Application of Noel*, 601 P.2d 1152, 1166 (Kan. 1979) (affirming an order denying release to a criminally committed patient). "We conclude the determination of whether the patient continues to be dangerous to himself, herself or others is a legal rather than a medical decision. . . . A medical opinion as to dangerousness, even if undisputed by other medical opinions, is not conclusive upon the court and must be weighed with the other evidence." *Id.*

ger.²⁰⁴ Those cases that reject psychiatric opinion as factually insufficient also stress the ultimate role of the court to find danger.²⁰⁵

These cases assume that predictive expertise lacks reliability; it comes as no surprise that they would find expert opinion insufficient on its own to sustain the stringent burden of proof. The fact of insufficiency adds another influence on the assessment of admissibility. If a jurisdiction has repeatedly held predictive testimony insufficient, a trial judge ruling on admissibility can assume that the proponent of the testimony will need to come prepared with other evidence. An expert prediction will not be the only brick; and a judge can consider using it (or not), knowing that it will not bear the whole weight of decision.

These cases also add a theme to the discussion. Courts, not mental health professionals, have the power to order commitments and bring specialized tools to the task: the adversarial presentation of evidence, the weighing of credibility, the balancing of societal interests, and the "tribunal's experience with the mainsprings of human conduct."²⁰⁶ One might

204. See, e.g., *Lindquist v. Bisch*, 542 N.W.2d 138, 140-41 (S.D. 1996) (affirming a commitment based solely on psychiatric testimony at trial, and rejecting a claim that the appellate court perform a de novo review of the evidence).

An individual's potential dangerousness is an essentially factual inquiry, founded on the application of the fact-finding tribunal's experience with the mainsprings of human conduct Assessing a person's dangerous propensities does not require the consideration of legal concepts and involve the exercise of judgment about the values underlying legal principles—the hallmarks of an essentially legal inquiry Furthermore, in assessing an individual's dangerous propensities, "the trial court is in the best position to determine the credibility of the testifying witnesses and weigh the evidence.

Id. at 140-41 (citations and quotation marks omitted). See also *People v. Devers*, No. A095661, 2002 WL 724931, at *3 (Cal. Ct. App. Apr. 25, 2002) (unpublished opinion affirming commitment). "Expert testimony is a substantial factor to be considered in assessing the sufficiency of the evidence. By itself, it provides a sufficient basis for the trial court's conclusion." *Id.*

205. *Riley v. State*, 396 N.W.2d 595, 598-99 (Minn. Ct. App. 1986) (affirming a trial court's refusal to release a criminal committee, in which the court found future danger despite contrary expert opinion).

In *Johnson v. State*, 331 N.W.2d 757 (Minn. 1983) the supreme court held that the district court was not required to rely on a psychiatrist's expert testimony that the petitioner's early release would not present danger to the public. Similarly, here the court was entitled to listen to and evaluate, but not necessarily accept the expert's testimony. Dr. Osekowsky's testimony was clearly admissible, but its weight and credibility were for the trial court.

Id. at 599. See also *State v. Putnoki*, 510 A.2d 1329, 1335 (Conn. 1986) ("Although psychiatric testimony as to the defendant's condition may form an important part of the trial court's ultimate determination, the court is not bound by this evidence It may, in its discretion, accept all, part, or none of the experts' testimony."); followed in *State v. Jacob*, 798 A.2d 974, 986 (Conn. App. Ct. 2002).

206. *Lindquist*, 542 N.W.2d at 141.

question the adequacy of the judicial forum as a predictor of danger.²⁰⁷ Yet it seems unlikely that society will transfer the function of civil commitment from courts to other professionals, or that it will abandon litigation as the vehicle for resolving the conflict between private and public values embodied in commitment cases. Courts have willingly accepted their role, and have developed some confidence in their competence to make the required findings. This assumption of institutional competence to predict danger, with or without expertise, constitutes another element of the context within which *Daubert's* fit criterion will operate.²⁰⁸

c. Other Evidence

If predictive testimony cannot bear the weight of fact-finding, how has the task of prediction and the unreliability of expertise affected other proof in commitments? Many courts have made no effort to structure the other proofs of danger in civil commitment, leaving it to fact-finders in individual cases to assess the strength of the other evidence in relation to the expert testimony.²⁰⁹ Yet some courts have imposed restrictions on this evidence, by categorizing different types of evidence which they require in addition to expert testimony. These categories include "overt" acts, "recent" acts, or "prognostic evidence." Still other courts accomplish the same end by requiring complex fact-finding with multiple proofs.

(i) Overt Acts

The "overt act" rule requires the proponent of commitment to prove some act of the respondent, committed before the proposed commitment, to serve as a basis from which to infer future dangerousness. The rule operates as a rule of sufficiency; it bars commitment without proof of the overt act. The rule often rests on a statutory footing;²¹⁰ however, it

207. Eric S. Janus & Paul E. Meehl, *Assessing the Legal Standard for Predictions of Dangerousness in Sex Offender Commitment Proceedings*, 3 PSYCHOL., PUB. POL'Y & L. 33, 41 (1997) (suggesting that the judges commit individuals whose probability of recidivism falls between thirty and seventy-five percent).

Even on the most optimistic assumptions, the actual probability standards used by the courts do not reach the 75% mark. With realistic but still optimistic assumptions, predictions of future violence could exceed the 50% "likely" rate. On pessimistic assumptions, courts are applying a standard that commits people with probabilities of recidivism as low as 30%.

Id.

208. See Part I.C.5 *infra* (discussing similarities between clinical predictions and judicial fact-finding as a basis for the judicial acceptance of these opinions).

209. See *supra* notes 188–200 (describing the requirement of other evidence in various states).

210. See, e.g., MONT. CODE ANN. § 53-21-126(2) (2002) ("Imminent threat of self-inflicted injury or injury to others must be proved by overt acts or omissions, sufficiently recent in time as to be material and relevant as to the respondent's present condition."); GA. CODE ANN. § 37-3-1(9.1) (1995) ("'Inpatient' means a person who is mentally ill and: (A) (i) Who presents a substantial risk of imminent harm to that person or others, as manifested by either recent overt acts or recent expressed threats of violence which present a probability of physical injury to that person or other persons . . .").

originated in decisions which rested on constitutional grounds.²¹¹ The rule is also a minority rule; the Supreme Court has never addressed it, and many states explicitly reject it.²¹²

The requirement of overt act evidence rests on a central assumption: that proof of such an act creates a stronger basis from which a fact-finder can infer future dangerousness. “[C]ivil confinement can be justified . . . if . . . dangerousness is based upon a finding of a recent overt act, attempt or threat to do substantial harm to oneself or another.”²¹³ More-

211. *Lessard v. Schmidt*, 349 F. Supp. 1078, 1093 (E.D. Wis. 1972), *vacated* 414 U.S. 473 (1974) (“we believe civil confinement can be justified in some cases if the proper burden of proof is satisfied and dangerousness is based upon a finding of a recent overt act, attempt or threat to do substantial harm to oneself or another”); *Lynch v. Baxley*, 386 F. Supp. 378, 391 (M.D. Ala. 1974) (“To confine a citizen against his will because he is likely to be dangerous in the future, it must be shown that he has actually been dangerous in the recent past and that such danger was manifested by an overt act . . .”); *Doremus v. Farrell*, 407 F. Supp. 509, 514–15 (D. Neb. 1975) (striking Nebraska commitment statute: “Due process and equal protection require that the standards for commitment must be (a) that the person is mentally ill and poses a serious threat of substantial harm to himself or to others; and (b) that this threat of harm has been evidenced by a recent overt act or threat. The threat of harm to oneself may be through neglect or inability to care for oneself.”); *Stamus v. Leonhardt*, 414 F. Supp. 439, 451 (S.D. Iowa 1976) (striking Iowa commitment statute: “This Court therefore holds that the commitment standards of Chapter 229 of the Code violated substantive due process by not requiring that subjects pose a serious threat to themselves or others, as evidenced by a recent overt act, attempt or threat”); *Suzuki v. Yuen*, 617 F.2d 173, 178 (9th Cir. 1980) (requiring overt act). *See also* *Cross v. Harris*, 418 F.2d 1095, 1100 (D.C. Cir. 1969) (upholding a sexual psychopath commitment statute, and indicating a preference for other evidence in addition to expert testimony); *See also* Note, *Overt Dangerous Behavior as a Constitutional Requirement for Involuntary Commitment of the Mentally Ill*, 44 U. CHI. L. REV. 562, 584 (1977) (discussing litigation asserting a constitutional requirement of overt acts).

212. *See* *People v. Sansone*, 309 N.E.2d 733, 740 (Ill. App. Ct. 1974) (declining to overturn an Illinois statute for lack of an “overt act” requirement); *United States ex rel. Mathew v. Nelson*, 461 F. Supp. 707, 712 (N.D. Ill. 1978) (declining to overturn the same Illinois statute); *Matter of Monroe*, 270 S.E.2d 537, 541 (N.C. Ct. App. 1980) (“This Court has not required ‘overt acts’ under the former standard of ‘imminent’ danger and the present statutory definition of ‘dangerous to others’ does not require a finding of ‘overt acts.’”); *Project Release v. Prevost*, 722 F.2d 960, 973 (2d Cir. 1983) (refusing to overturn a New York statute for lack of an “overt act” requirement); *In re L.R.*, 497 A.2d 753, 755–57 (1985) (declining to adopt an overt act requirement); *People v. Stevens*, 761 P.2d 768, 774 (Colo. 1988) (relying on *In re L.R.* and declining to adopt an overt act requirement). *Compare* *Matter of Gattson*, 593 P.2d 423, 426 (Kan. Ct. App. 1979) (declining formal adoption of an overt act requirement, but stating a belief that “a showing of present dangerousness will normally require evidence of a recent act, attempt, threat or omission of a serious nature.”), *with* *In re Harris*, 654 P.2d 109, 113 (Wash. 1982) (interpreting WASH. REV. CODE ANN. § 71.05.020 (West 2002) to require “a showing of a substantial risk of physical harm as evidenced by a recent overt act”); *In re LaBelle*, 728 P.2d 138, 144 (Wash. 1986) (refusing to extend the overt act requirement to proof that the respondent is “gravely disabled”).

Note also that, in 1996, the District Court in *Lynch v. Sessions*, 942 F. Supp. 1419, 1427 (M.D. Ala. 1996), dissolved the injunction under which Alabama had operated its civil commitment process, after the state had passed a new commitment statute. The new statute contained no “overt act” requirement, and subsequent Alabama cases refused to imply such a requirement. *See, e.g.*, *Webster v. Bartlett*, 709 So. 2d 1226, 1228 (Ala. Civ. App. 1997).

213. *Lessard*, 349 F. Supp. at 1078.

over, the requirement arises from concerns about the adequacy of expert predictions of danger, and reasserts the judicial role:

While the actual assessment of the likelihood of danger calls for an exercise of medical judgment, the sufficiency of the evidence to support such a determination is fundamentally a legal question. A mere expectancy that danger-productive behavior might be engaged in does not rise to the level of legal significance when the consequence of such an evaluation is involuntary confinement.²¹⁴

Even courts that have rejected the "overt act" requirement draw a clear connection between the unreliability of expert opinion and the need for evidence of an overt act. In *United States ex rel. Mathew v. Nelson*, an Illinois District Court rejected a due process challenge to an Illinois statute that lacked an overt act requirement. Appellants had cited the first generation critique of commitments, then in full flower.²¹⁵ The court recognized the force of this line of argument, but found no evidence that an overt act requirement would help: "No study has attempted to measure the extent to which the predictability of dangerousness is enhanced by a history of a recent overt act."²¹⁶ In rejecting a *per se* attack on the statute, the court advised respondents instead to attack the unreliability of predictive opinion on a case-by-case basis:

It may well be that in most cases the psychiatric determination necessary to support the finding of reasonable expectation that the statute requires could not be made in the absence of an overt act, just as it could not be made in the absence of other facts found in the patient's history or discovered in examining him. In those cases, the evidence will not justify a determination of dangerousness.²¹⁷

The "overt act" rule requires separate proof but does not affect the basis of the expert's opinion. Some courts approximate an overt act requirement by requiring the proponent to show that the expert based his opinion on some form of first-hand knowledge.²¹⁸ Other courts have gone further, by requiring the proponent of an expert prediction to prove the factual basis for the opinion through separate evidence. In Illinois, the

214. *Lynch v. Baxley*, 386 F. Supp. at 391 (footnote omitted).

215. See *supra* Part I.B.2 (discussing the scholarly critique of predictions).

216. *United States ex rel. Mathew v. Nelson*, 461 F. Supp. 707, 710 (N.D. Ill. 1978).

217. *Id.* at 711.

218. See, e.g., *Judicial Commitment of J.M.*, 560 So. 2d 100, 102-03 (La. App. 1990). The appellate court reversed a commitment in which the only evidence of defendant's behavior consisted of the testimony of a psychiatrist about acts which he had not seen. While the court noted that the medical history on which the psychiatrist relied included evidence admissible under the "medical diagnosis and treatment" exception to the hearsay rule, it refused to find them sufficient to support the commitment, in the absence of testimony by observers with first hand knowledge. *But see* *People v. Lane*, 581 P.2d 719, 722 (Colo. 1978) (affirming a commitment where the "only evidence at the hearing was the testimony of" a psychiatrist, who testified directly from the respondent's medical record at the hospital).

proponent of a psychiatric opinion on danger must prove the facts forming the basis for the expert testimony by clear and convincing evidence:

We recognize that the medical science of predicting future dangerousness is inexact and that a court is not required to wait until respondent harms himself or another before ordering involuntary admission. However, of primary concern in an involuntary admission proceeding is the right of the individual to be free from unjustified and unreasonable confinement. Both the facts upon which the medical opinion is based and the medical testimony upon which the decision to admit is based must be established by clear and convincing evidence.²¹⁹

Texas courts apply a similar rule: "Expert diagnosis alone is not sufficient to confine a patient for compulsory treatment. The expert opinion and recommendations must be supported by a showing of the factual bases on which they are grounded."²²⁰

(ii) *Recent Acts*

Where the jurisdiction does require an overt act, courts have sometimes imposed an additional requirement that the prior act be "recent,"²²¹ on the theory that the recency of the act makes it more probative of the likelihood of danger.²²² "In determining whether a person is dangerous, the focus must be on the subject's condition at the time of the hearing. Actions and statements of a person alleged to be mentally ill and dangerous which occur prior to the hearing are probative of the subject's present mental condition."²²³ No court uses a bright-line rule to determine how recent is recent. Instead, courts have phrased the rule in terms of whether "evidence of dangerousness [is] sufficiently probative to pre-

219. *In re Cochran*, 487 N.E.2d 389, 390-92 (Ill. Ct. App. 1985) (citations omitted) (reversing a commitment order for failure to prove basis of expert's opinion). *Accord* *Matter of Gregorovich*, 411 N.E.2d 981, 987 (Ill. App. Ct. 1980) (affirming commitment); *In re Cutsinger*, 186 Ill.App.3d 219, 542 N.E.2d 414, 419 (Ill. App. Ct. 1989) (reversing commitment).

220. *Johnstone v. State*, 961 S.W.2d 385, 388 (Tex. App. 1997) (citations omitted). *Accord* *Mezick v. State*, 920 S.W.2d 427, 430 (Tex. App. 1996); *see In re Breeden*, 4 S.W.3d 782, 784 (Tex. Ct. App. 1999). Note that these cases do not address an issue made explicit by recent revisions to the Federal Rules of Evidence: whether and when the proponent of expert testimony can introduce otherwise inadmissible evidence on the theory that it reflects the basis for the expert's opinion. FED. R. EVID. 703. To the extent that these cases require introduction as basis of inadmissible evidence (such as hearsay reports), these cases would significantly diverge from the Federal approach, which excludes inadmissible basis information, subject to a balancing of prejudice and probative value. *Id.*

221. The District Court in *Lessard v. Schmidt* required that "civil confinement can be justified in some cases if the proper burden of proof is satisfied and dangerousness is based upon a finding of a recent overt act . . ." 349 F. Supp. 1078, 1078 (E.D. Wis., 1972) (emphasis added).

222. *See, e.g., In re Interest of Blythman*, 302 N.W.2d 666, 671-73 (Neb. 1981) (affirming a recency requirement, but upholding a sexually violent predator civil commitment which relied on five-year-old acts).

223. *Id.* at 671.

dict future behavior and the subject's present state of dangerousness."²²⁴ Even states that have rejected a recency requirement note that trial courts will need to assess the probative value of the respondent's past history against the other available evidence of danger.²²⁵

(iii) *Prognostic Evidence*

Other courts regulate the strength of the evidence required for commitment by insisting that the proponent of commitment use past acts that are "prognostic" of future violence. This rhetoric seeks to mark off a sub-category of actions that justify a finding of future danger: "A court can use what has happened in the past as 'prognostic' evidence to help predict future conduct."²²⁶ The phrase "prognostic evidence" does not assume any particular degree of reliability, and may even acknowledge that the court's abilities to predict do not exceed those of experts: "Just as we recognize 'the fallibility of psychiatric opinions on the issue of whether a person meets the criteria for involuntary commitment . . . we recognize the fallibility of particular acts offered as proof of dangerous mental con-

224. *Id.* at 672. *But see* *In the Matter of D.D.*, 920 P. 2d 973, 975 (1996) (affirming a commitment order based on recent threats: "imminent threat of self-inflicted injury or injury to others must be evidenced by overt acts sufficiently recent as to be material and relevant to the person's present condition."); *Davis v. North Carolina Dept. of Human Res.*, 465 S.E.2d 2, 8 (N.C. Ct. App. 1995) (The court affirmed a commitment order based on acts over two months old:

We therefore construe the term "recent past" to mean "relevant past" and as such determine that the violent acts . . . within the six months prior to the district court hearing to be the "relevant past." These acts are relevant because they occurred close enough in time to the district court hearing to have probative value on the ultimate question before the court of whether there was a "reasonable probability that such [violent] conduct [would] be repeated." 1979 N.C. Sess. Laws ch. 915, § 1; see N. C. GEN. STAT. § 8C-1, Rule 401 (1992) (defining relevant evidence). We do not attempt to define the term with any greater degree of preciseness and each case must be viewed on its own facts in determining whether violent acts are relevant to the inquiry of involuntary commitment. The courts will be the ultimate judge of whether the conduct occurs within a relevant time.

Id. *But see* 50 PA. CONS. STAT. ANN. § 7301(b) (2001) ("Clear and present danger to others shall be shown by establishing that within the past 30 days the person has inflicted or attempted to inflict serious bodily harm on another and that there is a reasonable probability that such conduct will be repeated.").

225. *See, e.g., In re L.R.*, 497 A.2d 753, 756 (Vt. 1985) (rejecting respondent's claim that Vermont law incorporated a recency requirement, and noting that the other constitutional protections assure a sufficiently strong inference of danger: "Overt acts occurring shortly before the hearing may be given more weight than remote acts, but where the evidence is otherwise sufficient, recent acts are not a prerequisite to involuntary commitment.").

