

Winter 2003

Peer Dialogue: The Quagmire of Scientific Expert Testimony: Crumping the Supreme Court's Style

Paul R. Rice

Follow this and additional works at: <https://scholarship.law.missouri.edu/mlr>



Part of the [Law Commons](#)

Recommended Citation

Paul R. Rice, *Peer Dialogue: The Quagmire of Scientific Expert Testimony: Crumping the Supreme Court's Style*, 68 Mo. L. REV. (2003)

Available at: <https://scholarship.law.missouri.edu/mlr/vol68/iss1/8>

This Article is brought to you for free and open access by the Law Journals at University of Missouri School of Law Scholarship Repository. It has been accepted for inclusion in Missouri Law Review by an authorized editor of University of Missouri School of Law Scholarship Repository. For more information, please contact bassettcw@missouri.edu.

Peer Dialogue: The Quagmire of Scientific Expert Testimony: Crumping the Supreme Court's Style

*Paul R. Rice**

I. INTRODUCTION

In *The Trouble with Daubert-Kumho: Reconsidering the Supreme Court's Philosophy of Science*,¹ Professor David Crump may have successfully destroyed the intellectual facade constructed by the Supreme Court in its unfortunate decisions in *Daubert v. Merrell Dow Pharmaceuticals, Inc.*² and *Kumho Tire Co. v. Carmichael*.³ He has methodically outlined the parameters of scientific endeavor, explaining the importance of each method of research in a manner that is understandable to the scientifically challenged, and then persuasively demonstrated that the incomplete nature of the Supreme Court's doctrinal adherence to quantitative theory cannot rationally form the sole definition of "scientific knowledge."⁴

As Justices have performed poorly when they have taken on the mantle of historian,⁵ Professor Crump demonstrates that judges fare no better when

* Mr. Rice is at professor at the American University Washington College of Law and Director of its Evidence Project, www.wcl.american.edu/pub/journals/evidence/. He is the author of the casebook *EVIDENCE: COMMON LAW AND FEDERAL RULES OF EVIDENCE* (4th ed. 2001), the treatises *ATTORNEY-CLIENT PRIVILEGE IN THE UNITED STATES* (2d ed. 1999) and *ATTORNEY-CLIENT PRIVILEGE: STATE LAW* (2002) and a recent book entitled *BEST KEPT SECRETS OF EVIDENCE LAW: 101 PRINCIPLES, PRACTICES, AND PITFALLS* (2001).

1. 68 MO. L. REV. 1 (2003) [hereafter cited as Crump].

2. 509 U.S. 579 (1993).

3. 526 U.S. 137 (1999). Because the basic principles announced in *Daubert* for the application of Federal Rule of Evidence 702 to scientific evidence are also applicable to technological evidence, references to the *Daubert* ruling are generally intended to be applicable to the *Kumho Tire* decision as well.

4. In the *Daubert* opinion, the Court defined scientific knowledge in terms of "whether it can be (and has been) tested. 'Scientific methodology today is based on generating hypotheses and testing them to see if they can be falsified; indeed, this methodology is what distinguishes science from other fields of human inquiry.'" *Daubert*, 509 U.S. at 593 (quoting ERIC D. GREEN & CHARLES R. NESSON, *PROBLEMS, CASES, AND MATERIALS ON EVIDENCE* 645 (1983)).

5. See LEONARD W. LEVY, *ORIGINAL INTENT AND THE FRAMERS' CONSTITUTION* 300-13 (1988) (stating bluntly that "the Court has flunked history" and that the "judges exploit history by making . . . it yield results that are not historically founded"); Alfred H. Kelly, *Clio and the Court: An Illicit Love Affair*, 1965 SUP. CT. REV. 119, 122-25 (1965) (commenting on how Supreme Court Justices practice "law-office history," a

pretending to be scientists. Indeed, in Professor Crump's colorful language, under the deficiencies of the Court's dysfunctional test, courts seem to have been encouraged to "eat the television set and watch the orange."⁶

Professor Crump suggests that the decision on the admissibility of scientific expert testimony is really nothing more than a logical relevance decision, balanced against the potential for unfair prejudice. This is a proposition with which no one would disagree.⁷ In the *Daubert* opinion, the Court acknowledged that relevance "provides the baseline" for the admissibility of scientific evidence,⁸ but it construed Rule 702 as placing other limitations on the admissibility of scientific evidence. Rule 702 directs that scientific expert testimony is admissible if the witness's "scientific . . . knowledge will assist the trier of fact to understand the evidence or to determine a fact in issue."⁹ From this language the Court defines "science" and "knowledge" as ideas that are "ground[ed] in the methods and procedures of science"¹⁰ and known to be true or inferred from facts known to be true, and accepted as true on good grounds.¹¹