226. *In Interest of J.S.*, 545 N.W.2d 145, 149 (N.D. 1996) (affirming commitment; citing *In Interest of R.N.*, 513 N.W.2d 370, 372 (N.D. 1994) (respondent claimed that "a court should not speculate that she may be a risk in the future based on her past treatment history. However, a court is entitled to consider what has happened in the past as relevant 'prognostic' evidence of what is likely to occur in the future.")). *See also In re Renz*, 507 N.W.2d 76, 78 (N.D. 1993) (stating that patient's history may be predictor of dangerousness to self or others, and noting that "we consider prognostic evidence in other proceedings which bear on an individual's rights").

dition."²²⁷ In effect, the language once again reasserts the shift from psychiatric to judicial definitions of danger: "[T]he Legislature shifted from a focus on the necessarily imprecise element of psychiatric prognostication to an emphasis on the evidentiary underpinnings of the diagnosis; from that which is least capable of proof, to that which is most capable of proof."²²⁸

(iv) *Multiple Proofs*

Other states require multiple sources of proof (including but not limited to expert testimony) to support a conclusion of danger. This approach often finds expression in conclusory terms: the trial court should consider the "totality of the circumstances."²²⁹ But on occasion a court will spell out the multiple factors which undergird a "totality of the circumstances" test. In *In re Burton*, for example, the court explained:

Factors which are to be considered by the court in a commitment hearing include, but are not limited to, the following: (1) whether, in the court's view, the individual currently represents a substantial risk of physical harm to himself or other members of society; (2) psychiatric and medical testimony as to the present mental and physical condition of the alleged incompetent; (3) whether the person has insight into his condition so that he will continue treatment as prescribed or seek professional assistance if needed; (4) the grounds upon which the state relies for the proposed commitment; (5) any past history which is relevant to establish the individual's degree of conformity to the laws, rules, regulations and values of society; and (6) if there is evidence that the person's mental illness is in a state of remission, the court must also consider the medically suggested cause and degree of the remission and the probability that the individual will continue treatment to main-

227. *In re Fasi*, 567 A.2d 178, 182 (N.H. 1989) (holding that a civil commitment action does not permit a justification defense, since prior acts are offered not as the basis of the commitment order, but as "prognostic" proof).

228. *People v. Super. Ct. (Dodson)*, 196 Cal. Rptr. 431, 436 (Cal. Ct. App. 1983) (upholding a commitment statute against a constitutional challenge that it focused on the past rather than the future).

229. *See, e.g., People in Interest of King*, 795 P.2d 273, 275 (Colo. Ct. App. 1990) ("[B]ased on the totality of the evidence including the psychologist's professional opinion as to respondent's potentiality for danger, the court was justified in ordering continued involuntary treatment.").

tain the remissive state of his illness should he be released from commitment.²³⁰

Such a requirement reduces the probative importance of the expert opinion predicting danger, and therefore the risks caused by its admission, by emphasizing other evidence and assuring broad evidentiary support for the statutory finding. Requiring multiple proofs serves to counteract the weaknesses of predictive expertise.²³¹ The requirement of multiple proofs also bears interesting similarities to the approach of experts using guided clinical assessments that combine in-person appraisal with consideration of a wide range of objective "risk factors."²³²

Supplementary proof requirements form part of the evidentiary context within which to fit predictive expertise. The requirement of some other evidence, the effort to regulate the strength of the inference to be drawn from that evidence, and the effort to assure a varied, complex pool of information serve a vital purpose: they justify the fact-finder's consideration of predictive evidence by providing stronger and more diverse data against which to gauge the prediction. These rules can, of course, find justification in their own merits, regardless of the weaknesses of expert predictions. The fact remains, however, that many courts justify these approaches by reference to the weakness of predictive expertise.

230. *In re Burton*, 464 N.E.2d 530, 534 (1984) (affirming a commitment order); see also *State v. Mullins*, No. 73315, 1999 WL 148479, at *4 (Ohio Ct. App. Mar. 18, 1999). See also *In re Linehan*, 518 N.W.2d 609, 614 (Minn. 1994):

Where, [as here,] utter uncontrollability of sexual impulses is found, . . . the court, in predicting serious danger to the public, should consider six factors: (a) the offender's relevant demographic characteristics . . . ; (b) the person's history of violent behavior (with special attention to recency, severity, and frequency of violent acts); (c) the base rate statistics for violent behavior among individuals of this person's background; (d) the sources of stress in the environment; (e) the similarity of the present or future context to those contexts in which the person has used violence in the past; and (f) the person's record with respect to sex therapy programs.

Id. See also *In re Preston*, 629 N.W.2d 104, 115 (Minn. Ct. App. 2001).

231. See, e.g., *State v. Putnoki*, 510 A.2d 1329, 1335-36 (Conn. 1986):

Although a trial court may choose to attach special weight to the testimony of medical experts at a hearing to determine mental status, the ultimate determination of mental illness and dangerousness is a legal decision (citations omitted). Partly because definitions of dangerousness are necessarily vague . . . and partly because there are no "psychological or physical signs or symptoms which can be reliably used to discriminate between the potentially dangerous and the harmless individual," . . . psychiatric predictions of future dangerousness are tentative at best and are frequently conceded, even within the profession, to be unreliable. . . .

. . . In reaching its difficult decision, the court may and should consider the entire record available to it, including the defendant's history of mental illness, his present and past diagnoses, his past violent behavior, the nature of the offense for which he was prosecuted, the need for continued medication and therapy, and the prospects for supervision if released.

232. See Part I.B.1 *infra* (discussing guided clinical and adjusted actuarial predictions).

4. *Defining Danger*

So far, this Article has assumed a uniform standard of danger,²³³ a reasonable assumption given the uniform constitutional requirement to find danger. All states have enacted statutory definitions, however, and these definitions reflect distinctive approaches. Some jurisdictions, for example, use uncomplicated standards: a person who poses a danger to self or others as a result of mental illness.²³⁴ However, greater refinement is common with respect to some critical elements of proof: how soon the potential harm might occur (“immediacy”); how likely it is to occur (“probability”); and what and how severe the harm could be (“nature and severity”). Assessing the legal definitions of danger permits us to identify another element in *Daubert*’s fit requirement: the degree of similarity between the inferences inherent in the expert’s opinion and the inferences required by the legal standard. As we shall see, the statutory definitions of danger require fact-finding that is strikingly similar to the methodology used in forming clinical predictions of danger.

a. *Immediacy of Danger*

Some statutes require that the prospective danger be “imminent” or “immediate.” For example, Georgia defines a person subject to commitment as someone “who is mentally ill . . . and presents a substantial risk of imminent harm to that person or others [or] is so unable to care for that person’s own physical health and safety as to create an imminently life-endangering crisis.”²³⁵ Other formulations speak of “immediate danger”;²³⁶ “clear and present danger”;²³⁷ and “danger in the near future.”²³⁸

233. More general definitions of danger include an inherently predictive element in them, defining it to include the “[l]iability or exposure to harm or injury; the condition of being exposed to the chance of evil; risk, peril” or the “[p]ower (of a person. . .) to inflict physical injury.” IV OXFORD ENGLISH DICTIONARY 240–41 (2d ed. 1989).

234. See, e.g., ALASKA STAT. § 47.30.735(c) (2002) (“mentally ill and as a result . . . likely to cause harm to the respondent or others or . . . gravely disabled”); CAL. WELF. & INST. § 5256.6 (1998) (“as a result of a mental disorder or impairment by chronic alcoholism, a danger to others, or to himself or herself, or gravely disabled”); MD. CODE ANN., Health Gen § 10-617(3) (2000) (“presents a danger to the life or safety of the individual or of others”); NEV. REV. STAT. ANN. 433A.310 (2000) (“because of that [mental] illness, is likely to harm himself or others if allowed his liberty”); OR. REV. STAT. § 426.005(1)(d)(A) (1995) (“because of a mental disorder, is . . . (d)angerous to self or others”).

235. GA. CODE ANN. § 37-3-1 (9.1)(A)(i) (1995). See also MONT. CODE ANN. § 53-21-12 (2001) (“imminent threat of injury to the respondent or to others”); VA. CODE ANN. § 37.1-67.3 (1996) (“the person presents an imminent danger to himself or others as a result of mental illness”).

236. UTAH. CODE ANN. § 62A-12-234(10)(b) (2000) (“because of the proposed patient’s mental illness he poses an immediate danger of physical injury to others or himself”).

237. ALA. CODE § 22-52-10.4(a)(ii) (1997) (“[A]s a result of the mental illness the respondent poses a real and present threat of substantial harm to self and/or others.”); DEL. CODE ANN. tit. 16 § 5001 (1995) (“a real and present threat, based upon manifest indications, that such person is likely to commit or suffer serious harm”); OHIO REV. CODE ANN. § 5122.01(B)(2) (West 2001) (“a substantial risk of physical harm to others as manifested by evidence of recent homicidal or other violent behavior, evi-

These immediacy requirements reflect a concern about the indefiniteness of predictions; an “imminence” approach appears to assume that requiring danger in the near future is likely to produce more reliable predictions.²³⁹ To be sure, many courts have found that due process does not require “imminence”; in these cases, the relevant statutes produce sufficiently reliable results for constitutional purposes.²⁴⁰ Where it exists, however, an immediacy requirement narrows the range of permissible commitments, and thus creates a commitment scheme that constrains the uncertainties of expert prediction.

dence of recent threats that place another in reasonable fear of violent behavior and serious physical harm, or other evidence of present dangerousness”). Note that in at least one state, the statute further defines the language “clear and present danger” in such a way as to eliminate a requirement of imminence. ARK. CODE ANN. § 20-47-207(c)(1)(A) (1987) (“[C]lear and present danger” means “serious bodily injury . . . and . . . a reasonable probability that the conduct will be repeated.”).

238. FLA. STAT. ch. 394.467(1)(b) (2002) (“substantial likelihood that in the near future he or she will inflict serious bodily harm on himself or herself or another person”); LA. REV. STAT. ANN. § 28:2 (2001) (danger to others: “a person whose behavior or significant threats support a reasonable expectation that there is a substantial risk that he will inflict physical harm upon another person in the near future”); *but cf.* LA. REV. STAT. ANN. §28.2(4) (2001) (danger to self: “a person whose behavior, significant threats or inaction supports a reasonable expectation that there is a substantial risk that he will inflict physical or severe emotional harm upon his own person”); N.J. STAT. ANN. § 30:4-27.2 (1997) (“by reason of mental illness there is a substantial likelihood that the person will inflict serious bodily harm upon another person or cause serious property damage within the reasonably foreseeable future”).

239. *See State v. Krol*, 344 A.2d 289, 302 (N.J. 1975).

It is not sufficient that the state establish a possibility that defendant might commit some dangerous acts at some time in the indefinite future. The risk of danger, a product of the likelihood of such conduct and the degree of harm which may ensue, must be substantial within the reasonably foreseeable future.

Id. *See also supra* notes 133–36 (discussing amicus brief of American Psychiatric Association, which claimed that psychiatrists have greater skill at making short-term predictions).

240. *In re P.S.*, 702 A.2d 98, 104–05 (Vt. 1997) (rejecting a claim that the federal or state constitutions required a showing of present danger, and holding instead that a statute which required a showing of future danger comported with due process); *In re LaBelle*, 728 P.2d 138, 146 (Wash. 1986) (upholding a statute which permitted commitment of the “gravely disabled”, against a challenge that the constitution required a finding of “imminent” danger); *Moore v. Wyoming Medical Center*, 825 F. Supp. 1531, 39 (D. Wyo. 1993) (upholding a Wyoming statutory revision which had eliminated an “imminent” danger requirement in favor of a “substantial probability” requirement). In *Moore*, the District Court articulated a judicial role rationale for distinguishing between judicial and medical notions of danger:

[N]either the medical nor the legal profession has accurately defined or predicted “dangerousness.” . . . Nonetheless, courts must attempt to define and evaluate the contours of “dangerousness,” especially where, as in this case, a genuine issue of fact exists regarding whether the detainee was dangerous at the time the state detained her.

Id.

b. Probability of Danger

Some state statutes do not speak of the likelihood of danger at all, requiring only “danger of harm to self or others.”²⁴¹ Most states, however, do specify the strength of the required probability of harm. Formulations range in strength from a “likelihood” of harm;²⁴² through a “reasonable expectation” of harm;²⁴³ to a “probability” of harm; to a “substantial probability” of harm.²⁴⁴ No empirical relationship has been shown between different articulations of probability and the frequency of finding

241. See statutes cited *supra* note 234. As noted previously, some courts appear to finesse the probability issue entirely, by insisting that the proper inquiry is whether the respondent is “presently dangerous.” See discussion of recent overt acts *supra* at Part IV.D.3. Moreover, even statutes which state some degree of probability for danger to others often define danger to self in terms of a present condition. See, e.g., MICH. COMP. LAWS ANN. § 330.1401(b) (1999) (inability to meet physical needs: “a result of that mental illness is unable to attend to those of his or her basic physical needs such as food, clothing, or shelter that must be attended to in order for the individual to avoid serious harm”) (emphasis added). The ambiguity may derive from an ambiguity in the very concept of dangerousness; for example, one dictionary defines “dangerous” as “[f]raught with . . . risk; causing or occasioning danger; perilous, hazardous, risky, unsafe.” IV OXFORD ENGLISH DICTIONARY 242 (2d ed. 1989).

242. D.C. CODE ANN. § 21-545(a) (2001 & Supp. 2003) (“likely to injure himself or other persons”); HAW. REV. STAT. § 334-1 (1993) (“likely to do substantial physical or emotional injury”); IOWA CODE § 229.1 15.a (2000) (“likely”); KAN. STAT. ANN. § 59-2946(f)(3) (2002) (“likely”); N.M. STAT. ANN. § 43-1-3 (Michie 1978) (“more likely than not”).

243. 405 ILL. COMP. STAT. § 5/1-119 (1997) (“reasonably expected to inflict serious physical harm”); MICH. COMP. LAWS ANN. 330.1401(9)(a) (1999) (“reasonably . . . expected”); N.C. GEN. STAT. § 122C-3 (2001) (“reasonable probability”); N.D. CENT. CODE § 25-03.1-02 (2002) (“reasonable expectation”); 50 PA. STAT. ANN. tit. § 7301 (2001) (“reasonable probability”); S.D. CODIFIED LAWS § 27A-1-1(5)(a) (1999) (“reasonable expectation”).

244. CONN. GEN. STAT. § 17a-495 (1992) (“substantial risk”); FLA. STAT. ch. § 394.467(1)(a)2.b. (2002) (“substantial likelihood”); GA. CODE ANN. 37-3-1(9.1)(A)(i) (1995) (“substantial risk”); IDAHO CODE § 66-317(k)(1)-(2) (2000) (“substantial risk”); MINN. STAT. ANN. § 253B.02(2) (West 2003) (“substantial likelihood”); NEB. REV. STAT. § 83-1009(1) (1999) (“substantial risk”); N.H. REV. STATE. ANN. § 135-C:34 (1995) (“potentially serious likelihood”); N.Y. MENTAL HYG. § 9.01 (2002) (“substantial risk”); OHIO REV. CODE ANN. § 5122.01 (2001) (“substantial risk”); R.I. GEN. LAWS § 40.1-5-2 (1997) (“substantial risk”); TENN. CODE ANN. § 33-6-501 (2001) (“substantial likelihood”); WASH. REV. CODE § 71.05.020 (2002) (“substantial risk”); WIS. STAT. ANN. § 51.20 (1997) (“substantial probability”); WYO. STAT. ANN. § 25-10-101 (2001) (“substantial probability”). Compare LA. REV. STAT. ANN. § 28:2 (2001) (“a reasonable expectation that there is a substantial risk”) (note that in at least a few jurisdictions, the stated probability of danger for one kind of danger is stronger than for another. See, e.g., MASS. GEN. LAWS ANN. ch. 123 § 1 (West 1986) (gravely disabled: “a very substantial risk of physical impairment or injury to the person himself as manifested by evidence that such person’s judgment is so affected that he is unable to protect himself”), with MASS. GEN. LAWS ch. § 1 (West 1986) (danger to others: “substantial risk”); compare ME. REV. STAT. ANN. tit. 34-B § 3801 (West 1964) (“[a] reasonable certainty that severe physical or mental impairment or injury will result to the person. . .”), with ME. REV. STAT. ANN. tit. 34-B § 3801 (West 1964) (“substantial risk”); N.J. STAT. ANN. 30:4-27.2(h) (danger to self: “probable”), and N.J. STAT. ANN. 30:4-27.2(i) (danger to others or property: “substantial likelihood”).

danger.²⁴⁵ Appellate courts usually describe the evidence at hand, and then declare whether or not it meets the statutory standard.²⁴⁶

In doing so, courts use discretionary, non-statistical, and non-actuarial methods; no court has reduced the probability component of dangerousness to a statistical formula.²⁴⁷ Even courts that have passed on the admissibility of actuarial instruments for predicting danger have insisted that the legal determination reflects a balancing of policies, not the quantification of risks.²⁴⁸ Courts uniformly find that assessing the probability of danger requires a discretionary balancing of the individual's right to act freely and society's need for protection from harmful behavior.²⁴⁹ Such statements thus describe fact-finding on probability that requires the weighing of public and private interests, and not fact-finding dominated by the precise quantification of risks.

c. *Nature and Degree of Potential Harm*

Statutory definitions of danger also vary substantially in the nature and degree of prospective harm that the proponent must establish. For example, as to "danger to others," the simplest formulation leaves the nature of the harm undefined.²⁵⁰ Some states specify physical (or "bod-

245. See John Monahan & Eric Silver, *Judicial Decision Thresholds for Violence Risk Management*, 2 INT'L J. OF FORENSIC MENTAL HEALTH 2003, No.1 1-6 (2003) (describing the results of a survey of twenty-six trial judges, asking them to set the decision threshold for instituting short-term civil commitment as a "danger to others," using assessments of risk derived from the Macarthur Risk Assessment instrument).

246. See, e.g., *In re Barnard*, 616 N.E.2d 714, 730 (Ill. App. Ct. 1993) (affirming a commitment under a "reasonable expectation" standard); *In re Pollard*, No. A-92-863, 1993 WL 183594, at *8 (Neb. App. June 1, 1993) (unpublished opinion) (affirming a commitment on a "substantial risk" standard).

247. See *People v. Super. Ct. (Ghilotti)*, 44 P.3d 949, 953-54 (Cal. 2002) (addressing the meaning of the term "likely to engage in acts of sexual violence" as a standard for initial evaluation prior to formal commitment). The California Supreme Court discussed various verbal formulae for assessing the meaning of the term "likely," and settled on one which required initial evaluators to find "a *serious and well-founded risk*" of reoffense. *Id.* at 968. The court clarified that this did not mean "a greater than 50 percent chance the person would reoffend," *id.* at 970, but did mean "more than the mere *possibility*." *Id.* at 972. While the court noted that the term must be interpreted in light of the statutory purpose, it also advised interpretation "in light of the 'difficulties inherent in predicting human behavior' . . . particularly in mathematical terms." *Id.* at 971.

248. See *infra* Part I.B. (discussing cases which address statistical methods of assessing danger).

249. See, e.g., *Pollard*, 1993 WL 183594, at *8 (affirming a commitment under a substantial probability standard, and rejecting an argument that a psychiatric expert needed to testify to a particular degree of probability: "[a] medical expert's testimony need not be couched in the magic words 'reasonable degree of medical certainty or a reasonable probability,'" quoting *Shahan v. Hilker*, 488 N.W.2d 577, 580 (1992)); *Matter of Foster*, 426 N.W.2d 374, 378-79 (Iowa 1988) (interpreting the overt act requirement in Iowa's statute, and discussing the probability standard as a balancing of individual and societal interests); *People v. Stevens*, 761 P.2d 768, 774 (Colo. 1988) (reviewing other states' definitions of danger, and holding that due process requires that the "likelihood" of danger mean "a reasonable basis to believe that the individual's mental illness results in a present danger").

250. As keywords, statutes without definitions focus on "harm," ALA. CODE § 22-52-10.4 (1975), ALASKA STAT. § 47.30.735 (2002), IND. CODE ANN. § 12-7-2-53 (West 2001), MD. CODE ANN., HEALTH-

ily”) harm.²⁵¹ No state includes prospective threats, but many specify past threats as evidence of future harm.²⁵² Only a handful of states include neglect of others under the person’s care as a form of danger.²⁵³ In many states, sexual violence towards identified groups not only qualifies as harm, but often justifies an entirely separate commitment process.²⁵⁴ A small minority of states includes the prospect of psychic or emotional harm to others in their definition of danger.²⁵⁵ A comparable minority include danger to the property of others.²⁵⁶

“Danger to self” appears in all states as a form of danger. As with danger to others, this includes potential physical harm to the proposed patient, phrased either as a general standard²⁵⁷ or as more specific acts, typically suicide or self-mutilation.²⁵⁸ States usually deal with threats to

GEN. I § 10-632 (2000), MO. ANN. STAT. § 632.350 (West 2000), NEB. REV. STAT. § 83-1009 (1999), NEV. REV. STAT. ANN. 433A.310 (2000), 50 PA. CONS. STAT. § 7301 (West 2001), S.C. CODE ANN. § 44-17-580 (1976); VT. STAT. ANN. tit. 18, § 7101 (2000); “danger,” CAL. WELF. & INST. CODE § 5256.6 (West 1998), COLO. REV. STAT. ANN. § 27-10-109 (West 2002), N.H. REV. STATE. ANN. § 135-C:34, OR. REV. STAT. § 426.005 (1993) (“is dangerous”), VA. CODE ANN. § 37.1-67.3 (1959 Supp. 2002); or “injury,” MONT. CODE ANN. § 53-21-126 (2001).

251. “Physical”: ARIZ. REV. STAT. § 36-501 (2003), CONN. GEN. STAT. ANN. § 17a-495 (West 1992), GA. CODE ANN. § 37-3-1, HAW. REV. STAT. § 334-1 (1993), IDAHO CODE § 66-317 (2000), 405 ILL. COMP. STAT. ANN. § 5/1-119 (1997), IOWA CODE ANN. § 229.1 (2000), KAN. STAT. ANN. § 59-2946 (2002), KY. REV. STAT. ANN. § 202A.011 (1999), LA. REV. STAT. ANN. § 28:2 (West 2001), ME. REV. STAT. ANN. tit. 34-B § 3801 (West 1964), MASS. GEN. LAWS ANN. ch. 123 § 1 (West 1986), MICH. COMP. LAWS ANN. § 330.1401 (West 1999), MINN. STAT. § 253B.02 (West 2003), MISS. CODE ANN. § 41-21-61 (1999), N.Y. MENTAL HYG. § 9.01 (2002), OHIO REV. CODE ANN. § 5122.01 (West 2001), R.I. GEN. LAWS § 40.1-5-2 (1997), S.D. CODIFIED LAWS § 27A-1-1 (1999), UTAH CODE ANN. § 62A-12-234 (2000), WASH. REV. CODE § 71.05.020 (West 2002), W. VA. CODE § 27-1-12 (2001), WIS. STAT. § 51.20 (1997), WYO. STAT. ANN. § 25-10-101 (2001). “Bodily”: ARK. CODE ANN. § 20-47-207 (1997), FLA. STAT. ANN. ch. 394.467 (West 2002), N.M. STAT. ANN. § 43-1-3 (Michie 1978), TENN. CODE ANN. § 33-6-501 (2001). *Cf.* D.C. CODE ANN. § 21-545 (2001) (“injure himself or others”).

252. ARK. CODE ANN. § 20-47-207 (1997); CAL. WELF. & INST. CODE § 5300 (1998); MINN. STAT. § 253B.02 (West 2003).

253. KY. REV. STAT. ANN. § 202A.01 (1999); MICH. COMP. LAWS ANN. § 330.1401 (1999); OHIO REV. CODE ANN. § 5122.01 (West 2001); W. VA. CODE ANN. § 27-1-12 (2001).

254. *See supra* note 20 (discussing sexually violent offender commitment statutes) and Part II.B *infra* (assessing cases which consider the admissibility of predictive expertise in such statutes).

255. HAW. REV. STAT. § 334-1 (1993) (“substantial . . . emotional injury”); IOWA CODE ANN. § 229.1 (2000) (“serious emotional injury on members of the person’s family or others who lack reasonable opportunity to avoid contact with the person”).

256. DEL. CODE ANN. tit. 16, § 5001 (1995) (“harm . . . to property”); KAN. STAT. ANN. § 59-2946 (2002) (“substantial damage to another’s property”); N.D. CENT. CODE § 25-03.1-02 (2002) (“inflicting significant property damage”).

257. *See supra* note 234.

258. N.M. STAT. ANN. § 43-1-3 (Michie 1978) (“suicide or . . . serious bodily harm to himself by violent or other self-destructive means”); N.C. GEN. STAT. § 122C-3 (2001) (“suicide” and “self-mutilation”); N.D. CENT. CODE § 25-03.1-02 (2002) (“suicide” but not physical harm to self); 50 PA. ANN. STAT. tit. § 7301 (2001) (“suicide” and “mutilation”); S.D. CODIFIED LAWS § 27A-1-1 (1999) (“inflict serious physical injury upon himself”).

self in the same way as threats to others.²⁵⁹ Only one state permits civil commitment based on psychic harm to self.²⁶⁰

Neglect of oneself (as opposed to active violence) supports commitment in most states. The traditional approach includes neglect as a type of “danger to self.”²⁶¹ An increasing number of states permit commitment on a finding that the person is “gravely disabled.”²⁶² The content of this term of art varies from state to state. Its central meaning references a handful of different sorts of risks, including “inability or failure to provide . . . the essential human needs of food, clothing, shelter, and medical care.”²⁶³ Second, “gravely disabled” references a lack of judgment resulting from mental illness which in turn can expose the defendant to unwarranted, severe risks of harm from others—essentially a passive recklessness towards identifiable dangers posed by other people.²⁶⁴ Third, the term can reference loss of the mental functioning necessary for accomplishment of routine daily tasks.²⁶⁵

259. See *supra* note 252 and accompanying text.

260. See LA. REV. STAT. ANN. § 28:2 (2001) (“inflict . . . severe emotional harm upon his own person”). Some states include the prospect of mental deterioration, ME. REV. STAT. ANN. tit. 34-B § 3801 (West 1964) (“severe . . . mental impairment or injury . . . to the person”); N.D. CENT. CODE § 25-03.1-02 (2002) (“substantial deterioration in mental health which would predictably result in dangerousness”). See also *infra* notes 262–65 (discussing statutes allowing proof that the defendant is “gravely disabled”).

261. ARK. CODE ANN. § 20-47-207 (Michie 2001); FLA. STAT. ANN. § 394.467 (West 2002); GA. CODE ANN. § 37-3-1 (1995); 405 ILL. COMP. STAT. ANN. 5/1-119 (West 1996); IOWA CODE ANN. § 229.1 (West 2000); KAN. STAT. ANN. § 59-2946; ME. REV. STAT. ANN. tit. 34-B, § 3801; MASS. GEN. LAWS ANN. ch. 123 § 1 (West 1986); MICH. COMP. LAWS ANN. § 330.1401 (West 1999); MINN. STAT. ANN. § 253B.02 (West 2003); MISS. CODE ANN. § 41-21-61 (West 1999); MONT. CODE ANN. § 53-21-126 (2001); NEB. REV. STAT. § 83-1009 (1999); N.M. STAT. ANN. § 43-1-3; N.C. GEN. STAT. § 122C-3; N.D. CENT. CODE § 25-03.1-02; OHIO REV. CODE ANN. § 5122.01 (West 2001); OKLA. STAT. ANN. tit. 43A § 1-103 (West 2001); OR. REV. STAT. ANN. § 426.005 (Michie 1995); S.D. CODIFIED LAWS § 27A-1-1; TENN. CODE ANN. § 33-6-501 (2001); TEX. HEALTH & SAFETY CODE ANN. § 574.034 (Vernon’s 2003); VT. STAT. ANN. tit. 18 § 7101 (2000); W. VA. CODE ANN. § 27-1-12 (Michie 2001); WIS. STAT. ANN. § 51.20 (1997); WYO. STAT. ANN. § 25-10-101 (Michie 2001).

262. ALASKA STAT. § 47.30.735 (Michie 2002); ALASKA STAT. § 47.30.755 (Michie 2002); CAL. WELF. & INST. CODE § 5150 (West 1998); COLO. REV. STAT. ANN. § 27-10-109 (West 2002); CONN. GEN. STAT. ANN. § 17a-495 (West 1992); HAW. REV. STAT. § 334-60.2 (1993); IDAHO CODE § 66-329 (Michie Supp. 2003); IND. CODE § 12-26 (West 2002); LA. REV. STAT. ANN. § 28:54 (West 2001); N.M. STAT. ANN. § 43-1-3 (Michie 1975) (“grave passive neglect”); WASH. REV. CODE ANN. § 71.05.040 (West 2002).

263. COLO. REV. STAT. ANN. § 27-10-102(5)(A)(I) (West 2002). See also CAL. WELF. & INST. CODE § 5008(h)(1)(A); CONN. GEN. STAT. ANN. § 17a-495 (West 1992); HAW. REV. STAT. § 334-1 (1993); IDAHO CODE § 66-317 (Michie Supp. 2003); IND. CODE ANN. § 12-7-2-96 (West 2002); LA. REV. STAT. ANN. § 28:2.

264. COLO. REV. STAT. ANN. § 27-10-102(5)(A)(II) (West 2002) (“[I]acks judgment in the management of his resources and in the conduct of his social relations to the extent that his health or safety is significantly endangered and lacks the capacity to understand that this is so”).

265. See, e.g., WASH. REV. CODE § 71.05.020(14) (West 2002) (“manifests severe deterioration in routine functioning evidenced by repeated and escalating loss of cognitive or volitional control over his or her actions and is not receiving such care as is essential for his or her health or safety”).

The standard formulation of “danger to self and others” omits mention of any specific degree of severity; where present, such requirements take two forms. The statute may allow evidence of conduct that matches recognized felony-level crimes of physical violence: for example, murder, battery, or sexual assault.²⁶⁶ More commonly, the statute uses qualifying adjectives or adverbs, which require that the threatened harm be “serious”²⁶⁷ or “substantial.”²⁶⁸

d. *Assessment of Definitions*

The composite picture that emerges from this review displays distinctive characteristics of fact-finding about future dangerousness.²⁶⁹ The essentials involve a determination by the fact-finder that the respondent’s mental illness causes a risk of harm to self or others. Thus, mental illness, causation, risk (or probability), and harm form the central proofs. Variations include the degree or probability of the risk, the strength of its short-term likelihood, and the nature and severity of the harm. Most terms are decidedly indeterminate: they speak, for example, of “substantial probability” or “severe harm.” To be sure, the different types of harm seem more well-defined; here, the statutes permit overlapping proofs of harm, so that if evidence fails on one, the petitioner may still prevail on another. Fact-finding on danger thus entails simultaneous assessments of multiple possible harms, using indeterminate rhetorical standards to reach a decision on the size of the risk, the potential for harm, and the severity of harm involved.

This composite picture prompts useful insights. First, the indeterminacy of the critical elements allocates a zone of discretion to the fact-finder on the desired conclusion. As the courts have often noted, this zone responds to an adversarial balancing of interests: the state’s desire to protect and to treat the individual against the individual’s interest in remaining at liberty. In assessing probability, imminence, and severity, and in selecting among harms, the fact-finder not only can, but must, bal-

266. See, e.g., TENN. CODE ANN. § 33-6-501 (“the person has threatened or attempted homicide or other violent behavior”); WYO. STAT. ANN. § 25-10-101 (“physical harm to other individuals as manifested by a recent overt homicidal act, attempt or threat or other violent act, attempt or threat”).

267. ARIZ. REV. STAT. ANN. § 36-501 (West 2003); ARK. CODE ANN. § 20-47-207 (West 2003); DEL. CODE ANN. tit. 16, § 5001 (1995); FLA. STAT. ANN. § 394.467 (West 2002); 405 ILL. COMP. STAT. ANN. 5/1-119 (West 1996); MICH. COMP. LAWS ANN. § 330.1401 (West 1999); MO. ANN. STAT. § 632.350 (West 2000); NEB. REV. STAT. § 83-1009 (1999); N.H. REV. STAT. ANN. § 135-C:34 (1996); N.J. STAT. ANN. § 30:4-27.2 (West 1981); OKLA. STAT. ANN. tit. 43A, § 1-103 (West 2001); S.C. CODE ANN. § 44-17-580 (Law Co-op 1985); S.D. CODIFIED LAWS § 27A-1-1 (Lexis 1999); TENN. CODE ANN. § 33-6-501; TEX. HEALTH & SAFETY CODE ANN. § 574.034 (Vernon 2003); W. VA. CODE ANN. § 27-5-4 (Michie 2001).

268. ALA. CODE § 22-52-10.4 (West 2003); CAL. WELF. & INST. CODE § 5300 (West 1998); HAW. REV. STAT. § 334-1; KAN. STAT. ANN. § 59-2946 (Supp. 2002); KY. REV. STAT. ANN. § 202A.011 (West 2003).

269. See also Grant H. Morris, *Defining Dangerousness: Risking a Dangerous Definition*, 10 J. CONTEMP. LEGAL ISSUES 61, 67-71 (1999) (describing different statutory methods of defining danger).

ance these competing interests. The legal definitions thus do not structure the decision primarily as a search for accuracy in prediction, much less for statistical quantification of risk. Rather, they seek to justify judicial action: either psychically and physically coercive compassion; or potentially harmful freedom.

Second, the statutory definitions seem to suggest a method for fact-finding on danger that parallels the methods of clinical assessors. In most states, explicit language requires proof by historical facts and expert testimony; no state adopts a statistical or actuarial methodology for finding danger. Fact-finders must thus infer danger from the prior actions (or inactions) of the respondent, coupled with interpretive testimony from an expert. Even without an "other act" requirement, the rules of sufficiency reviewed earlier typically require some showing of conduct outside of expert prediction. Such an approach expresses traditional evidentiary values. Courts make inferences from what has already happened (and from the best available skilled insight) about the significance of those events for the future.

Third, legal decision-making on dangerousness encourages a clinical method; it is itself clinical decision-making. It focuses on the individual, and seeks to identify the unique features of that individual that might justify intervention. It does not refer to groups with which the individual may share characteristics, nor to what members of that group might do or be shown statistically to have done. It entails complex decision-making, with uncertain rhetorical standards applied to overlapping points of factual focus. It assumes inaccuracy and unreliability, but requires a strong degree of moral conviction that the decision about danger justifies the choice of severe intervention or risky inaction.

Fact-finding on danger thus accepts the discretion to balance competing interests, the reliance on both direct observation and specialized opinion, and the imperative of decision-making to justify coercive action against an individual. Clinical and actuarial prediction relate to this fact-finding in different ways, and neither fits perfectly: the scientific goal of statistically accurate prediction diverges widely from the judicial mandate of balanced and restrained use of state power.

5. *Conclusion: Predictive Testimony in Civil Commitments*

But *Daubert* does not require a perfect fit. This section on commitment law suggests how courts might assess the variability in *Daubert's* fit requirement. First, the fit inquiry should ask how far courts have integrated concerns over the quality of given expertise into the structure and content of the legal process in question. This integration might consist of matters as fundamental as the constitutional adequacy of evidence or process; as technical as the allocation of the risk of error through burdens

of proof; and as pragmatic as the case law definition of facts deemed sufficient to support the required legal finding. With predictions of danger, courts have shaped the commitment process to some extent because of the unreliability of expert opinion.

Second, we can infer that "fit" means in part the degree to which the particular expertise supplements and constrains the uncertainty inherent in the relevant fact-finding. Prediction of future behavior is inherently uncertain. Predictive expertise may not be all that reliable; but when it comes to prediction, nothing else is better. This inference relies on both evidentiary necessity and on cool assessment of alternatives: information which, despite its flaws, offers a useful improvement on unaided judicial fact-finding.

Third, we can infer that fit requires assessing how prevalent both the expert witness and the given opinion are within the process. Mental health professionals and their views on danger dominate the commitment process from initiation to treatment to termination of the order. This fact assures that trial judges will be both familiar with the experts and their methodologies, and will have regularly heard argument from advocates seeking to admit (or to attack) their conclusions.

Finally, we can infer that fit requires comparing the inferential process embodied in the expertise to the inferential processes embedded in the fact-finding. Such a comparison allows an intriguing conclusion for predictive expertise: that courts would be more likely to find common ground with clinical than with actuarial methods, both as a matter of methodology and as a matter of decision. Taken to its logical extreme, this point might mean that courts would have greater difficulty admitting actuarial assessments into evidence.

As we shall see, however, appellate courts have never excluded either form of opinion. This makes sense, especially if we see reliability and fit existing in a dynamic relationship under *Daubert*: the questionable reliability of clinical judgments would be offset by their extraordinarily good fit; while the greater reliability of actuarial assessments would overcome a lesser fit with the demands of individualized fact-finding. In effect, the balancing of reliability and fit should work differently for different kinds of expertise, even within the same case.

II. *DAUBERT* (AND *FRYE*) AND DANGER: ADMISSIBILITY OF EXPERT PREDICTIONS

Appellate courts uniformly admit expert psychiatric predictions in civil commitments, whether based on clinical prediction or actuarial as-

sessments.²⁷⁰ Prior to *Daubert*, few cases discussed the question; since *Daubert*, every appellate court to have reviewed the question has admitted expert predictions. This unanimous and unflinching acceptance of these opinions by courts is surprising. How can this expertise, so thoroughly questioned and cautiously advanced, receive no appellate disapproval, in a *Daubert* regime focused on scientific reliability? In those states that have adhered to *Frye*, how can this expertise, so roundly critiqued by its own practitioners, qualify under a “general acceptance” standard? *Daubert* allows more consistent answers to these questions than does *Frye*, but only if we understand it as a test that looks to something more than mere scientific validity.

This Part discusses the admissibility question in detail. Section A reviews the case law, assessing how predictive expertise has fared in those cases that openly address admissibility. Despite the uniform results, the rationales vary widely. In particular, while all courts applying *Daubert* have conceived of it as a “scientific reliability” test, each of these courts has had to stretch this standard to apply it to predictive testimony. *Frye* courts have faced an even more daunting task; they have had to explain either how predictive testimony satisfies the general acceptance standard, or why it should receive no special testing at all. Section B offers a sounder approach to the acceptance of predictive expertise, combining the notions of reliability and fit. This approach acknowledges the strong differences between the clinical and actuarial approaches, and explains how *Daubert* offers a doctrinally sound method for admitting the expert testimony that each produces.

Before this, one short tangent is in order: How does the treatment of predictive expertise compare to the evidentiary handling of other forms of psychiatric, behavioral, or other “soft” expertise? Both before and after *Daubert*, this kind of expertise has come before the courts in many different forms.²⁷¹ *Daubert* left open whether its test applied to expertise

270. One author has suggested that a series of recent trial court decisions indicates a trend towards excluding these opinions. See Donna Cropp Bechman, *Sex Offender Civil Commitments: Scientists or Psychics?*, 16 CRIM. JUST. 24, 31–32 (2001) (identifying trial court decisions in Iowa, Arizona, Florida and Missouri which have excluded predictive opinion in sex offender commitments.) However, in the first two states, appellate opinion has more recently indicated a firm acceptance of such opinions; we discuss each below. Neither Florida nor Missouri have as yet seen an appellate resolution of the issue.