"selection of data favorable to the position being advanced without regard to or concern for contradictory data or proper evaluation of the relevance of the data proffered"); Edward C. Walterscheid, *The Use and Abuse of History: The Supreme Court's Interpretation of Thomas Jefferson's Influence on the Patent Law*, 39 IDEA 195, 197 (1999) (demonstrating, through an examination of Thomas Jefferson's influence on the field of patent law, how the Supreme Court's use of history is often "incomplete, over selective, misleading, and biased").

6. See Crump, *supra* note 1, at 41.

7. This is the balancing approach that has been advocated by Professor McCormick in EVIDENCE 363-64 (1954) and Professor John W. Strong in *Questions Affecting the Admissibility of Scientific Evidence*, 1970 U. ILL. L. FORUM 1 (1970).

8. "We interpret the legislatively enacted Federal Rules of Evidence as we would any statute. . . . Rule 402 [which makes all relevant evidence admissible unless otherwise excluded by a specific rule] provides the baseline." *Daubert*, 509 U.S. at 587 (citation omitted).

9. FED. R. EVID. 702. When the *Daubert* opinion was written, Rule 702 provided that "[i]f scientific, technical, or other specialized knowledge will assist the trier of fact to understand the evidence or to determine a fact in issue, a witness qualified as an expert by knowledge, skill, experience, training, or education, may testify thereto in the form of an opinion or otherwise." *Id.*

After the *Daubert* decision, the Advisory Committee on the Federal Rules of Evidence added the following clause to the end of the rule: "if (1) the testimony is sufficiently based upon reliable facts or data, (2) the testimony is the product of reliable principles and methods, and (3) the witness has applied the principles and methods reliably to the facts of the case." *Id.* This was intended to be a codification of the principles outlined in the *Daubert* opinion.

10. *Daubert*, 509 U.S. at 590.

11. While the Court's reasoning is logical, it appears to have been constructed from whole cloth. When Rule 702 was promulgated, it was designed to accomplish two objectives. First, to lower the standard from "necessity" to simple relevance. Second,

Therefore, as the Court further explains, “in order to qualify as ‘scientific knowledge,’ an inference or assertion must be derived by the scientific method.”¹² This ensures that the science is valid, therefore meeting evidentiary standards of reliability. This, the Court insists, requires that the testimony “be supported by appropriate validation, *i.e.*, ‘good grounds,’ based on what is known.”¹³ Later the Court explains that this ordinarily requires that the principle or methodology has been appropriately validated through testing, independent peer review, assessment of the rate of error, and ultimately, whether it is generally accepted in the relevant scientific community.¹⁴

to codify the broad range of means by which the necessary expertise can be acquired.

The language cited by the Court was not intended as an indirect means of establishing a substitute for the *Frye* “general acceptance” test that regulated only the use of novel scientific evidence. Despite the increasing concerns about the *Frye* test, there is not a single word in either the rule or its legislative history to even suggest that the purpose of Rule 702 was to refine standards for the admissibility of every type of science, all technology, and any other conceivable areas of speciality, simply by the use of the word “knowledge.” Both the structure and language of the rule reinforce the conclusion that the rule was only codifying the status quo. As written, the clause “scientific, technical or other specialized knowledge” does not appear to be a restriction on the type of knowledge that must be possessed, as much as a universal way of expressing the range of areas in which witnesses can have special expertise. In contrast to the Court’s strained interpretation and use of this language, it is far more probable that the clause was designed only to regulate *how* one can qualify as an expert (from “knowledge, skill, experience, training, or education”), and *when* expert testimony is admissible (when it “will assist the trier of fact to understand the evidence or determine a fact in issue”).