271. In addition to predictive testimony, a non-exclusive list of the uses of psychiatric or psychological expertise includes the existence of a syndrome or a diagnostic category in explanation of a particular party's behavior, see *Henson v. State*, 535 N.E.2d 1189, 1192 (Ind. 1989) (rape trauma syndrome); *Commonwealth v. Dunkle*, 602 A.2d 830, 833–34 (Pa. 1992) (child sexual abuse syndrome); *Fowler v. State*, 958 S.W.2d 853, 862–64 (Tex. App. 1997) (consistency with behaviors of a domestic violence victim); *Frenzel v. State*, 849 P.2d 741, 743 (Wyo. 1993) (child sexual abuse accommodation syndrome); the presence or absence of a mental state, proof of which was required by relevant law, such as insanity or specific intent, see *Douglas v. United States*, 386 A.2d 289, 296 (D.C. 1978)

in the behavioral or psychic sciences, or indeed, to any opinions based on non-scientific expertise.²⁷² *Kumho Tire* resolved this by applying *Daubert* to all expertise.²⁷³ However, many jurisdictions decline to follow *Daubert*, and continue to apply *Frye*. In effect, the behavioral, psychic, and soft sciences now receive review under three different approaches: application of *Daubert*; application of *Frye*; or exemption from any special testing.

Predictive expertise has been subject to these same three approaches; what differs are the results. Judicial opinion, split on virtually every other form of behavioral or psychic expertise, has so far unanimously accepted predictive expertise in civil commitments.²⁷⁴ This con-

(mental capacity); *Cecil v. Commonwealth*, 888 S.W.2d 669, 674-75 (Ky. 194) (addressing expert testimony on defendant's insanity); *In re Estate of Dokken*, 604 N.W.2d 487, 494 (S.D. 2000) (testamentary capacity); the uncertainty and unreliability associated with witness's capacity for accurate testimony, see *McCleery v. City of Bakersfield*, 216 Cal. Rptr. 852, 856-58 (Cal. Ct. App. 1985) (addressing expertise on the reliability of eyewitness testimony); *Campbell v. State*, 814 P.2d 1, 6-7 (Colo. 1991) (same); *Jordan v. State*, 928 S.W.2d 550, 552-53 (Tex. Ct. App. 1996) (same); the accuracy and persistence of memory, especially of traumatic events, see *Isely v. Capuchin Province*, 877 F. Supp. 1055, 1056 (E.D. Mich. 1995) (same); *Logerquist v. McVey*, 1 P.3d 113, 115 (Ariz. 2000) (assessing admissibility of expert testimony on repressed memory); *Wilson v. Phillips*, 86 Cal. Rptr. 2d 204, 206-08 (Cal. Ct. App. 1999) (same); and the usefulness of hypnosis, see *Commonwealth v. Nazarovitch*, 436 A.2d 170, 171 (Pa. 1981) (hypnotically refreshed testimony).

It is well beyond the scope of this Article to offer a complete analysis of all of these cases.

272. *Compare Compton v. Subaru of Am., Inc.*, 82 F.3d 1513, 1518-19 (10th Cir. 1996) (*Daubert* limited to novel scientific evidence or the hard physical sciences), *cert. denied*, 519 U.S. 1042 (1996), with *Moore v. Ashland Chem., Inc.*, 126 F.3d 679, 685-88 (5th Cir. 1997) (*Daubert* applies to clinical medicine, but specific factors discussed in *Daubert* do not apply), and *Watkins v. Telsmith, Inc.*, 121 F.3d 984, 990-91 (5th Cir. 1997) (applying *Daubert* to expert testimony in engineering), and *Cook v. Am. S.S. Co.*, 53 F.3d 733, 738 (6th Cir. 1995) (applying *Daubert* to technical and other specialized knowledge), and *Tyus v. Urban Search Management*, 102 F.3d 256, 263 (7th Cir. 1996) (*Daubert* applies to social science experts, in this case sociologists and statisticians), and *United States v. Cordoba*, 104 F.3d 225, 230 (9th Cir. 1997) (*Daubert* doesn't apply to expert testimony regarding the modus operandi of drug traffickers because not based on scientific knowledge), and *United States v. Webb*, 115 F.3d 711, 716 (9th Cir. 1997) (*Daubert* doesn't apply to expert testimony as to why people typically hide guns in the engine compartments of their cars), and *Coleman v. Exxon Chem. Corp.*, 162 F. Supp. 2d 593, 617 (S.D. Tex. 2001) (affirming exclusion of expert statistical evidence, but applying *Daubert*), and *Ohio ex rel. Montgomery v. Louis Trauth Dairy, Inc.*, 925 F. Supp. 1247, 1252-53 (S.D. Ohio 1996) (applying *Daubert* to expert statistical testimony).

273. *Kumho Tire Co. v. Carmichael*, 526 U.S. 137, 147-49 (1999); see also *Elcock v. Kmart Corp.*, 233 F.3d 734, 744 (3d Cir. 2000) ("*Kumho Tire* . . . extended the rigorous gatekeeping function assigned to trial judges by *Daubert* . . . to cases involving non-scientific testimony."); See also FED. R. EVID. 702 advisory committee's note to 2000 Amendment ("[T]he Court in *Kumho* clarified that this gatekeeper function applies to all expert testimony, not just testimony based in science.").

274. Only an occasional quixotic, non-binding opinion has questioned this result. In *Fernandez v. State*, 564 S.W.2d 771, 773, 776 (Tex. Crim. App. 1978), the majority upheld the exclusion of polygraph evidence; the dissent critiqued this exclusion by contrasting it with the admission of predictive testimony, despite its flaws. In *Inmates of Allegheny County Jail v. Wecht*, 1996 WL 474106, at *1 (3d Cir. Aug. 22, 1996), *opinion withdrawn and case set for en banc hearing*, 93 F.3d 1146 (3d Cir. 1996), a majority of a the three judge panel determined that prisons fell within relevant statutes preventing dis-

sensus confirms that predictive testimony is an outlier, even among cognate areas of specialized knowledge. Given this state of affairs, something more than straight reliability must explain the courts' willing use of the expertise.

A. ADMITTING PREDICTIONS OF DANGER

1. *Daubert and Scientific Reliability*

Only a few states have used *Daubert* to assess the admissibility of predictive testimony. The most extensive treatment has occurred in Texas, primarily in death penalty cases. Texas originally followed *Frye*, but in the early 1990s rejected that standard in criminal cases, and after *Daubert*, aligned both civil and criminal evidence standards with the *Daubert* approach.²⁷⁵ Texas also has a long history of appellate litigation over the use of predictive testimony in the penalty phase of capital trials,²⁷⁶ including in cases that involved testimony by "Dr. Death," an ex-

ability-related discrimination, and remanded the case for a determination whether violence-prone offenders could participate in community-based rehab programs. *Id.* at *19.

In partial dissent, Judge Becker opposed remand for determination of the claimants' likelihood of danger, on the grounds that predictive science could produce no reliable answers, and thus that remand served no useful purpose: "In my view, this is a meaningless exercise." *Id.* at *19.

Finally, in *Flores v. Johnson*, 210 F.3d 456 (5th Cir. 2000), the court, acting per curiam, affirmed the rejection of a habeas petition from a Texas death penalty sentence, as an impermissible collateral attack. In a special concurrence, Judge Garza reviewed predictive testimony using the *Daubert* standard and opined that *Daubert* should lead to the exclusion of such testimony, and an overturning of *Barefoot*. *Id.* at 462-66. See *infra* text accompanying note 297 (discussing Judge Garza's argument about the reliability standard of *Daubert*).

275. *Kelly v. State*, 824 S.W.2d 568, 573 (Tex. Crim. App. 1992) (overturning *Frye* and requiring the proponent of "novel scientific evidence" to establish "(a) the underlying scientific theory must be valid; (b) the technique applying the theory must be valid; and (c) the technique must have been properly applied on the occasion in question." Note that sections (a) and (b) of the quoted standard focus on the validity of the scientific theory and its application, while section (c) focuses on its application to the dispute, arguably a relevance standard. The same court later held that the *Kelly* standard applied to all forms of expert testimony, novel or not. *Hartman v. State*, 946 S.W.2d 60, 62-63 (Tex. Crim. App. 1997).

See also *E.I. du Pont de Nemours & Co., Inc. v. Robinson*, 923 S.W.2d 549, 556 (Tex. 1995) (confirming *Kelly*, approving *Daubert*, and requiring a proponent "to show that the expert's testimony is relevant to the issues in the case and is based upon a reliable foundation"). The court reviewed and approved the *Daubert* factors, adding to this non-exclusive list "(2) the extent to which the technique relies upon the subjective interpretation of the expert; . . . [and] . . . (6) the non-judicial uses which have been made of the theory or technique." *Id.* at 557 (citations omitted).

276. See *Barefoot v. Estelle*, 463 U.S. 880, 896 (1983) (affirming the use of psychiatric predictions of danger in capital penalty proceedings, despite their unreliability, against a due process challenge); *Estelle v. Smith*, 451 U.S. 454, 469 (1981) (requiring *Miranda* warnings and notice to counsel in advance of any psychiatric examination which might produce testimony about future dangerousness in death penalty litigation). Several cases assess the sufficiency of expert predictions for a given death sentence. See *Chambers v. State*, 568 S.W.2d 313, 324 (Tex. Crim. App. 1978) (sufficient); *Crawford v. State*, 617 S.W.2d 925, 933 (Tex. Crim. App. 1980) (no expert testimony, still sufficient); *Cantu v. State*, 842 S.W.2d 667, 674-75 (Tex. Crim. App. 1992) (sufficient); *Fuller v. State*, 827 S.W.2d 919, 935

pert with a long-standing habit of predicting future dangerousness with close to absolute certainty.²⁷⁷

In early cases, the Texas courts held predictive expertise admissible without any special testing.²⁷⁸ By the early 1990s, Texas courts had begun to discuss the validity of psychiatric predictions and continued to admit them for a variety of reasons.²⁷⁹ Finally, in *Nenno v. State*, the Texas Court of Criminal Appeals reviewed the admissibility of predictive testimony under its version of *Daubert*.²⁸⁰ The court initially saw the question as whether these standards applied to “nonscientific expert testimony (i.e. that involving technical or other specialized knowledge),” and answered with a “qualified yes.”²⁸¹ It stressed the flexibility and complexity

(Tex. Crim. App. 1992) (sufficient). *Contra* *Cook v. State*, 858 S.W.2d 467, 474–75 (Tex. Crim. App. 1993) (rejecting attacks on the expert’s knowledge and on the use of hypotheticals).

277. See *Barefoot*, 463 U.S. at 899; see also *Fuller v. Johnson*, 114 F.3d 491, 497 (5th Cir. 1997) (rejecting an attack on the testimony of Dr. James P. Grigson). The Fifth Circuit noted that Dr. Grigson had been labeled “Dr. Death” by the media, and had received a series of unfavorable articles on his career. It also noted that “[t]he American Psychiatric Association has reprimanded him twice for his testimony, and it has filed an amicus curiae brief with the Supreme Court urging the Court to prohibit his predictions because the association finds the predictions unreliable.” *Id.* at 497. Defendant had claimed that because of this coverage, including academic criticism of the doctor’s past conclusions, the doctor had lied in predicting danger with absolute certainty. The Court held that such an argument went to weight, and not to admissibility. *Id.* See also *U.S. v. Rojas*, 15 M.J. 902, 923–24 (NMCMR 1983) (affirming admission of Dr. Grigson’s testimony, and noting his response to a question about the A.P.A.’s position in opposition to the forensic use of predictive opinions, that “he disagreed with this opinion which he felt was held by a small percentage of the APA membership”); *Gardner v. Johnson*, 247 F.3d 551, 556 n.6 (5th Cir. 2001) (noting that Dr. Grigson’s sobriquets included “the hanging psychiatrist”; describing his absolute certainty about the future violent behavior of Randall Dale Adams, a man later found to have been convicted by the use of falsified evidence; and describing his expulsion from membership of the American Psychiatric Association and the Texas Society of Psychiatric Physicians “for arriving at a psychiatric diagnosis without examining the individuals in question and for indicating, while testifying as an expert witness, that he could predict with 100 percent certainty that the individuals would engage in future violent acts”). *Contra* *Chambers v. State*, 568 S.W.2d 313, 324 (Tex. Crim. App. 1978) (holding Dr. Grigson’s testimony sufficient without special challenge); *Fuller v. State*, 829 S.W.2d 191, 195 (Tex. Crim. App. 1992) (citing *Barefoot* to reject a challenge to Dr. Grigson’s testimony that “absolutely there is no question, no doubt, whatsoever, that [Appellant] . . . will commit future acts of violence”); *Cook v. State*, 858 S.W.2d 467, 475 (Tex. Crim. App. 1993) (rejecting an attack on Dr. Grigson’s knowledge of the defendant’s background); *Carter v. State*, 851 S.W.2d 390, 391–94 (Tex. App. 1993) (holding the doctor’s testimony admissible against a general admissibility challenge).

278. See *Fuller v. State*, 829 S.W.2d at 195 (Tex. Crim. App. 1992) (noting earlier cases that held such testimony admissible if the trial court determined them relevant, followed by a line of cases which, in addition, required a showing of the expert’s qualifications).

279. *Id.* (citing *Chambers v. State*, 568 S.W.2d 313, 324 (Tex. Crim. App. 1978)). See also *Carter v. State*, 851 S.W.2d at 393 (holding predictive testimony admissible if it helps the jury determine a fact in issue and is otherwise relevant; citing *Duckett v. State*, 797 S.W.2d 906, 914 (Tex. Crim. App. 1990); *Massey v. State*, 933 S.W.2d 141, 157 (Tex. Crim. App. 1996) (holding predictive testimony admissible where the expert had strong professional credentials)).

280. *Nenno v. State*, 970 S.W.2d 549, 560–62 (Tex. Crim. App. 1998) (emphasis omitted).

281. *Id.* at 560.

of the trial court's gatekeeping function, especially as to evidence built on the soft sciences, including "the social sciences or fields that are based primarily upon experience and training as opposed to the scientific method."²⁸² The court held that the requirement of reliability applies with less rigor to soft rather than to hard sciences. It articulated alternate criteria for assessing reliability: "(1) whether the field of expertise is a legitimate one, (2) whether the subject matter of the expert's testimony is within the scope of that field, and (3) whether the expert's testimony properly relies upon and/or utilizes the principles involved in the field."²⁸³ On the strength of these criteria, the court held admissible clinical predictions by a doctor with extensive experience studying the risk of future sexual victimization of children.²⁸⁴ Notably, *Nenno* addressed predictive testimony in a death penalty case; no Texas case has assessed predictive testimony in civil commitments. The appellate cases have simply assumed such evidence to be admissible in civil commitments without discussion.²⁸⁵

An appellate court in Iowa, however, has used *Daubert*²⁸⁶ to assess predictive testimony in a civil commitment. The court first issued a panel

282. *Id.* at 561.

283. *Id.* The court indicated its belief that "[t]hese questions are merely an appropriately tailored translation of the *Kelly* test to areas outside of hard science," and indicated that factors such as *Daubert's* four criteria did not necessarily apply to non-hard scientific expertise. *Id.*

284. *Id.* at 562. Specifically, it found that research into such behavior constituted a "legitimate field of expertise," that the witness' testimony was within the scope of that field, and that the witness had used methods characteristic of those practicing in this field. These methods included: "interviews, case studies, and statistical research . . . [study of] a thousand cases that concerned the issue of future dangerousness in some fashion . . . [and of] . . . solved cases to attempt to understand the dynamics of what occurred; . . . personal interviews with inmates convicted of child sex offenses, examining the inmates' psychological records, and examining the facts of the offenses involved." *Id.* The court acknowledged the lack of any peer review of this expertise, but held that went more to weight than to admissibility. *Id.*

285. *See, e.g.,* *Goldwait v. State*, 961 S.W.2d 432, 434-35 (Tex. App. 1997) (affirming a commitment order based on evidence which included testimony from two doctors who offered clinical testimony predicting danger; held, state had met its burden of clear and convincing evidence, without any discussion of the admissibility of these opinions).

286. Technically, Iowa uses a flexible approach, which includes, but is not restricted to the *Daubert* standard. In *Leaf v. Goodyear Tire & Rubber Co.*, the Iowa Supreme Court indicated that Iowa trial courts did not have to use *Daubert* in assessing expert testimony. 590 N.W.2d 525, 533 (Iowa 1999). However, they *could* do so:

trial courts may find it helpful, particularly in complex cases, to use one or more of the relevant *Daubert* "considerations" in assessing the reliability of expert testimony. Therefore, trial courts may, in their discretion, consider the [four *Daubert* factors] if deemed helpful in a particular case.

Id. (citation omitted). *See also* *State v. Atwood*, 602 N.W.2d 775, 783 (Iowa 1999) (applying an abuse of discretion standard to trial court rulings on admissibility); *In re Detention of Holtz*, No. 01-0243, 2002 WL 663683 (Iowa Ct. App. Apr. 24, 2002) (on file with author) *withdrawn and superceded by In re Detention of Holtz*, 653 N.W.2d 613 (Iowa Ct. App. 2002) (en banc).

opinion in April 2002 that excluded opinion evidence based on actuarial assessments. However, in September 2002, the court (through the same judge) issued an en banc decision that admitted the same evidence.²⁸⁷ The April opinion used *Daubert* to assess testimony based on various actuarial assessment tools; none of them satisfied this *Daubert* screening.²⁸⁸ The defendant had offered rebuttal witnesses whose only purpose was to address the weaknesses of the various actuarial methods under a *Daubert* standard.²⁸⁹ The court concurred with this expert: “[T]here was no foundation to show proper testing of these instruments, peer review or publication and the evidence was insufficient to show acceptance of these actuarial risk assessment instruments in the scientific community.”²⁹⁰ The opinion included cites recent literature stressing the limitations on these instruments.²⁹¹

287. *In re Detention of Holtz*, 653 N.W.2d 613, 619–20 (Iowa App. 2002).

288. *Holtz*, 2002 WL 663683, at *2, *5 (superceded opinion; on file with author) (assessing the RRASOR, the Static-99, the Minnesota Sex Offender Screening Tools (MnSOST) and the Minnesota Sex Offender Screening Tool—Revised (MnSOST-R)). Neither the parties nor the court assessed how these actuarial tools might combine with clinical opinions on danger. See *supra* Part II.B.1.b (identifying various different actuarial scoring instruments for assessing risk).

289. *Holtz*, 2002 WL 663683, at *3–*4 (on file with author). One witness testified that:

[t]he reliability is unknown and has not been established in scientific review, and I think that their reliability is uncertain, first of all, because it hasn't gone through that process, and, secondly, because the task of those tests is very difficult, is to try to predict future behavior, and that's obviously a very difficult thing to do. So I think those tests are new, unproven, and the reliability is unknown.

Id. at *4. Another rebuttal witness confirmed this assessment:

Now, I think common sense and certainly our professional views tell us that predicting future behavior of individuals is a bit tricky anyway because we know that's going to be a difficult thing. But when we develop tests to do that, we have to do so in a way that ensures that they have some reasonable degree of accuracy in doing so. And none of these tests has been subjected to any kind of evaluation that would allow you to say with a degree of certainty that they can predict somebody's future behavior. They're simply too new.

Id. This witness noted not only the lack of any credible peer review, but also the impossibility of such a review:

They've never been published in a way that would allow you to evaluate their predictive accuracy. So not only they are not accepted, it's impossible for them to be accepted right now. It would be impossible for some years until people actually subject the research on which these things were based to some kind of independent evaluation.

Id. Finally, this witness indicated that, even if valid, the tests have limited utility.

[I]t is acceptable to use the actuarial instruments as a way of trying to determine the presence of various kinds of risk factors. . . . [W]hat is not acceptable at this time is 'adding up those numbers to get some kind of a—score that you can then change into a prediction of the future.'

Id. at *5.

290. *Id.*

291. *Id.* The cites are as follows:

Donna Cropp Bechman, *Sex Offenders Civil Commitments: Scientists or Psychics?*, 16-SUM CRIM. JUST. 24, 26–30 (2001); Grant H. Morris, *Defining Dangerousness: Risking A Dangerous Definition*, 10 J. CONTEMP. LEGAL ISSUES 61, 89–92 (1999); John Q. La Fond & Bruce J. Winick, *Foreword Sex Offenders and The Law*, 4 PSYCHOL., PUB. POL'Y, & L. 3, 6–8 (1998); Eric S. Janus & Paul E. Meehl, *Assessing the Legal Standard for Predictions of Dangerous-*

Curiously, in withdrawing this opinion, the same judge wrote a second opinion, issued en banc, that contained virtually the same review of facts, procedure and precedent.²⁹² The en banc opinion added a description of testimony from the actuarial expert, who had conceded the newness of the tests, but stressed that he had used them only as part of an overall clinical assessment.²⁹³ The en banc opinion omitted its earlier citation to hostile literature, and emphasized instead that “[o]ur research has revealed no state appellate court decision which has found actuarial instruments inadmissible at [sexually violent predator] proceedings.”²⁹⁴ The opinion then approved admission, but only as part of a thorough clinical prediction:

ness in Sex Offender Commitment Proceedings, 3 PSYCHOL., PUB. POL’Y, & L. 33 (1997). Even those who cautiously endorse the actuarial instruments acknowledge that the reliability and validity has not been sufficiently established. *See, e.g.*, R. Karl Hanson, *What Do We Know About Sex Offender Risk Assessment?*, 4 PSYCHOL., PUB. POL’Y, & L. 50, 52–53 (1998) (stating that the actuarial instruments are modestly accurate); Grant T. Harris et al., *Appraisal and Management of Risk in Sexual Aggressors: Implications For Criminal Justice Policy*, 4 PSYCHOL., PUB. POL’Y, & L. 73, 90–91 (1998) (recognizes there is a risk for false positives, individuals predicted as likely to reoffend who do not); Judith V. Becker & William D. Murphy, *What We Know and Do Not Know About Assessing and Treating Sex Offenders*, 4 PSYCHOL., PUB. POL’Y, & L. 116, 126–27 (1998) (stating there is still number of false positives which could lead to some individuals being detained who would not reoffend).

Id.

292. *In re Detention of Holtz*, 653 N.W.2d 613, 614–16 (Iowa Ct. App. 2002) (en banc).

293. *Id.* at 617, 619–20. The opinion substantially shortened its summary of the rebuttal testimony. *Id.* at 617–18.

294. *Id.* at 619 (quoting *In re Commitment of R.S.*, 773 A.2d 72, 96 (N.J. Super. Ct. App. Div. 2001) (citing *People v. Ward*, 83 Cal. Rptr. 2d 828, 832 (Cal. Ct. App. 1999) (“In civil commitment cases where the trier of fact is required by statute to determine whether a person is dangerous or likely to be dangerous, expert prediction may be the only evidence available.”); *People v. Poe*, 88 Cal. Rptr. 2d 437, 440 (Cal. Ct. App. 1999) (use of RRASOR upheld); *Garcetti v. Super. Ct.*, 102 Cal. Rptr. 2d 214, 241 (Cal. Ct. App. 2000) (use of PCL-R, RRASOR and Static 99 upheld); *In re Detention of Walker*, 731 N.E.2d 994, 998 (Ill. App. Ct. 2000) (use of RRASOR upheld); *In re Detention of Strauss*, 20 P.3d 1022, 1024 (Wash. Ct. App. 2001) (use of MnSOST, RRASOR and VRAG upheld); *In re Detention of Campbell*, 986 P.2d 771, 779 (Wash. 1999) (reliance on actuarial and clinical assessment proper and weight to be given evidence is question for the jury). *See also* *State, ex rel. Romley v. Fields*, 35 P.3d 82, 89 (Ariz. Ct. App. 2001) (“use of actuarial models by mental health experts to help predict person’s likelihood of recidivism is not the kind of novel scientific evidence or process to which Frye applies”); *Commonwealth v. Reese*, No. CIV.A 00-0181-B, 2001 WL 359954 at *9 (Mass. Super. April 5, 2001) (“[S]tatistics, in general, are better predictors of future sexual dangerousness than clinical judgments.”)). The cited cases use different standards of review, and are analyzed in this Article under their appropriate sections.

[W]e are not concluding that actuarial risk assessment instruments are reliable per se or have our approval when used alone and not in conjunction with a full clinical evaluation. We note this was not the situation or issue presented in the instant case. The instruments were used in conjunction with a full clinical evaluation and their limitations were clearly made known to the jury.²⁹⁵

The en banc opinion provides no explanation for the change.²⁹⁶

Only one member of one court, Judge Emilio Garza, has indicated a willingness to reject predictive testimony under a *Daubert* standard, albeit in a solo concurrence to a majority opinion that turned on an unrelated issue.²⁹⁷ Using the four original *Daubert* factors as guidelines, Judge Garza found: (1) no coherent testing of individual opinions or of general clinical methods has been performed; (2) no peer review of individual opinions, along with negative peer assessments of clinical methodology; (3) a minimum error rate of fifty percent; and (4) an overwhelmingly negative professional consensus on predictive reliability.²⁹⁸ The judge's argument seem tendentious; other decisions by other courts have reached more balanced and also more favorable assessments of clinical opinion. Moreover, Judge Garza found fault with a thin clinical opinion, not a richer, more textured mix of clinical and actuarial information.²⁹⁹ Nonetheless, if we assume a narrow view of *Daubert*, there is much to be said for Judge Garza's approach. If *Daubert* focuses solely on scientific

295. *Holtz*, 653 N.W.2d at 619–20 (emphasis omitted).

296. The court's en banc opinion offers some clues for the switch, but none are determinative. First, one can view the increased stress on the proponent's expert as a recognition of the abuse of discretion standard of review; given the existence of some evidence to support the trial court, reversal was not warranted. Second, the more thorough assessment of primary legal authority may have been conclusive. Third, the stress on the interplay of clinical and actuarial methods may have inspired greater confidence than actuarial methods standing alone. Finally, both the quoted testimony and the cited authority stress the lack of any feasible alternative proof to sustain a judicial prediction of danger. This Article will return to these points later. See *infra* Part II.B.

297. *Flores v. Johnson*, 210 F.3d 456, 458–70 (2000) (Garza, J., specially concurring) (affirming a Texas conviction and death sentence of a Mexican national, and rejecting claims based on ineffective assistance of counsel and violation of rights protected by the Vienna Convention). His colleagues in the per curiam opinion rejected his argument in a brief footnote focused on the lack of any constitutional issue. *Id.* at 456 n.1 (“It is the inescapable fact that a lay jury is asked to judge future dangerousness. We cannot then reject as constitutionally infirm the admission into evidence of the same judgment made by a trained psychiatrist.”). The concurring opinion reviews testimony by Dr. James Griffith, a doctor with a record of predictive testimony in Texas death cases equal to that of Dr. Grigson. This testimony is clinical, not actuarial. Indeed, one gravamen of Judge Garza's complaint focuses on the fact that the doctor never interviewed the defendant, but based his opinion on listening to testimony at trial and on hypothetical questions. *Id.* at 458.

298. *Id.* at 464–65.

299. See *In re Detention of Holtz*, 653 N.W.2d 613, 617 (Iowa Ct. App 2002) (en banc decision). See also *Commonwealth v. Reese*, No. CIV.A 00-0181-B, 2001 WL 359954 at *9 (Mass. Super. Ct. Apr. 3, 2001) (comparing clinical and actuarial methods with “guided clinical judgment,” which incorporated both methodologies).

validity, his discussion lays a strong foundation for excluding predictive expertise.³⁰⁰

Neither the Texas nor the Iowa opinions, however, rest solely or even primarily on scientific reliability. The *Nenno* opinion applies *Daubert* with less rigor to experience-based expertise, rejects the use of traditional scientific methods, and asks solely whether the relevant field of expertise is “legitimate,” a term to which the Court assigns no content.³⁰¹ The en banc opinion in *Holtz* replaces the panel’s earlier, scientifically focused rejection of predictive expertise with a mix of precedent and deference to fact-finder discretion.³⁰² Taken together, the justifications offered in these decisions rely not on science but rather on the mechanics and pragmatism of dispute resolution: the need for legitimacy; an appropriate deference to fact-finders; the reliance on legal precedent; and a trust in adversarial processes to limit the force of weak expertise. Not coincidentally, the same or similar considerations permeated my earlier review of how the use of predictive testimony has shaped civil commitment proceedings overall.³⁰³ At the very least, these opinions debunk the need for a purely scientific reading of *Daubert* and *Kumho Tire*, and confirm that courts can and will assess the fit of testimony within the broader purposes and processes of litigation, before excluding it as unreliable.

2. *Frye* and General Acceptance

At first glance, any discussion of *Frye* may seem tangential to the application of *Daubert* to predictive expertise. For at least a few reasons, however, a consideration of *Frye* promotes understanding of *Daubert*’s role in this setting. First, *Daubert* itself retains a concern for general acceptance. Moreover, more states have coped with predictive expertise in civil commitments using a *Frye* approach than a *Daubert* approach. These states have had difficulty applying *Frye*, and some have eliminated *Frye* entirely as a test for predictive expertise. Finally, the ways in which the *Frye* jurisdictions have adapted their standard confirm a central con-

300. *But see* John Monahan, *Violence Risk Assessment: Scientific Validity and Evidentiary Admissibility*, 57 WASH. & LEE L. REV. 901, 910–15 (2000) (applying the *Daubert* factors to predictive testimony, and determining that a sound argument for general admissibility exists).

301. *Nenno v. State*, 970 S.W.2d 549, 561 (Tex. Crim. App. 1998). If the field is found “legitimate,” the expert’s opinion need only fall within the scope of that field and use the principles and methodologies of that field. *Id.* The Court completely fails to offer any standard for distinguishing legitimate from illegitimate fields of inquiry.

302. *In re Detention of Holtz*, 653 N.W.2d at 619–20. Indeed, this opinion expresses the view that more scientifically based objections to predictive testimony “went to the weight the evidence should receive as opposed to the issue of admissibility.” *Id.* at 619.

303. *See supra* Part I.C (discussing the accommodation of predictive unreliability in the content, process, and structure of civil commitments).

tention: that admissibility of expertise requires as much of an assessment of its fit to a given dispute, as it does of an inquiry into its scientific validity.

Two states have explicitly applied the “general acceptance” standard to predictive testimony: Washington and New Jersey. In the leading Washington case, the proposed patient elicited testimony at trial about the lack of general acceptance in the relevant community,³⁰⁴ on appeal, the Washington State Psychiatric Association joined as *amicus curiae*, making the same argument.³⁰⁵ The court disagreed, and affirmed the trial court’s ruling that the evidence was admissible. The court articulated its long-standing standard for general acceptance, in terms of “whether the evidence being offered is based on established scientific methodology.”³⁰⁶ The court’s reasoning, however, focused on how well-established predictive expertise had become *in the courts*, rather than in the scientific community. It noted its earlier acceptance of predictions on constitutional grounds, and repeated the warning that rejecting predictions would “eviscerate” the civil commitment process.³⁰⁷ In effect, the court concluded that the strong fit between commitment cases and psychiatric/psychological testimony overcame any unreliability in the predictions:

The sciences of psychology and psychiatry are not novel; they have been an integral part of the American legal system since its inception. Although testimony relating to mental illnesses and disorders is not amenable to the types of precise and verifiable cause and effect relation petitioners seek, the level of acceptance is sufficient to merit consideration at trial.³⁰⁸

304. *In re Young*, 857 P.2d 989, 1016–18 (Wash. 1993) (affirming admission of psychiatric opinions predicting dangerousness under Washington’s version of the *Frye* standard). Over a decade earlier, Washington had ruled that use of psychiatric predictions in civil commitments did not violate due process standards. *In re Harris*, 654 P.2d 109, 111 (Wash. 1982) (rejecting a due process challenge to the use of psychiatric predictions to justify civil commitment). Anticipating *Barefoot*, the Washington Court found little in the existing law to overturn reliance on predictive expertise on constitutional law. *Id.* In its view, psychiatric predictions were central to commitments; rejection of such opinions “would eviscerate the entire law of involuntary commitment as well as render dubious the numerous other areas where psychiatry and the law intersect.” *Id.* The Court also noted the relationship between procedural protections and constitutional adequacy, noting that courts could constrain the risk of error from predictions “by requiring demonstration of a substantial risk of danger and by imposing procedural safeguards and a heavy burden of proof.” *Id.* To this end, it construed the relevant statute to require a showing of a “recent overt act” as a prerequisite to a finding of danger. *Id.* at 113. *In re Young*, 857 P.2d at 1016.

305. *Id.*

306. *Id.* (quoting *State v. Cauthron*, 846 P.2d 502, 507 (Wash. 1993)).

307. *In re Young*, 857 P.2d at 1016–17 (quoting *In re Harris*, 654 P.2d at 111). It also noted several legislative enactments which it determined implied an acceptance of psychiatric predictions.

308. *In re Young*, 857 P.2d at 1017. Later Washington cases have followed *Young*, both as to clinical predictions, *In re Detention of Campbell*, 986 P.2d 771, 779 (Wash. 1999) (affirming admission of predictive expertise in a sex offender case); *Pedersen v. State*, No. 43031-9-1, 2000 WL 426460, at *5

In New Jersey, one appellate court applied *Frye* to actuarial assessments³⁰⁹ against the back-drop of a well-developed case law on civil commitment³¹⁰ and a long-standing acceptance of the *Frye* standard.³¹¹ New Jersey courts had long applied *Frye* to testimony based on the behavioral sciences and psychiatry,³¹² holding that the proponent of expertise can prove general acceptance through expert testimony, authoritative scientific and legal writings, or judicial opinions.³¹³

The appellate division in *In re Commitment of R.S.* undertook a *de novo* review of the admissibility of expert evidence based on actuarial as-

(Wash. Ct. App. Apr. 17, 2000) (affirming admission, despite renewed claim that the methods were not generally accepted); *State v. In re Detention of Soliz*, No. 44127-2-I, 2000 WL 965007, at *12-*13 (Wash. Ct. App. July 3, 2000) (affirming admission, against a claim of unreliability for the lack of actuarial methods); and as to the use of actuarial assessments; *In re Detention of Thorell*, 2000 WL 222815, at *5-*6 (Wash. Ct. App. Feb. 22, 2000) (affirming admission of the VRAG, the RRASOR, and the SORAG); *In re Detention of Strauss*, 20 P.3d 1022, 1025 (Wash. Ct. App. 2001) (affirming admission of actuarial methods). A few of these cases held that questions about the "general acceptance" of the inferential methods were a matter of weight and not admissibility, and could be accurately assessed by the finder of fact. See *In re Detention of Strauss*, 20 P.3d at 1026; *Pedersen*, 2000 WL 426460, at *5; *Soliz*, 2000 WL 965007, at *12-*13.

309. *In re Commitment of R.S.*, 773 A.2d 72, 89-95 (N.J. Super. Ct. App. Div. 2001) (affirming admission of various actuarial assessment instruments). No New Jersey case has explicitly discussed clinical predictions: many cases permit these without question. Many of these cases express the institutional role argument: "The final determination of dangerousness lies with the courts, not the expertise of psychiatrists and psychologists. Courts must balance society's interest in protection from harmful conduct against the individual's interest in personal liberty and autonomy." *In re Registrant G.B.*, 685 A.2d 1252, 1256 (N.J. 1996) (holding admissible expert testimony in a Megan's Law case). See also *In re A.I.*, 696 A.2d 77, 81-82 (N.J. Super. Ct. App. Div. 1997) (affirming high risk designation for petitioner under Megan's Law, despite contrary expert opinion); *In re J.L.J.*, 509 A.2d 184, 186 (N.J. Super. Ct. App. Div. 1985) (affirming general commitment, despite allegedly unanimous psychiatric opinion indicating no danger).

310. See *State v. Krol*, 344 A.2d 289, 296 (N.J. 1975) (setting constitutional minima for civil commitment proceedings). The state also has a sexually violent offender commitment process, N.J. Stat. Ann. § 30:4-27.24 to -27.38 (West 2003) ("New Jersey Sexually Violent Predator Act" or "SVPA"), see also John Kip Cornwell et al., *The New Jersey Sexually Violent Predator Act: Analysis and Recommendations for the Treatment of Sexual Offenders in New Jersey*, 24 SETON HALL LEGIS. J. 1, 5-8 (1999) and the original version of Megan's Law, N.J. STAT. ANN. § 2C:7-1 to :7-11 (West 1999) (the "Registration Law and Community Notification Law") (requiring registration of known sex offenders, and notification to any community in which such an offender resides), a process which requires some predictive testimony.

311. While it applies *Daubert* in toxic tort litigation, it uses *Frye* for all other purposes, both civil and criminal. *State v. Doriguzzi*, 760 A.2d 336, 341 (N.J. Super. Ct. App. 2000) (noting the displacement of *Frye* by *Daubert* in federal courts, but also noting the continued viability of *Frye* in New Jersey, except for toxic tort litigation); *In re Commitment of R.S.*, 773 A.2d 72, 89 (N.J. Super. Ct. App. 2001) (noting the same).

312. *State v. Fortin*, 745 A.2d 509, 513-15 (N.J. 2000) (behavioral expertise subject to *Frye* standard); *State v. Cavallo*, 443 A.2d 1020, 1024 (N.J. 1982) (psychiatric testimony subject to same standard of admissibility as other expert testimony); *State v. Free*, 798 A.2d 83, 92 (N.J. Super. Ct. App. 2002) (reversing a court order admitting expert psychiatric testimony on false confessions where trial court applied a *Daubert* rather than a *Frye* standard).

313. *State v. Harvey*, 699 A.2d 596, 621 (N.J. 1997).

sessments.³¹⁴ Seven witnesses testified about “general acceptance” at trial, and the parties submitted briefs identifying both scientific literature and legal precedent for and against the admission of the actuarial instruments in question. As in Washington (and Iowa), the court relied largely on prior judicial admission of these opinions as an indicator of general acceptance.³¹⁵ Yet its discussion often veered away from general acceptance, and focused both on the reliability of the underlying science and on the utility of the evidence to the dispute.³¹⁶ The court reviewed each witness’s testimony in detail,³¹⁷ and noted two primary realities: the constitutional adequacy of predictive opinion,³¹⁸ and the incremental value of predictions to fact-finding on danger, given its “sufficient reliability.”³¹⁹

The court held the *Frye* standard did not require unanimous acceptance: “[A] party need not necessarily show there is a unanimous belief in the absolute infallibility of the techniques that underlie the scientific

314. *In re Commitment of R.S.*, 773 A.2d 72, 86 (N.J. Super. Ct. App. 2001). The court limited its eventual acceptance of the expertise to cases in which it was used as part of clinical testimony. *Id.* at 90.

315. *Id.* at 95–97. The court cited to cases from Washington (discussed in this section), California (discussed at Part II.A.3 *infra*), Minnesota and a series of cases which have accepted actuarial assessments “without comment” (both discussed at Part II.B.2 *infra*).

316. *Id.* at 88 (“The sole question, then, is whether actuarial instruments as indicators of sexual offender recidivism have achieved a state of the art so that an expert’s testimony based in part upon them is sufficiently reliable.”).