In reality, the Court has done little more than engage in common law rule-making—a right that Congress took away from judges when the Federal Rules of Evidence were codified. When Supreme Court Justices want to change the rules, they should be required go through the cumbersome and unresponsive Advisory Committee process that the rest of us are compelled to tolerate. Perhaps then the flawed process would be changed. For a discussion of the inadequacies of the Advisory Committee process, and particular problems with the Federal Rules of Evidence Advisory Committee, see generally Paul R. Rice, *Advisory Committee on the Federal Rules of Evidence: Tending to the Past and Pretending for the Future?*, 53 HASTINGS L.J. 817 (2002); Paul R. Rice & Neils-Erik William Delker, *Federal Rules of Evidence Advisory Committee: A Short History of Too Little Consequence*, 191 F.R.D. 678 (2000); Paul R. Rice, *Rule Changes Raise Questions*, NAT’L L.J., Nov. 13, 2000, at A20; Paul R. Rice, *Expert Overhaul Needed: Outdated Federal Rules Require More Than Judicial Tinkering*, LEGAL TIMES, Jan. 31, 2000, at S1; Paul R. Rice, *Bring on the Reformers: Evidence Code Cries Out for More Than Cautious Tinkering*, LEGAL TIMES, October 19, 1998, at 28; Paul R. Rice, “Good Enough” Is Not Good Enough for Evidence Rules, letter to the editor responding to article by Professor Capra, NAT’L L.J., Dec. 2, 1998.

12. *Daubert*, 509 U.S. at 590.

13. *Id.*

14. *Id.* at 593-94. Theoretically *Daubert* converted the exclusive “general acceptance” standard of *Frye* to a factor to be considered by the presiding judge.

Professor Crump parts company with the Court when it imposes what he believes to be an overly restrictive definition of “scientific knowledge” on the scientific validation and relevancy determination. Because Rule 702 requires judges to decide questions of admissibility only on whether proffered opinion evidence would “assist” jurors in understanding the evidence that has been presented or determining a fact in issue, he argues that the Supreme Court’s limitation of the definition of “science” to the quantitative method is too restrictive. His argument is persuasive, and particularly compelling because the *Daubert* opinion expanded the judicial screening responsibility from novel scientific principles and methodologies (the limit of the previously controlling *Frye* “general acceptance” test) to all expert scientific testimony.¹⁵

When the previously controlling *Frye* standard was followed, after general acceptance was demonstrated, and the science was no longer considered novel, judges determined admissibility by the Crump standard, “logical relevance balanced against potential unfair prejudice.” Therefore, by recommending that admissibility be determined on purely logical relevance grounds, Professor Crump is actually advocating nothing new. He simply wants to get back to basics—a giant step backwards, if you will, to the pre-*Frye* way of determining the admissibility of all evidence. The only difference, which makes his proposal more compelling today, is the fact that the standard for the admissibility of all expert testimony has been lowered. Now the testimony need only be helpful.¹⁶

Aside from his critique of the majority’s definition of “science,” what Professor Crump seems to be arguing is not much different from what was argued by Chief Justice Rehnquist in his prescient dissent. He thought that

15. *Frye v. United States*, 293 F. 1013 (D.C. Cir. 1923). The District of Columbia Circuit imposed a “general acceptance in the relevant scientific community” standard for the admissibility of all *novel* scientific evidence. *Id.* at 1014. Subsequently, the *Frye* standard was adopted throughout the country, and remained the controlling standard for the admissibility of novel scientific evidence in literally every state and federal jurisdiction until the *Daubert* decision. In *Daubert*, the Court stated:

Although the *Frye* decision itself focused exclusively on “novel” scientific techniques, we do not read the requirements of Rule 702 to apply specifically or exclusively to unconventional evidence. Of course, well-established propositions are less likely to be challenged than those that are novel, and they are more handily defended. Indeed, theories that are so firmly established as to have attained the status of scientific law, such as the laws of thermodynamics, properly are subject to judicial notice under Federal Rule of Evidence 201.

Daubert, 509 U.S. at 592 n.11.

16. Under Rule 702 of the Federal Rules of Evidence, testimony is admissible if it would “assist the trier of facts to understand the evidence.” FED. R. EVID. 702. By contrast, the common law, under which *Frye* was created, required the expert testimony to be “necessary” to the jurors’ understanding of the evidence or their ability to accurately resolve a factual issue.

general observations about the parameters of “science” and “knowledge” were not necessary in order to decide the questions before the Court, and cautioned that when judges delve into “matters far afield from the expertise of judges,” their “reach can so easily exceed [their] grasp.” Subsequently, his admonitions were borne out in the utter confusion that has resulted from judges attempting to use the four illustrative factors as “technical hurdles” or “tests to be rigorously surmounted” rather than flexible criteria.¹⁷ Neither Chief Justice Rehnquist nor Professor Crump wants to provide details on how reliability is to be assessed—leaving that determination to a case-by-case assessment based on underlying logic of relevance and unfair prejudice.