317. *In re Commitment of R.S.*, 773 A.2d at 76–84. The state offered four witnesses, and R.S. three; all seven witnesses agreed to a series of propositions: (1) Actuarial assessments were not “psychological” tests, in the sense that they did not test a specific personality trait or mental disorder of the respondent, *id.* at 92; (2) Actuarial assessments serve as tools to assess risk; they rely on static features the respondent’s personality, and identify the risk of reoffense for a group which shares those features, *id.* at 77; (3) Thus, actuarial assessments did not in fact predict that the specific respondent would commit a sexual offense in the feature, but only allowed the inference that someone with characteristics like the respondent’s had a specified risk of reoffense, *id.* at 79–80; (4) As risk assessment tools, many but not all of the actuarial instruments at issue had been statistically validated, *id.* at 78; (5) Finally, the experts agreed generally as to how the various instruments ranked in order from most to least statistically valid. *Id.* at 78–85. The experts disagreed about general acceptance, on the validity of tests as predictive tools, and also use of these tools as aids to judicial fact-finding. *Id.*

318. *Id.* at 89–91. The court reviewed *Barefoot’s* constitutional approval of clinically-based predictions, noting that the case had approved clinical predictions, and stating that, given this, it would certainly have approved actuarial assessments. The court also noted that prior New Jersey precedent had approved as “reliable” or “helpful” the scale for assessing risk embedded in New Jersey’s Megan’s Law (the RRAS), which all of the experts in this case had agreed to be the least validated form of risk assessment tool. “Since the RRAS satisfied the requirements of due process and fundamental fairness . . . , we conclude it also satisfies these constitutional elements in the present matter.” *Id.* at 88.

319. *Id.* at 91. The court saw sufficient reliability as heavily dependent on the “context of the proceedings involved,” *id.* (quoting *State v. Cavallo*, 443 A.2d 1020, 1026 (N.J. 1982)), and determined that the expert testimony in instant case supported a conclusion of reliability: “The extensive expert testimony . . . concerning validation studies, cross-validation studies, reliability studies, correlation coefficients, and clinically-derived factors attests to such reliability in this context, where the actuarials are not used as the sole or free-standing determinants for civil commitment.” *Id.*

evidence. . . . [The proponent must show] that . . . the relevant scientific community widely, but perhaps not unanimously, accepts [the technique] as reliable."³²⁰ The court narrowed its definition of the relevant community primarily to authors of articles favorable to actuarial assessments.³²¹ It rejected articles offered by the defendant by discounting the credibility and weight which the court would ascribe to the authors' conclusions.³²² New Jersey precedent, in its view, did not require "a specific number of articles to satisfy the test of general acceptance."³²³ All it sought was "a consensus of acceptance," which the court found in the existence of many "serious" articles, in listings of workshops, and in the availability of sources on the Internet.³²⁴

In re Commitment of R.S. reflects the only published *de novo* decision on the facts underlying admission of predictive testimony.³²⁵ To its credit, the opinion offers a fairly clear explication of the statistical and inferential validity of actuarial tools. It does an adequate job of applying *Daubert's* reliability concerns to predictive testimony. But it said that it applied *Frye*; and as a *Frye* analysis, the case distorts the controlling standard, and illustrates three primary weaknesses of applying *Frye* to expertise as disputed as predictive opinion:

- (1) *Degree of professional acceptance.* The New Jersey court played with numbers by limiting its pool of favorable articles and by padding its numbers with references to workshops and Internet citations. The court did not explain why it chose one pool of consensus and rejected others.

320. *Id.* at 89.

321. *Id.* at 92–93. The court's narrow focus on articles submitted by the parties is, of course, a natural and reasonable outgrowth of the adversarial system. At the same time, it seems a flaw in the *Frye* standard, as applied to scientific literature, to rest an assessment of that literature on the capacities and resources of the parties to such proceedings.

322. For example, as to one of the respondent's articles, the court formed an "impression . . . that while the author has extensive experience testifying as an expert witness in commitment hearings, he has little experience actually developing or testing risk assessment techniques." *Id.* at 94. As another example, the Court used a single article written by one of the respondent's witnesses, Dr. Randy Otto, as a means of rebutting the weight to be ascribed to that witness' own testimony. *Id.* at 93.

323. *Id.* at 94.

324. *Id.* at 94–95.

325. Despite its limits, the case strongly influenced the Iowa appellate court to reverse course completely, from excluding actuarial assessments as unreliable science to affirming their admission as sufficient for judicial purposes. See *In re Detention of Holtz*, 653 N.W.2d 613, 619 (Iowa Ct. App. 2002) (discussed in Part II.A.2 *supra*). The only other state to have applied *Frye* directly, predictive opinions came without discussion under *Frye*. *State v. Villez*, 942 P.2d 522, 538 (Haw. 1997) (holding expert psychiatric predictions of danger admissible in civil commitment case, without any *Frye* analysis). Hawaii uses a *Frye* standard, see *State v. Montalbo*, 828 P.2d 1274, 1279–81 (Haw. 1992), but the court in *Villez* did not apply the "general acceptance" standard. 942 P.2d at 538. Instead, it held that, in reaching an opinion, the expert used materials "reasonably relied upon by experts in the field of clinical psychology . . ." *Id.*

- (2) *General acceptance for forensic purposes.* The New Jersey court noted the frequency with which actuarial assessments are used for purposes of “screening” or for “risk assessment” decisions during treatment; but pure actuarial assessment in fact reflects a small portion of predictive expertise.³²⁶ The court did not explain how and why it chose to accept a particular usage as persuasive.³²⁷
- (3) *Strength of professional disagreement.* The New Jersey court referred frequently to its chosen community as “professionals who assess sex offenders for risks of reoffense”;³²⁸ but the court narrowed the pool to only those experts who favor their own expertise.³²⁹ The court offered no rationale for this narrowing.

The general acceptance standard seems inadequate where deep divisions about a given expertise exist in the expert community. One might expect that *Frye* might require exclusion in such a case. But the New Jersey and Washington cases do the opposite. They transform the *Frye* standard into something very much like a “sufficient reliability” test, which focuses most notably on the mild but distinct incremental helpfulness of these opinions to the fact-finding process. The concerns expressed in these cases parallel the approach of the *Daubert* courts; they in effect balance reliability and fit under the guise of a “sufficient reliability” test.

326. See *supra* note 122 (citing an acknowledgment that the vast majority of assessments of danger occur clinically, and that few assessors use purely actuarial methods).

327. More specifically, a court would assess whether all persons engaged in the particular practices (e.g., danger assessments) use such assessments; to determine the entire range of opinion within the whole group; and to justify explicitly why one group’s opinion has been selected as “acceptance” over another.

328. See, e.g., *In re Commitment of R.S.*, 773 A.2d 72, 90 (N.J. Super. Ct. App. Div. 2001).

329. Two variations of this approach are plausible: either explicitly to describe the “particular field” as including only proponents of the particular science; or to define the field using definitions unrelated to the science, but in such a way as to include only proponents of the science. Both are logically flawed, explicitly results-oriented, and inconsistent with a *general* acceptance standard. See Donna Cropp Bechman, *Sex Offender Civil Commitments: Scientists or Psychics?*, 16 CRIM. JUST. 24, 31 (2001) (“the ‘relevant scientific community’ for purposes of analyzing [actuarial instruments] under *Frye* or *Daubert* is not only comprised of the clinicians who administer the instruments . . .”).

3. *Exempting Predictive Expertise*

Many states use neither *Daubert* nor *Frye*'s general acceptance standard to gauge the admissibility of expert testimony.³³⁰ Indeed, many states explicitly exempt predictive expertise from any special testing,³³¹ applying instead other rules on expert testimony and standard tests for relevance. This section first discusses the rationale for exempting predictive expertise, and then assesses how it might fare under traditional relevance analysis.

Many *Frye* states apply the test to "novel scientific evidence," but exempt certain categories of expertise from the *Frye* standard, notably medical or psychiatric expertise. For example, California courts have carved out an exemption from the state's *Frye* test³³² for medical and psychiatric testimony. While expertise generally creates a false "aura of certainty" that might overly influence a fact-finder,³³³ the concerns do not apply to a medical or psychiatric expert. Rather, "jurors may temper their acceptance of his [sic] testimony with a healthy skepticism born of

330. Some states have adopted neither *Daubert* nor *Frye*, but have instead chosen their own distinct standard for assessing the admissibility of expert testimony. See, e.g., *Harper v. State*, 292 S.E.2d 389, 395 (1982) (affirming exclusion of "truth serum" medication; standard for admitting novel scientific testimony is "whether the procedure or technique in question has reached a scientific stage of verifiable certainty, or . . . whether the procedure 'rests upon the laws of nature'"). No state has yet applied such a state-specific test to predictive expertise.

One might also argue that the rationales of such cases as *In re Detention of Holtz*, 653 N.W.2d 613 (Iowa App. 2002), or *In re Commiunent of R.S.*, 773 A.2d 72 (N.J. Super. Ct. App. Div. 2001), create what are in effect new state standards that do not comply rigorously with the mandates of *Daubert* (in Iowa) or *Frye* (in New Jersey). Yet these cases are more accurately seen as adjustments (or distortions) of the relevant standards, rather than as entirely new standards.

331. *Kumho Tire* extended *Daubert* to all types of expertise; if applied in all *Daubert* jurisdictions, this approach eliminates any exemption for predictive expertise. See *supra* Part II.A.3 (discussing *Kumho Tire*'s clarification that *Daubert*'s holding applies to "scientific, technical or other specialized knowledge").

332. *Kelly*, 549 P.2d 1240, 1251 (Cal. 1976) (reversing admission of testimony based on spectrographic voiceprint analysis, and affirming *Frye* as the prevailing test for 'novel' scientific evidence). The California Supreme Court has specifically retained its test and rejected *Daubert* as unpersuasive. *People v. Leahy*, 882 P.2d 321, 331 (Cal. 1994) (rejecting *Daubert* and reaffirming *Kelly/Frye* as the controlling standard). In *Leahy*, the court responded to criticism of *Frye* by reasserting its rationale from *Kelly*, and by noting some purported advantages of *Frye* over *Daubert*: the outsourcing of establishing new scientific procedures from the courtroom to the laboratory; the ability of trial courts to avoid *de novo* review each time expertise appears; and finally, the availability of "a battery of well-qualified scientific and medical personnel" as witnesses. *Id.* at 327-30. The opinion stresses the presumed inexperience of trial judges in assessing scientific principles. *id.* at 331, and clarified the unanimity required for general acceptance: "the views of a typical cross-section of the scientific community, including representatives, if there are such, of those who oppose or question the new technique." *Id.* at 336 (citing *Kelly*, 549 P.2d at 1248).

333. See *Kelly*, 549 P.2d at 1245.

their knowledge that all human beings are fallible.”³³⁴ Thus, *Frye* does not apply “even when the witness is a psychiatrist and the subject matter is as esoteric as the reconstitution of a past state of mind or the prediction of future dangerousness”³³⁵

California courts have exempted predictive testimony from the *Frye* test virtually without dissent.³³⁶ Non-evidentiary cases offer no barrier to predictive expertise,³³⁷ and evidentiary cases repeatedly confirm the exemption of psychiatric predictions from *Frye*. For example, in *People v. Ward*, the defendant in a sexually violent predator commitment proceeding argued that clinical predictions lacked general acceptance.³³⁸ The

334. *People v. McDonald*, 690 P.2d 709, 724 (Cal. 1984) (reversing a conviction where trial court excluded psychiatric testimony on the fallibility of eyewitness identification).

335. *Id.* While California courts have uniformly admitted predictive expertise, they have reached disparate results for other uses of psychiatric testimony. For example, in *People v. Stoll*, the California Supreme Court affirmed the admission of expert testimony offering a diagnosis of defendant’s mental illness, to support the inference that defendant could not have acted as charged. *People v. Stoll*, 783 P.2d 698, 710 (Cal. 1989). The court held that the testimony satisfied the exemption because it contained nothing which might unduly prejudice the jury with an unwarranted aura of certainty. *Id.* The Court found the uncertainty of such diagnoses (which combined clinical and psychometric assessments) untroubling: “[T]his process is a learned professional art, rather than the purported exact “science” with which *Kelly/Frye* is concerned” *Id. Contra* *People v. Bledsoe*, 681 P.2d 291, 301 (Cal. 1984) (applying *Kelly/Frye* to exclude expert testimony concerning child sexual abuse accommodation syndrome); *People v. Shirley*, 723 P.2d 1354, 1383–84 (Cal. 1982) (applying *Kelly/Frye* to exclude post-hypnotic testimony of a hypnotised witness).

336. *See* *People v. Murtishaw*, 631 P.2d 446, 471 (Cal. 1981) (reversing death sentence due to improper admission of predictive testimony by a psychopharmacologist). The Court in *Murtishaw* did not apply *Kelly/Frye*, but excluded the predictions because the probative value of “unreliable predictions” was outweighed by its substantial prejudicial impact. *Id.* at 470–71. *See infra* notes 358–61 and accompanying text (discussing relevance analysis of predictive expertise). Indeed, the court specifically distinguished civil commitments from death penalty cases:

In most of the other cases in which courts have upheld admission of opinion testimony forecasting future violence . . . the trier of fact is required by statute to determine whether a person is “dangerous” In such cases expert prediction, unreliable though it may be, is often the only evidence available to assist the trier of fact.

Id. at 469.

337. Neither constitutional nor procedural rules required the exclusion of predictive expertise. *People v. Bennett*, 182 Cal. Rptr. 473, 478 (Cal. Ct. App. 1982) (affirming the use of expert predictive testimony without reference to *Kelly* or *Frye*); *People v. Henderson*, 166 Cal. Rptr. 20, 27 (Cal. Ct. App. 1980) (affirming admission of expert testimony on danger in a sexual offender commitment case); *People v. Mapp*, 198 Cal. Rptr. 177, 181–82 (Cal. Ct. App. 1983) (affirming continued commitment, and noting the centrality of expert predictions of danger in commitment cases); *People v. Super. Ct. (Blakely)*, 70 Cal. Rptr. 2d 388, 398 (Cal. Ct. App. 1997) (remanding case for trial on state’s claim for continued commitment: “the instant statutory scheme contemplates that in said proceeding, the trier of fact shall be aided by the expert testimony of psychologists or psychiatrists”). These non-evidentiary cases created the procedural and substantive context within which the imperfections of predictive testimony could survive. *See Henderson*, 166 Cal. Rptr. at 27 (“Although admittedly those techniques do not produce certainty, the significance of this failure to meet an ideal of perfection is a consideration for the trier of fact in weighing the effect of the testimony.”)

338. *People v. Ward*, 83 Cal. Rptr. 2d 828, 832 (Cal. Ct. App. 1999) (affirming the admission of clinical predictions of future dangerous sexual conduct).

court disagreed and distinguished predictive testimony from evidence “involving novel devices or processes” and expert medical testimony.³³⁹ It stressed that, “in civil commitment cases, where the trier of fact is required by statute to determine whether a person is dangerous or likely to be dangerous, expert prediction may be the only evidence available.”³⁴⁰

California courts have reached the same result in evaluating actuarial assessments. In *Garcetti v. Superior Court*, the trial court had excluded an actuarial assessment tool used as the basis for expert’s prediction in a sexually violent predator case.³⁴¹ The appellate court reversed, noting the long-held exemption for opinions by doctors and psychiatrists, and extended the exemption to the methods used by experts in formulating their predictions. The court reasoned that “it does not matter if the psychiatrist used clinical or actuarial models or even whether the psychiatrist followed the *Diagnostic and Statistical Manual* published by the American Psychiatric Association, since experts are not restricted to one methodology or another in rendering predictions on future dangerousness.”³⁴² Holding that the exemption from *Frye* applied, the court was remarkably unconcerned with the weaknesses of expert predictions. “[I]t is of no consequence,” the court wrote, “if a difference of opinion exists among professionals relating to which methodology should be utilized. . . . It is also of no consequence that the reliability of the instrument being right is only 70 percent according to validity or accuracy rates.”³⁴³

339. *Id.* (citing *People v. Stoll*, 783 P.2d 698, 712 (Cal. 1989) (psychiatric prediction is art, not science)).

340. Indeed, at least a few cases have suggested that expert psychiatric predictions are not only the only evidence, but legally required and even legally sufficient in their own right to sustain a commitment order. *See, e.g., People v. Devers*, No. A095661, 2002 WL 724931, at *3 (Cal. Ct. App. Apr. 25, 2002).

341. *Garcetti v. Super. Ct.*, 102 Cal. Rptr. 2d 214, 216–17 (Cal. Ct. App. 2000).

342. *Id.* at 238.

343. *Id.* Later cases follow this result. In *People v. Woods*, No. C037203, 2001 WL 1649216, at *5–*8 (Cal. Ct. App. Dec. 20, 2001), an appellate court reviewed admission of two actuarial assessment tools (the “RRASOR” and the “Static 99”). The court noted that the instruments had been used as part of a clinical assessment; each was “merely the starting point in the expert’s analysis . . . [and] . . . not an infallible prediction of the likelihood to reoffend.” *Id.* at *6. The court held that, since “[b]oth experts testified the assessment instruments used were imperfect predictive tools, and that other factors were also considered in reaching an opinion,” the exemption from *Kelly / Frye* still applied. *See also People v. Hayes*, No. A093285, 2002 WL 462277, at *5–*6 (Cal. Ct. App. Mar. 26, 2002) (affirming admission of an actuarial assessment tool “Static 99” when used as part of a clinical assessment).

But see People v. Williams, No. D035886, 2001 WL 1464186, at *4–*7 (Cal. Ct. App. Nov 19, 2001) (applying *Kelly* to an actuarial assessment tool (HCR 20) and determining that the tool was “generally accepted”). *Williams* made no mention of the exemption; the only California decision to apply the traditional test to any part of a predictive opinion, the case appears to be an anomaly.

A similar approach to exempting predictive expertise appears in Arizona, another pro-*Frye*³⁴⁴ and anti-*Daubert* state.³⁴⁵ As the Arizona Supreme Court explained in *Logerquist v. McVey*:

Frye is inapplicable when a qualified witness offers relevant testimony or conclusions based on experience and observation about human behavior for the purpose of explaining that behavior. . . .

....

. . . [E]xpert evidence based on a qualified witness' own experience, observation, and study is treated differently from opinion evidence based on novel scientific principles advanced by others.³⁴⁶

The Arizona court thus distinguished opinions resting on the expert's own experience, observation, and study from opinions based on principles advanced by others and applied by the expert to the facts of the relevant case.³⁴⁷ The court stated that any disagreement in the relevant community went to weight, and not to admissibility.³⁴⁸ The court did

344. *State v. Valdez*, 371 P.2d 894, 898 (Ariz. 1962) (excluding polygraph evidence due to lack of general acceptance). Note that, unlike California, Arizona uses an evidentiary code closely patterned after the Federal Rules. See, e.g., ARIZ. R. EVID. 702.

345. *Logerquist v. McVey*, 1 P.3d 113, 133-34 (Ariz. 2000) (reversing a trial court's application of the *Frye* test to exclude psychiatric testimony in support of a claim of repressed memory of sexual abuse). The Arizona Supreme Court had deferred until 2000 the decision whether to incorporate the new federal standard. See *State v. Bible*, 858 P.2d 1152, 1183 (Ariz. 1993) (reserving decision on the adoption of *Daubert*, given the number of unresolved questions about its application); *State v. Johnson*, 922 P.2d 294, 296 (Ariz. 1996) (same: "The federal courts have not yet had a fair opportunity to apply *Daubert*; thus it is too early to properly evaluate it."). It criticized *Daubert* for what it described as an unwarranted shift of decision-making authority from the jury to the trial judge, enabling trial judges to substitute their judgment about the reliability (and even the credibility) of expertise for that of the jury. *Logerquist*, 1 P.3d at 130-32. It also noted: the relative inexperience of trial judges in assessing scientific reliability, *id.* at 129; the burdensome, time-consuming nature of pre-trial hearings in a *Kumho* regime, *id.*, and the likelihood that *Daubert* would produce more arbitrary results, thus more uncertainty than the *Frye* standard, *id.* at 125.

346. *Logerquist*, 1 P.3d at 123.

347. *Id.* A later Arizona court would refer to this distinction as a distinction between inductive and deductive reasoning. *State ex rel. Romley v. Fields*, 35 P.3d 82, 88 (Ariz. Ct. App. 2001) ("the inductive-deductive dichotomy in *Logerquist*"). In *Logerquist*, the court relied both on California precedent, and on a range of Arizona cases, including: *State v. Lindsey*, 720 P.2d 73, 77 (Ariz. 1986) (behavior patterns of incest victims: *Frye* not applicable, evidence admitted); *State v. Roscoe*, 700 P.2d 1312, 1319-20 (Ariz. 1984) (dog handler's interpretation of tracking dog's scent identification: *Frye* not applicable, evidence admitted); *State v. Varela*, 873 P.2d 657, 663-64 (Ariz. Ct. App. 1993) (child sexual abuse accommodation syndrome: *Frye* not applicable, evidence admitted); *State v. Tucker*, 798 P.2d 1349, 1355 (Ariz. Ct. App. 1990) (behavior of child molesters and victims generally: *Frye* not applicable, evidence admitted). *Logerquist*, 1 P.3d at 119-21. With these, the Court contrasted *State v. Hummert*, 933 P.2d. 1187, 1192-93 (Ariz. 1997) (statistical assessments of probabilities derived from DNA statistics: *Frye* applicable, evidence excluded).

348. *Logerquist*, 1 P.3d at 124.

not address predictive testimony, although it did allude to the unreliability of expert predictions in arguing against the adoption of *Daubert*.³⁴⁹

The distinction between the expert's own observation and the expert's interpretation of separately derived data creates a possible difficulty for actuarial assessments. Barely eighteen months after *Logerquist*, an Arizona court faced exactly that challenge.³⁵⁰ The trial court relied on *Logerquist's* distinction to exclude testimony based on actuarial assessments, which contained "data derived by a technique or principle developed by others."³⁵¹ The appellate court reversed. It noted that the lower court's reliance on *Logerquist* was understandable, but found its reliance too mechanical. *Logerquist's* holding applied to any expert behavioral evidence.³⁵² Actuarial assessments did not have "an aura of infallibility . . . they are subject to interpretation and their predictive value is far less than 100%. In addition, the testifying expert must still explain to the fact-finder why he or she believes that a particular individual will likely re-offend or not re-offend."³⁵³

Only California and Arizona have explicitly exempted expert psychiatric predictions of danger from special testing under a *Frye*- or *Daubert*-like analysis.³⁵⁴ But few cases in any state impose special standards on the admissibility of expert predictive testimony in commitment cases; indeed, few discuss admissibility at all. Including California and Arizona, only seven states have articulated rationales for the handling of

349. *Id.* at 126. The court assumed that *Daubert* would exclude predictive testimony, and saw that result as in conflict with the Supreme Court's approval of psychiatric predictions in *Barefoot*. *Id.* at 126-27. How, the Arizona court asked, could the Supreme Court's acceptance of predictive expertise in the constitutional case square with its new evidentiary standard of scientific reliability? *Id.* at 127. The Arizona court speculated that perhaps the Supreme Court might apply different standards in civil and criminal cases, or perhaps the Supreme Court "had reason to see things differently in the ten years that elapsed between *Barefoot* and *Kumho*." *Id.* See also Michael H. Gottesman, *From Barefoot to Daubert to Joiner: Triple Play or Double Error*, 40 ARIZ. L. REV. 753, 753 (1998), cited in *Logerquist*, 1 P.3d at 127.

350. State *ex rel.* Romley v. Fields, 35 P.3d 82, 88 (Ariz. Ct. App. 2001) (holding that *Frye* did not apply to clinical testimony based on actuarial assessments of future sexually dangerous behavior). The case arose after "[d]ozens of individuals . . . filed motions requesting *Frye* hearings to contest the admissibility of expert opinion testimony on recidivism based on actuarial instruments in SVPA hearings." *Id.* at 84.

351. *Id.*

352. *Id.* at 88 (citing *People v. Ward*, 83 Cal. Rptr. 2d 828 (Cal. Ct. App. 1999); *People v. Stoll*, 783 P.2d 698 (Cal. 1989)).

353. *Romley*, 35 P.3d at 89.

354. A Florida appellate court has also noted (in dicta) that it would also have exempted predictive testimony *Frye*. *Westerheide v. State*, 767 So. 2d 637, 657 (Fla. Dist. Ct. App. 2000) (*Frye* claim concerning clinical predictions withdrawn before appeal, case resolved on other grounds, 831 So. 2d 93, 112 (Fla. 2002)). Interestingly, the lower court noted that actuarial assessments "may very well be subject to a *Frye* analysis." *Id.*

expert predictive testimony;³⁵⁵ all other state cases that discuss predictive testimony use rationales unrelated to its admissibility.³⁵⁶ Traditional evidentiary analysis would identify two sources of additional challenges: other provisions of the rules for expert testimony, and standard screening for relevance.³⁵⁷

A good example of the latter appears in a Colorado decision,³⁵⁸ in which the Colorado Supreme Court affirmed admission of predictive expertise using standard screening for relevance and prejudicial impact.³⁵⁹

355. The four states already discussed also include Texas and Iowa (*Daubert*), and Washington and New Jersey (*Frye*). See *infra* notes 358–61 and accompanying text (discussing the last state, Colorado.)

356. See cases cited *supra* under Part I.C (discussing the impact and influence of psychiatric testimony on civil commitment proceedings). How should one interpret the dearth of case law on point? Why have advocates mounted so few sustained challenges on evidentiary grounds, especially given the persistent scholarly and psychiatric criticism of predictive testimony? To be sure, advocates may prefer to use the academic criticism to contest its weight, or even to offer contradictory testimony, hoping to dilute the persuasive impact of predictive testimony with their own conflicting expertise.

It also seems plausible that *Barefoot* has had a suppressing effect on both trial courts and practitioners. We have seen in this and preceding subsections how often courts have referenced *Barefoot*'s constitutional accommodation of predictive expertise. See *infra* Part II.C.1 (discussing the admissibility of predictive testimony under due process standards).

It seems less plausible to assume that only these six states (and a handful of others) have admitted predictive testimony in commitment cases, while all others have excluded them. Not only is this inconsistent with the reasoning of the cited cases, and with *Barefoot*; it also seems inconsistent with the central role that predictive testimony plays in the commitment process.

357. See *Logerquist v. McVey*, 1 P.3d 113, 132 (Ariz. 2000). The court there said:

The Rules of Evidence, and Rule 702 itself, erect barriers to admission of all opinion evidence: the evidence must be relevant, the witness must be qualified, and the evidence must be the kind that will assist that jury.

[T]he rules also permit trial judges to reject even relevant evidence that meets the Rule 702 test if the probative value is "substantially outweighed by the danger of unfair prejudice, confusion . . . or misleading the jury, or by considerations of undue delay, waste of time . . ." ARIZ. R. EVID. 403.

Id. (citations omitted); *State ex rel. Romley v. Fields*, 35 P.3d 82, 89 (Ariz. Ct. App. 2001) (following *Logerquist* and citing Arizona Rules 702, 703 and 403 as baseline tests for admissibility of predictive testimony). See also *infra* Part II.B (discussing traditional challenges to expert qualifications).

358. *Vialpando v. People*, 727 P.2d 1090, 1093–96 (Colo. 1986) (reversing the exclusion of the petitioner's expert witnesses using a traditional evidentiary analysis). The trial court in this case had excluded witnesses offered by the petitioner for release from a criminal commitment, despite the fact that he had allowed the prosecution to offer witnesses testifying to the petitioner's future dangerousness. The Court made no mention of *Frye*, although Colorado uses that standard, *Fishback v. People*, 851 P.2d 884, 890 (Colo. 1993) (affirming Colorado's use of the *Frye* standard), and recognizes an exemption from *Frye* for "social science or experienced-based opinions," *Schultz v. Wells*, 13 P.3d 846, 849 (Colo. App. 2000) (accident reconstruction evidence: *Frye* not applicable, evidence excluded), citing *Brooks v. People*, 975 P.2d 1105, 1115 (Colo. 1999) (interpretation of dog-tracking evidence: *Frye* not applicable, evidence admitted) citing cases on either side of the distinction, *Schultz*, 13 P.3d at 849–50.

359. *Vialpando*, 727 P.2d at 1094 ("whether . . . the proffered evidence relates to a fact that is of consequence to the determination of the action . . . whether . . . the proffered evidence makes the existence . . . [of that fact] more probable or less probable . . . [and] whether . . . the probative value of the evidence is substantially outweighed by the danger of unfair prejudice, confusion of the issues, or misleading the jury"), citing COLO. R. EVID. 401–403. See also FED. R. EVID. 401–403.

The predictive testimony satisfied this screening: “[T]he potential dangerousness of a defendant is a critical element. . . . The proffered evidence tended to make the existence of a ‘fact of consequence to the determination of the action’—the absence of future dangerousness—more probable with the evidence than without it.”³⁶⁰ Weighing probative value against potential prejudicial impact, the court noted the centrality of the danger issue, the conflict in testimony between experts, and the presence (in the excluded testimony) of “foundational” information that the jury could use to assess all of the expertise.³⁶¹

Does predictive testimony really require expertise? If behavioral evidence lacks any “aura of certainty,” the fact-finder can evaluate its content in light of his or her own experience, and lay testimony on the topic should be admissible. Such an approach would eliminate even the standard testing for credentials and basis imposed by most jurisdictions on expert testimony, and leave only relevance testing.³⁶² But no jurisdiction has approved lay predictions of danger to the exclusion of expert testimony. At best, courts have indicated that, as a matter of evidentiary sufficiency, commitment courts may rely on lay opinion and discount expert testimony on danger.³⁶³ None go so far as to approve a lay witness’ opinion about prospective danger. Given the persuasive and functional

360. *Id.* at 1095.

361. *Id.* at 1096. The unfairness of excluding the patient’s experts while admitting the state’s had a strong influence on the court in *Vialpando*. One might construct an argument that the probative dangers of expert predictive testimony substantially outweighs its probative value. The unreliability of predictive testimony should severely reduce its probative value under a 403 analysis; moreover, the certainty with which predictions can be phrased, coupled with the potential confusion arising from imprecise use of actuarial assessments, pose strong probative dangers. Yet no case has excluded predictive testimony from civil commitment on these grounds; and the very centrality of the danger issue in commitments, coupled with the difficulty of predicting danger generally, may well counterbalance the probative risks of such testimony. See *People v. Murtishaw*, 631 P.2d 446, 469 (Cal. 1981) (excluding predictive expertise from the sentencing phase of capital cases, but distinguishing civil commitments on these grounds).

362. At the same time, many of the functional advantages of expert testimony would disappear. See *supra* Part I.A. (discussing differences between lay and expert testimony).

363. See *Hill v. State*, 358 So. 2d 190, 206–09 (Fla. Dist. Ct. App. 1978) (discussing the balance between lay and expert testimony in commitment cases, and indicating that since “psychiatrists’ predictions of future dangerousness are increasingly subject to doubt, . . . lay testimony on an issue may be more weighty than that of experts”); *People v. Hockenberry*, No. A095277, 2002 WL 1000075, at *2–*3 (Cal. Ct. App. May 16, 2002) (affirming admission of testimony about future danger from an individual who was neither a psychiatrist or a psychologist); *People v. Sword*, 34 Cal. Rptr. 2d 810, 819 (Cal. Ct. App. 1994) (affirming a determination of dangerousness in the absence of any expert psychiatric testimony on danger); *In re Melton*, 597 A.2d 892, 898 (D.C. 1991) (affirming the admission of expert testimony on danger despite defendant’s challenge to the experts’ qualifications because “there is nothing inherently unattainable [even] about a [lay judicial officer’s] prediction of future criminal conduct for purposes of preventive detention”) (internal quotations omitted). All of these cases affirm the sufficiency of lay observation testimony to support a trial determination of danger; none explicitly approve admission of lay testimony on danger.

advantages of expert testimony, it seems unlikely that any advocate would choose to try.

In conclusion, exempting predictive expertise from special testing under *Frye* seems puzzling at best. It rests on an untested, probably untestable empirical assertion: that expert testimony about behavioral matters lacks “the aura of authority” for a fact-finder that justifies any special testing. To exempt strongly controversial expertise from *Frye*’s focus on general acceptance also seems questionable.

Yet exemption offers two useful insights into the evidentiary problems posed by predictive expertise. Predictive testimony may lack “general acceptance,” or for that matter, “scientific reliability,” and that failure may justify exclusion. Yet predictions can never be made with certainty; predictive testimony may thus never attain “general acceptance” or “scientific validation.” Faced with a legal mandate to decide danger, courts will admit whatever help they can, and adjust or make exceptions to evidentiary rules where useful information does not and cannot satisfy their requirements. While doctrinally questionable, this approach has strong pragmatic appeal, one which underlies not only the exemptions from *Frye* but also the “fit” component of *Daubert*.

Second, these courts contend that fact-finders can assess the weaknesses of behavioral testimony, and that objections about reliability go to weight, not to admissibility. These contentions may express a deeper insight: that the inferential processes through which “experts” predict danger parallel the mental process through which fact-finders reach a finding of dangerousness. Fact-finders and experts alike must draw inferences from history, testimony, first-hand exposure, and specialized inference (some of it statistical) to reach conclusions about danger. In exempting predictive testimony, these courts essentially accept a kinship with the predictive experts, which permits courts to adapt the expertise to the difficult exercise of decision-making in this context.

B. FIT AND RELIABILITY OF EXPERT PREDICTIONS IN CIVIL COMMITMENTS

The preceding sections argue that every court that has accepted predictive expertise has either distorted or refused to apply stated evidentiary standards. One might well end the analysis here, and chalk up the results to the courts’ desire to protect the public from violent, mentally ill people. Such capitulation to a desired result, however, seems both ill-advised and unnecessary. A richer understanding of the *Daubert* standard permits the construction of a coherent rationale for admitting predictive expertise, without abandoning the rule of law. This reassessment focuses on the twin concerns of fit and reliability, and articulates a fuller content for each. To limit reliability to purely scientific validation, and to

equate fit with baseline logical relevance raises severe problems with any opinion that lacks (and which is likely always to lack) validation by scientific method. It thus seems appropriate to ask whether fit and reliability mean something more complex than simple relevance and scientific validity.

The notion of fit includes logical relevance; but this Article has already suggested that “fit” entails other questions: (1) the degree to which courts have accommodated concerns over the expertise within the structure and content of the relevant legal process; (2) the degree to which the particular expertise supplements and eases the inherent difficulty of fact-finding on an issue; (3) the prevalence of the particular expertise within the relevant process; and (4) the similarity of the inferential process embodied in the opinion to the inferential processes required for fact-finding.³⁶⁴ Where answers to these questions indicate strong fit for a given expertise, fit may then counter reliability and overcome concerns about scientific validity.

Reliability also means something different and broader than scientific validity. Expertise must use “the same level of intellectual rigor that characterizes the practice . . . in the relevant field,”³⁶⁵ but that does not mean that a court must use the same methods, or apply them to the same purpose. Instead, it means that the ideas, methods, and application embodied in the opinion must be sufficiently reliable for use in the courtroom. Trial courts may adjust the standards of validity scientists might use to the quite different standards that apply in deciding hard cases with limited data and limited time.

The richer notion of “fit” and the dispute-oriented definition of reliability explain and justify the courts’ approaches to predictive testimony. The next subsections explore these standards in the context first of predictive opinions, and then of actuarial assessments.³⁶⁶

364. See *supra* Part I.C (articulating different ways in which predictive testimony fits with civil commitment process).

365. *Kumho Tire Co. v. Carmichael*, 526 U.S. 137, 152 (1997).

366. These sections also discuss briefly, and in passing, the challenges facing courts which continue to adhere to *Frye*.

I. *Clinical Predictions*

Clinical predictions satisfy *Daubert's* fit requirement in a special, and perhaps even unique way in civil commitment proceedings.³⁶⁷ In this setting, courts (and legislatures) have accommodated concerns with expert unreliability in numerous ways: through the development of constitutional standards, heightened burdens of proof, special rules of sufficiency, and the patterning of evidence. Acknowledging the pervasive distrust of predictive expertise, courts have nonetheless noted its incremental utility in carrying out the fundamentally impossible task of precise prediction. Clinical decision-makers and judgments permeate the legal process for commitment—at initiation, in testimony, in treatment, and in termination phases. Courts have skeptically, but persistently, used predictive opinions as a touchstone in the shaping of the commitment process.

Clinical assessments also have a strong affinity with judicial fact-finding on danger. Clinicians assess danger through in-person contact and appraisal of the individual, documentary review, reconstruction of patient history, and even (these days) statistical assessments. Clinicians use this pool of information to reference prior experience with similar cases, to identify and constrain intuitive and subjective judgments, and to formulate predictions of whether the individual will become dangerous. This description applies with equal force to judicial fact-finders. Courts have on occasion attempted to articulate the mix of information that they require in reaching the legal conclusion of danger,³⁶⁸ even using terms that parallel clinical assessments. It thus should come as no surprise that courts find nothing novel or foreign in clinical assessments, or that they believe that fact-finders are able to assess the reliability of the inferences that clinical predictions invite.

However unreliable clinical predictions may be as a science, they have sufficient minimal reliability to satisfy *Daubert's* dispute-oriented reliability standard. To be sure, no one can articulate exactly how a clinical expert will reach a conclusion on danger. This reality results in part from the mix of different kinds of data; and in part from the different kinds of reasoning, including intuition, that predictions require. Yet the constitutional minima of civil commitments permit—indeed require—predictions,³⁶⁹ and the courts have fashioned a process that accommodates this mandate by assuring both a broad pool of information and am-

367. This paragraph summarizes conclusions reached through more thorough analysis in Part I.C., *supra*.

368. See *supra* Part I.C.3 (discussing complex fact-finding as a method to make predictions in civil commitments).

369. See *supra* Part I.C.1 (describing the constitutional requirement of mental illness and danger).

ple help from those with specialized insight. Formed through a process virtually identical to fact-finding, and applied in a context designed to accommodate its weaknesses, expert clinical predictions satisfy *Daubert's* minimal standards of reliability.³⁷⁰

To say these things, however, is not to advocate a *per se* rule of admission for clinical expertise. Proponents should use well-qualified experts, with backgrounds that reveal some experience in clinical assessment of danger.³⁷¹ Experts should come prepared to acknowledge the weaknesses in the inferences underlying a given opinion. They should identify and distinguish among the different data and methods used in formulating the opinion. Where the opinion rests solely on intuition or guesswork, or fails to account for conflicting data, or ignores ambiguities or uncertainties, or substitutes bias for informed and skilled assessment, opponents should attack. Trial courts should exclude testimony that does not rise to a minimal level of utility for dispute resolution purposes, and even where it does, should exclude such evidence when it introduces prejudicial or confusing content.³⁷² If a proponent can overcome these

370. States using a *Frye* standard to assess clinical predictions face a more difficult justification. Strict application of *Frye* to clinical predictions should lead to exclusion. Few if any scientists would testify that clinical predictions have "general acceptance." See *Barefoot v. Estelle*, 463 U.S. 880, 920-21 (1983) (Blackmun, J., dissenting); *In re Young*, 857 P.2d 989, 1017 (Wash. 1993). Courts have either used a severely distorted version of general acceptance, including reliability and familiarity as factors, *supra* Part II.A.2, or by exempting clinical predictions entirely from the demands of *Frye*. See, e.g., *People v. Ward*, 83 Cal. Rptr. 2d 828, 832 (Cal. App. 1999) (affirming the admission of clinical predictions of future dangerous sexual conduct); *supra* Part II.A.3. In other words, there is nothing in the strict logic of "general acceptance" which would permit the admission of predictive expertise. *Frye* courts should thus exempt it from *Frye* testing, relying on the test's traditional limitation to "novel" scientific evidence. Given the deep intertwining of predictive opinion with commitment process, this expertise arguably seems less "novel" than, say, actuarial assessments.