As noted above, the “scientific knowledge” overlay was created by the majority in *Daubert* because of the language of Rule 702. However, in the process of wrapping themselves in the cloak of “scientific knowledge,” simply because the nature of evidence under discussion was scientific, the Justices, as predicted by Chief Justice Rehnquist, lost their bearings. Beyond their strangled definition of “science,” which Professor Crump so effectively debunks, the Court also lost any semblance of common sense.

The Justices went off on their “scientific knowledge” tangent without *full* recognition that science is not—and should not be—used in the courtroom in the same way that it is used in the laboratory.¹⁸ Scientific truths reflected in

17. See Crump, *supra* note 1, at 17.

18. While the Court did recognize that the “quest for truth in the courtroom and the quest for truth in the laboratory” are different, it did so only to justify stringent standards of scientific verification and reliability. *Daubert*, 509 U.S. at 596-97.

Petitioners . . . suggest that recognition of a screening role for the judge that allows for the exclusion of “invalid” evidence will sanction a stifling and repressive scientific orthodoxy and will be inimical to the search for truth It is true that open debate is an essential part of both legal and scientific analyses. Yet there are important differences between the quest for truth in the courtroom and the quest for truth in the laboratory. Scientific conclusions are subject to perpetual revision. Law, on the other hand, must resolve disputes finally and quickly. The scientific project is advanced by broad and wide-ranging consideration of a multitude of hypotheses, for those that are incorrect will eventually be shown to be so, and that in itself is an advance. Conjectures that are probably wrong are of little use, however, in the project of reaching a quick, final, and binding legal judgment—often of great consequence—about a particular set of events in the past. We recognize that, in practice, a gatekeeping role of the judge, no matter how flexible, inevitably on occasion will prevent the jury from learning of authentic insights and innovations. That, nevertheless, is the balance that is struck by Rules of Evidence designed not for the exhaustive search for cosmic understanding but for the particularized resolution of legal disputes.

Id. The Court only recognized the difference in the nature of the search. The more important distinction that the Court failed to recognize was the fact that the laboratory

scientific principles are used in scientific research to discover further scientific truths. Through these discovered truths, additional research is done and more truths are discovered. Consequently, depending on the context, science will insist on very high levels of accuracy and reliability (confidence levels). In fact, most quantitative sciences impose something in the neighborhood of a 95%

is searching for further truths from the scientific truth reflected in the principle being used, whereas the courtroom is simply trying to use the principle, in conjunction with the remainder of the evidence in the record, to resolve a dispute by a preponderance of the evidence. Nothing remotely resembling a scientific truth is being sought.

confidence level.¹⁹ In evidence parlance, this might be the equivalent of

19. This was illustrated by Professor Allen relative to the analysis in *Daubert* of Bendectin and its causal relationship to the plaintiff's injuries.

The plaintiffs in *Daubert* suffer from limb-reduction birth defects allegedly caused by their mothers' use of Bendectin during pregnancy. The affirmative scientific support for their assertion came solely from animal studies and chemical structural analyses of the chemical. A number of epidemiological studies of the effect of Bendectin have been done, but none of them found statistically significant correlations between the use of Bendectin and birth defects. Undeterred, the plaintiffs offered a reanalysis of the data of the epidemiological studies by well credentialed experts. In essence, their argument was that the structure of standard epidemiological studies favors reducing Type I errors at the expense of making more Type II errors. A Type I error is an erroneous finding of a causal relationship where there is none between the drug and birth defects, and a Type II error is an erroneous finding of no causation. Each of the studies used the standard confidence level of 95%, which means that the chances of erroneously rejecting the null hypothesis (and thus erroneously concluding that there is a causal link) is less than 5%. But accepting such a standard increases dramatically the probability that the conclusion of no causation is false. For example, by one calculation, even if Bendectin caused a doubling in the rate of birth defects, the probability that the published studies would have yielded a statistically significant outcome is less than 20%.

The skewing of mistakes against erroneous findings of causation may seem odd to lawyers, but is plausible in the context of science. In tests such as those involving Bendectin, the issue generally is whether the drug causes a particular result, such as increasing the chances of birth defects. Causation of this sort is virtually always highly complex, and thus essentially random outcomes occur relatively frequently making it difficult to sort out the precise effect of a single variable. This is the standard difficulty of arguing from correlation to causation. In order to protect against an erroneous finding of causation, most quantitative sciences have set high standards for their experiments, such as the 95% confidence level. That data do not meet this standard, and thus that the null hypothesis is not rejected, does not mean there is no causal link; it means that this experiment did not find one, judged by the 95% confidence level. Other experiments may come out differently.

Ronald J. Allen, *Expertise and the Daubert Decision*, 84 J. CRIM. L. & CRIM. 1157, 1164-65 (1994) (footnotes omitted); see also Michael D. Green, *Expert Witnesses and Sufficiency of Evidence in Toxic Substances Litigation: The Legacy of Agent Orange and Bendectin Litigation*, 86 N. U. L. REV. 643 (1992).

Professor Allen goes on to explain that this approach to experimentation is driven by the emphasis on Karl Popper's theory on falsifiability (which the Court in *Daubert* adopted in its definition of science knowledge), rather than attempting to verify hypotheses. The focus of falsifiability is appropriate for science because the concern of science:

is to get a good approximation of reality over the long run. From that perspective, erroneous findings that some proposition is supported are worse

establishing reliability, and therefore admissibility, beyond a reasonable doubt.

Such a burden of persuasion has never before been sanctioned for determining the admissibility of evidence, even the voluntariness of confessions in a criminal proceeding.²⁰ Indeed, even when science is used in commercial endeavors, the level of reliability insisted upon is directly linked to the consequences of error to life and property. To avoid liability, for example, many industries probably insist on a much higher standard of reliability than would be necessary in judicial proceedings where legal responsibility is being determined by a preponderance of the evidence. While I do not know this to be true, I would expect that the acceptable error rates of the sciences and technologies involved in creating grades of concrete and their substitutes are probably much higher than those involved in the manufacturing of synthetic heart valves. And even the higher error rates that are generally acceptable for concrete or its substitutes will likely be much lower when the concrete-like substances are used in the construction of bridges and overpasses, where defects are more likely to cause a loss of life.

Because science employed in the courtroom need only “assist” jurors in coming to a decision by only a preponderance of the evidence, the tolerable limits of unreliability are much higher. Indeed, the standard for determining the admissibility of all evidence is preponderance—the judge need only be convinced, by a preponderance of the evidence, that the elements of the applicable evidence rules governing admissibility have been established. The problem that remains unaddressed is how does the preponderance standard for admissibility relate to scientific evidence after the *Daubert* opinion, in which judges have been instructed to weigh the same factors that the scientific community would weigh, when assessing scientific principles under consideration?

than conclusions that no significant results were obtained. Erroneous findings of support would tend to generate belief in the truth of the proposition, which in turn may infect the research program, whereas erroneous findings of no support will primarily generate further efforts at falsification.

Allen, *supra*, at 1166.

As Professor Allen explains, herein lies the problem in using science in the courtroom. The law admits evidence under a preponderance standard and decides disputes under the same standard. The law imposes a decisional rule that is:

designed to generate about the same number of erroneous verdicts for plaintiffs as for defendants. . . . The law assumes that an erroneous rejection of the null hypothesis that defendant is not liable (thus erroneously finding for the plaintiff) is equivalent to an erroneous rejection of the hypothesis that defendant is liable (thus finding for the defendant).

Id. In other words, the law historically has attempted to scatter the errors, while science attempts to “reduce the number of false findings.” *Id.*

20. See *Lego v. Twomey*, 404 U.S. 477, 489 (1972).

In *Daubert* the Court instructed judges to consider four factors: (1) whether the principles and methodologies under consideration have been tested; (2) the results of peer review in professional publications; (3) the potential rate of error; and (4) whether they have achieved general acceptance in the relevant scientific community.²¹ Does this perpetuate the same high standards of reliability insisted upon by the scientific community?

The Justices might argue that, rather than imposing science's high standards of validity on the trial judge, the *Daubert* decision was doing just the opposite. Because the *Frye* general acceptance test was being abandoned in *Daubert* for a discretionary test to be applied by the presiding judge, the Justices might argue that their decision relaxed the restrictive scientific standards, rather than perpetuating them, thereby permitting far more evidence to flow into the trial.²² While this argument may be theoretically accurate, in practical reality it has little substance.

II. THE GOSSAMER FACTORS OF *DAUBERT*

While the *Daubert* decision offered four nonexclusive factors that should be considered by the presiding judge in screening scientific evidence, it is unclear how this delineation will help judges move away from reliance on people in the relevant science (a *Frye* type approach) when judges generally do not have the expertise to evaluate and apply them.²³

21. *Daubert*, 509 U.S. at 593-94.

22. In discussing whether the *Frye* "general acceptance" test had been assimilated into the Rules, the Court indicated that both the Rules and the Court's decision were liberalizing the admissibility of scientific evidence.