371. The assurance provided by experience relates to one of the analytical processes underlying clinical expertise: the ability to compare a specific pool of data for one person to comparable pools of data gathered while assessing others. In essence, this process replicates informally what actuarial assessments seek to do with statistical formality: compare the individual to a group of other similar individuals, and draw inferences about the individual from that comparison. As Grove and Meehl put it:

the clinician . . . attempts to do a subjective, impressionistic, in-the-head job of actuarial computation. . . . [T]he clinician's brain is functioning as merely a poor substitute for an explicit regression equation or actuarial table. Humans simply cannot assign optimal weights to variables, and they are not consistent in applying their own weights.

Grove & Meehl, *supra* note 76, at 315. The description (and critique) applies with equal force to judicial fact-finders, with the difference that fact-finders have the societal mandate to engage in precisely such an analysis.

372. Phrased more technically, the opponent of psychiatric predictions should challenge *both* on FED. R. EVID. 702 grounds (as interpreted by *Daubert*) and on FED. R. EVID. 403 grounds (relating to the balance of probative value and prejudicial effect). While it is beyond the scope of this paper to assess the relationship between the two rules, I contend that the 702 analysis adds a layer of assessment distinct from and additional to that imposed by FED. R. EVID. 403. *Daubert* relates specifically to expert opinion: the drawing of inferences from complex data; the use of scientific or other specialized methods to organize and structure both data and inference; and the special influence that such inferen-

challenges, however, the joint fit-reliability standard within *Daubert* provides ample ground for admitting clinical predictions.

2. *Fit and Reliability for Actuarial Assessments*

Actuarial assessments prompt a different dynamic balancing of *Daubert*'s twin factors. Actuarial assessments have emerged more recently and their judicial use has so far focused on commitments for sexually violent offenders. Still, *Daubert* should permit admission of these opinions as well, at least for carefully limited purposes.

How well do actuarial assessments fit the demands of civil commitments? Courts have had limited occasion to process concerns about this expert testimony into civil commitment proceedings. They have either limited the inferences drawn from statistical methods;³⁷³ or tried to define the potential uses of various "risk factors" as factual elements sufficient to reach a finding of danger.³⁷⁴ Moreover, it is difficult to assess the prevalence of free-standing actuarial assessments, especially outside commitments for sexual offenders.³⁷⁵ At the same time, actuarial assessments do seem to ease at least one part of the task of making predictions. They permit a fact-finder to understand better how shared risk factors can lead to valid inferences about risks; and (in the opposite direction) might also lead to the deflation of stereotypes about other groups with whom the defendant may share attributes.

How similar are the inferential processes of actuarial assessment to those required for fact-finding in commitment cases? Judicial fact-finding uses multiple sources of data and disparate analytical methods to reach an individualized prediction of danger.³⁷⁶ By contrast, actuarial assessments focus on behaviors of a separate group of diverse individuals who share certain characteristics with a defendant, make statements about

tial methods (if undisclosed or unchallenged) may have on uninformed fact-finders. These concerns justify handling under a rule separate from the fundamental, pervasive balancing of probative value against prejudicial impact.

373. See *In re Young*, 857 P.2d at 1016-17; *In re Commitment of R.S.*, 773 A.2d 72, 94 (N.J. Super. Ct. App. Div. 2001).

374. See *In re Commitment of R.S.*, 773 A.2d 72, 88 (N.J. Super. A.D. 2001) (upholding the use of actuarial instruments "as a factor in the overall prediction process"—"allocating weight to risk factors in accordance with scientific literature and expertise is an acceptable method of predicting future criminal sexual behavior"). This absorption of risk factors into legally acceptable (and perhaps even sufficient) prerequisites for finding danger goes beyond questions of admissibility, and may raise significant concerns about commitment standards which rely solely on "risk factors" for a conclusion of danger. See also Part II.C.3 *supra* (describing complex fact finding as an accommodation for clinical predictions).

375. See *supra* note 81 (citing Grove and Meehl for the proposition that most assessors of danger use clinical and not actuarial methods).

376. See *supra* Part I.B.1 (describing similarities between clinical prediction and judicial fact-finding on danger).

degrees of risk that a defendant shares with that group, and depend exclusively on statistical methodologies for their validity.³⁷⁷

There are three important differences between judicial fact-finding and actuarial methods: in sources, in inferential methods, and in the purpose of the evaluation itself. As to sources, statistical methods focus on similarities between the defendant and the relevant group, and on the behavioral consequences tested by the actuarial assessors for that group. By contrast, judicial fact-finders must make decisions about individuals. Accordingly, they must consider characteristics in the defendant that fall outside the statistically significant similarities, and assess behaviors different in both kind and intensity from those in the statistician's group. In effect, actuarial assessments focus on other people doing other things, while a fact-finder takes the defendant as found, and assesses traits and behaviors in multiple combinations, not just those presented by the actuarial test.³⁷⁸

The differences in the inferential processes are equally fundamental: actuarial assessments make inferences at once broader and narrower than individuated fact-finding requires. To mention the statistical risk is certainly relevant,³⁷⁹ but actuarial assessments make statistical statements about broad pools of individuals, and about risks averaged over that group. Moreover, actuarial assessments address only a well-defined pool of factors that may or may not provide an exhaustive portrait of the individual on trial. In effect, actuarial assessments are inevitably crude, because they cannot measure the full congeries of traits for the person on trial.³⁸⁰

377. See *supra* Part I.B.1.6 (describing actuarial assessment methodologies).

378. Cf. JOHN MONAHAN ET AL., *RETHINKING RISK ASSESSMENT* 130–35 (2002) (identifying “two primary reasons . . . given in support of allowing clinicians the option to use their judgment to revise actuarial violence risk assessment estimates”). This study suggests two comparable risks of inaccuracy for pure actuarial assessments: “questionable validity generalization,” i.e., the use of instruments developed based on a group with one demographic to assess the risk of violence in groups with other demographics; and “rare risk or protective factors,” i.e., the presence of unusual or rare factors in a given case which “precisely because they are rare—will not have been properly taken into account in the construction of the actuarial instrument.” *Id.* at 132.

379. See FED. R. EVID. 401.

380. As to the kinds of behavior recognized as dangerous by the assessment tool, there is some indication that existing tools are underinclusive, and do not fully reflect the range of behaviors recognized as meeting legal definitions of danger. If true, this would lead the assessment tool to reach conclusions that in fact understated the risk factors for a group with the given characteristics. See Randy K. Otto, *On the Ability of Mental Health Professionals to “Predict Dangerousness”: A Commentary on Interpretations of the “Dangerousness” Literature*, 18 *LAW & PSYCHOL. REV.* 43, 54–59 (1994).

These disparities between fact-finding on danger and actuarial assessments can, and perhaps will be narrowed to the point where they do not raise any evidentiary concerns. See Erica Beecher-Monas (2001 article).

Finally, actuarial methods differ from fact-finding in design and function. Actuarial analysis seeks to summarize and to explain how a certain group of people have acted in the past, expressed as a differential risk. It advises and informs, but makes no decision about a dispute. By contrast, juridical fact-finding seeks an explicable basis on which to take action that balances strongly held and conflicting individual and societal interests. It must thus justify itself by reference to the individual, and necessarily calls for more than just generalized statements about how other groups have acted in the past.

Despite this lack of a close fit, actuarial assessments have provably greater reliability than clinical predictions when offered for properly limited purposes. The system of statistical assessment can be explained and justified; it has strong internal coherence. Moreover, the weaknesses of assessments are subject to ready testing through the adversarial process, especially with reference to disparities between the actuarial conclusion and the contingent, contextual reality of a defendant. If properly limited, then, the extensive literature validating both the significance of the risk factors, and the probability statements of the actuarial tools, more than satisfy *Daubert's* dispute-oriented reliability concern.³⁸¹

Some courts seem to accept actuarial assessments only by admitting them as part of clinical predictions. The cases neither address the issue squarely nor develop it fully, but they do suggest some trends. The majority view would admit actuarial assessments when integrated into clinical predictions.³⁸² No court has yet addressed the admissibility of an

381. Actuarial assessment becomes unreliable when offered for other purposes. For example, some proponents of assessment tools argue that they produce predictions of probable danger of *this* defendant, and even that an actuarial assessment demonstrates that the defendant has a "propensity for violence." See, e.g., R. Karl Hanson, *What Do We Know About Sex Offender Risk Assessment?*, 4 PUB. POL'Y & L. 50, 67 (1998) ("[B]oth actuarial and mixed clinical / actuarial methods can be expected to reliably identify a small subgroup of offenders with an enduring propensity to reoffend.") (emphasis supplied), cited favorably in *In re Commitment of R.S.*, 773 A.2d 72, 93-94 (N.J. Super. A.D. 2001) (favorably assessing Professor Hanson's views generally). Offered and argued for this purpose, actuarial assessments fall well below *Daubert's* reliability standard: by definition, the methods produce neither predictions for individuals nor descriptions of mental states.

At an extreme, use of risk factors in this way could result in the creation of an evidentiary presumption of danger, in the form of: "If you find [certain risk factors], then you may/must infer danger." This Article does not discuss this potential for new presumptions. But the discussion of commitment process earlier in this Article makes it seem unlikely that courts will substitute legal presumptions for complex fact-finding on danger. See Part I.C *supra* (discussing judicial accommodation of predictive uncertainty in commitment jurisprudence).

382. See, e.g., *In re Detention of Strauss*, 20 P.3d 1022, 1026-27 (Wash. App. 2001) (admitting results from the MnSOST, RRASOR, and the VRAG tests as part of "an overall risk assessment"); *State ex rel Romley v. Fields*, 35 P.3d 82, 89 (Ariz. Ct. App. 2001) (admitting actuarial instruments when offered as part of clinical testimony); *People v. Woods*, No. C037203, 2001 WL 1649216, at *5-*6 (Cal. Ct. App. Dec. 20, 2001) (affirming use of the RRASOR and the Static 99 as part of a clinical assessment); *People v. Hayes*, No. A093285, 2002 WL 462277, at *6 (Cal. Ct. App. Mar. 26, 2002) (af-

actuarial assessment presented outside the context of a clinical prediction. However, one or two opinions have stated the view that actuarial assessments are in fact “better than” clinical predictions.³⁸³ The division in the courts parallels, and perhaps reflects, the division in the scientific community.³⁸⁴

Daubert does not require this linkage between the two differing forms of predictive expertise; it admits both.³⁸⁵ Actuarial assessments sat-

firming use of the Static 99 when used as part of a clinical assessment). *Compare* *Westerheide v. State*, 767 So. 2d 637, 656–57 (Fla. Dist. Ct. App. 2000) (distinguishing between clinical predictions formed without any reliance on actuarial methods, and clinical predictions resting at least in part on such methods; the latter would be subject to *Frye*, while the former would not), *with* *Green v. State*, 826 So. 2d 351, 353–54 (Fla. Dist. Ct. App. 2002) (admission of expert testimony involving the use of actuarial methods as part of clinical predictions held harmless, without ruling on whether admission was error).

383. *Comm'r v. Reese*, No. CIV.A 00-0181-B, 2001 WL 359954, at *8–*10 (Mass. Super. Ct. Apr. 5, 2001) (finding no probable cause to continue detention of a defendant alleged to be sexually violent). The trial court identified three possible methods of predicting the likelihood of sexual reoffense: clinical judgment; statistical analysis; and “guided clinical judgment.” *Id.* at *9. It noted that “statistical risk progression scales . . . have substantially greater predictive accuracy than clinical judgment,” and that “. . . statistics, in general, are better predictors of future sexual dangerousness than clinical judgments.” *Id.* at *9. *See also* *In re* Commitment of R.S., 773 A.2d 72, 93–95 (N.J. Super. Ct. App. Div. 2001) (favorably assessing Professor Hanson’s views); *In re* Registrant C.A., 679 A.2d 1153, 1170 (1996) (affirming use of a statistical methodology in Megan’s Law cases: “the use of actuarial concrete predictors is at least as good, if not in most cases better, in terms of reliability and predictability than clinical interviews”). *See also* *In re* Detention of Holtz, 653 N.W.2d 613, 613, 619 n.5 (Iowa Ct. App. 2002) (citing *Reese*’s appraisal of statistical methods of prediction without comment). *See also* *Barefoot v. Estelle*, 463 U.S. 880, 922 n.4 (1983) (Blackmun, J., dissenting) (“[S]tatistical prediction is clearly more reliable than clinical prediction.”).

Taken to an extreme, this view might not only result in separate admission of actuarial assessments, but also reverse the court’s historical tolerance of clinical predictions, perhaps by requiring all clinical predictions to rely on statistical methodologies. *See also* Erica Beecher-Monas & Edgar Garcia-Rill, *Danger at the Edge of Chaos: Predicting Violent Behavior in a Post-Daubert World*, 24 *CARDOZO L. REV.* 1845, 1897–1900 (2003) (arguing that *Daubert* requires the exclusion of clinical predictions, but that actuarial instruments may improve judgments of juries in death penalty sentencing proceedings).

384. *See supra* Part I.B (describing the division in the research community between advocates of clinical, guided clinical and pure actuarial approaches). The parallel is not precise: no court has accepted pure actuarial assessments, and a faint trend exists towards allowing them *only* when coupled with clinical assessments, a trend which appears to reflect a consensus in the treatment community; the scientific research community appears more severely divided.

See also JOHN MONAHAN ET AL., *RETHINKING RISK ASSESSMENT* 133–34 (2002) (analogizing prediction of violence to predicting the weather). This analogy notes that weather forecasters rely heavily on carefully collected meteorological data, processed by sophisticated programming, and presented to local meteorologists for use; the local observer preserves the ability to alter the prediction “if he or she looked out the window and saw threatening clouds approaching.” More affirmatively, the study notes empirical studies that show that “clinical” methods improve “actuarial” weather predictions: “[t]he clinically revised predictions of temperature and precipitation are consistently more valid than the unrevised actuarial ones.” *Id.* at 134.

385. Again, this does not mean a *per se* rule of admissibility. Qualifying the expert; monitoring the application of the actuarial tool to the defendant; challenging mismatches of risk factors; or noting the presence of other factors not accounted for by the relevant tools: all serve as bases for challenge. Trial

isfy *Daubert's* twin standards of reliability and fit, albeit with a different dynamic than with clinical predictions.³⁸⁶ The concerns about fit may help explain the considerably greater lengths to which courts have had to go in admitting actuarial assessments, at least compared to clinical opinions. These concerns will most probably also lead courts to accept actuarial assessments primarily as part of clinical assessments for the foreseeable future.

CONCLUSION

This Article has described a new model for assessing expertise, extracting it from the puzzling receptivity of courts to notoriously unreliable predictive testimony about future dangerousness. This model requires courts to assess the validity of expertise in other disciplines, and to understand the degree of rigor imposed by that discipline on a given opinion. But the model reformulates *Daubert's* reliability concern as one focused on the demands of judicial fact-finding. The model requires courts to fit the opinion to the case at hand, which entails a richer inquiry than a mere search for bare relevance. Fit surely includes the four considerations described in this Article—integration, incremental value, prevalence, and similarity of inference—but others may exist as well.

So described, the model has potential significance in three different ways. For mental health law and practice, this Article concludes that courts have consistently gotten it right; it aligns current doctrine with decades of consistent judicial opinion. Predictive testimony should be admissible in civil commitment cases under *Daubert*. Does this basic conclusion take the mental health profession off the hook? The answer is “no,” not least because rejecting a *per se* rule of exclusion does not require a *per se* rule of admission. Predictions must still receive separate testing in the mechanics of trial. The relevant test is different from scientific validation, but contains its own rigors and ambiguities.³⁸⁷ Nor does

judges should also exclude actuarial assessments if offered as individualized predictions or as descriptions of internal propensities. Indeed, FED. R. EVID. 403 (or an equivalent balancing of probative dangers with probative value) might also lead to exclusion when actuarial assessments come in as a description of a propensity.

386. The leading case to use *Frye* deviated from strict adherence to a “general acceptance” standard and substituted a reliability standard together with a relatively careful limitation of its probative value to the narrower purposes described here. *In re Commitment of R.S.*, 773 A.2d 72, 92–94 (N.J. Super. App. Div. 2001). While significantly more scientific literature accepts actuarial assessments than clinical predictions, courts using *Frye* must still make a choice about which portion and what percentage of the expert community to use when gauging “general acceptance.” (The same problem appears for predictions which combine clinical and actuarial methods.). See *supra* Part I.A.2 (critiquing the judicial use of *Frye*) and note 370 (recommending exemption from *Frye* for clinical predictions).

387. See *supra* Parts II.C.1–2 (discussing grounds for successful challenging of particular predictive opinions, both clinical and actuarial).

the rule endorsed here endanger future improvements in predicting dangerousness. Mental health professionals have pragmatic pressures within their own discipline that drive the need for improvement. They are unlikely to accept judicial standards for their practices, which encompass different concerns and purposes than dispute resolution.

The dynamic model proposed here also raises useful questions, both for theoretical and empirical researchers. Exactly what is the relationship between fit and reliability? Is it possible to map a coherent picture of that dynamic? In the case of predictive expertise, the two concerns act in inverse proportion: The strength of the fit overcomes weaknesses in the reliability of predictive testimony, and vice versa. But that may not be true in all cases, and may in fact reflect unique features of predictive testimony as well as of civil commitments. The analysis set forth here, including its inquiry into constitutional baselines, burdens of proof, standards of sufficiency, patterns of evidence, and definitions of central facts, may help guide research about the limits of expert testimony in other areas of law.

This model also has intensely practical consequences for the federal courts, and for those states that have made the switch to *Daubert*.³⁸⁸ Trial judges can require proponents of unfamiliar expertise to show not only the degree of rigor a particular field requires, but also the utility of that standard for the particular category of case. How closely will the inferences drawn by the expert parallel those required of the fact-finder? Have courts already accommodated the opinion within the structure and procedure of the related law? Into just what context will the expert speak, with regard to substantive rules, burdens of proof, standards of sufficiency, and the mix of other evidence? How complex or intractable is the fact the expert opinion addresses, and how does the expertise add to fact-finding on that point? As these questions become settled for a given type of opinion in a given type of case, the number of hearings, and the disparity in results should abate, both at trial and at the appellate level.

The model described here reinforces a more fundamental point. The *Daubert* cases deal not solely with science, nor even with the reliability of expertise standing alone. Rather, these cases focus on how new ideas are integrated into dispute resolution processes and (more specifically) on how to use advances in empirical and theoretical inquiry to help satisfy

388. The discussion of reliability and fit here may also be of use in *Frye* jurisdictions, although it seems doubtful. This Article suggests that the *Frye* standard should have difficulty with a highly disputed opinion that nonetheless integrates well with the demands of dispute resolution in a given case. Despite that suggestion, it seems unlikely that any state now adhering to *Frye* will revise that decision in light of this Article.

the judicial imperative to decide cases. The cases set the terms on which fact-finders borrow from other disciplines; but, in so doing, they require fact-finders (and their judicial shepherds) to retain the discipline and pragmatic judgment acquired from years of resolving previous disputes. Judicial decision-makers must ask not only whether new knowledge can be justified *in its own terms*, but also whether, when, and how new knowledge has a role to play in advancing the just and expedient resolution of conflict.

In short, predicting danger does require expertise, but not primarily mental health expertise, even in its newer, more quantified form. Rather, courts must bend the separate, but distinctly similar disciplines of mental assessment and behavioral prediction to the service of judicial expertise. Authoritative prediction of danger remains a judicial task and a judicial discipline. Courts and advocates should search other disciplines for help with this vital task. They should accept that help, even if less than perfect, so long as it fits the needs of the case.