The drafting history makes no mention of *Frye*, and a rigid "general acceptance" requirement would be at odds with the "liberal thrust" of the Federal Rules and their "general approach of relaxing the traditional barriers to 'opinion' testimony." . . . Given the Rules' permissive backdrop and their inclusion of a specific rule on expert testimony that does not mention "general acceptance," the assertion that the Rules somehow assimilated *Frye* is unconvincing. . . . That austere standard, absent from, and incompatible with, the Federal Rules of Evidence, should not be applied in federal trials.

Id. at 588-89 (citations omitted).

23. Indeed, the difficulties that trial judges will encounter in fulfilling their "gatekeeping" role was noted by Chief Justice Rehnquist in his dissenting opinion in *Daubert*:

I defer to no one in my confidence in federal judges; but I am at a loss to know what is meant when it is said that the scientific status of a theory depends on its "falsifiability," and I suspect some of them will be, too. I do not doubt that Rule 702 confides to the judge some gatekeeping responsibility in deciding questions of the admissibility of proffered expert testimony. But I do not think it imposes on them either the obligation or the authority to

For example, most judges employing these factors will have no independent basis for evaluating the tests that were employed to assess the validity and reliability of the scientific principles and methodologies in question. Perhaps minor variations in protocols could produce radically different results. They will not have the expertise to determine an acceptable error rate, or to measure the adequacy of the controls that were in place to insure that the principles and methodologies that were employed produced accurate results. If people in relevant scientific communities had conducted peer reviews and reported negatively, it is highly improbable that judges would ignore those reviews based on their own assessment of the science.

By the time most courts begin considering the fourth factor, general acceptance, the answer is a forgone conclusion because the relevant community has already been consulted repeatedly on the first three factors. As a consequence, more often than not, in its practical application, *Daubert* will be little more than *Frye* in drag.²⁴ With neither the time nor the ability to perform this task, judges will simply “retool” *Frye* by anointing a single expert and substituting that expert’s opinions for the relevant scientific or technological

become amateur scientists in order to perform that role. I think the Court would be far better advised in this case to decide only the questions presented, and to leave the further development of this important area of law to future cases.

Id. at 600-01 (Rehnquist, J., concurring in part and dissenting in part). Indeed, the warnings of the Chief Justice were prescient. The factors delineated by the majority have become immutable hurdles in preliminary screening efforts.

24. PAUL R. RICE, BEST-KEPT SECRETS OF EVIDENCE LAW: 101 PRINCIPLES, PRACTICES, AND PITFALLS 177 (2001); PAUL R. RICE, EVIDENCE: COMMON LAW AND FEDERAL RULES OF EVIDENCE § 8.03[B][1] (4th ed. 2001). This conclusion may be supported by a recent study by the Rand Institute for Civil Justice, Lloyd Dixon & Brian Gill, *Changes in the Standards for Admitting Expert Evidence in Federal Civil Cases Since the Daubert Decision* (2001). In Section 5.3 the authors concluded:

The role that general acceptance should play in a judge’s decision to exclude expert evidence was at the heart of the *Daubert* decision, and we here evaluate how this role has changed over time. It is not obvious how *Daubert* might have affected the importance placed on general acceptance. . . . [T]here is uncertainty a priori about the emphasis judges would place on general acceptance after *Daubert*. On the one hand, even though general acceptance is one of the *Daubert* factors, an increasing focus on direct examination of the method underlying the evidence may limit general acceptance’s importance. On the other, general acceptance may be used by judges as a convenient indicator of reliability and be critical to admissibility.

Id. § 5.3. Discussing the results of the study, the authors concluded that “evidence found not to be generally accepted was, if anything, more likely to be found unreliable after *Daubert* than before.” *Id.* In assessing the results of this study, it is important to remember that admissibility of scientific evidence that was not novel did not require the judge to assess the degree to which it was generally accepted in the relevant science.

communities. The only differences now are that: (1) the judge's definition of the relevant scientific community may not be as transparent, (2) *Daubert* is now much bulkier dressed in this new garb because judges are applying it to *all scientific evidence*, not just to novel scientific principles and technologies, and (3) judges are now laying out the technical criticisms and concerns of the expert as if they were their own—in much the same way that some do with the logic of the opinions drafted by their law clerks.

In an article by Justice Stephen Breyer,²⁵ he acknowledged the degree to which judges rely on guidance from the scientific community. He discussed the many ways that judges have managed this gatekeeping function. He noted that some have appointed specially trained law clerks and special masters. Others have appointed independent experts, in addition to those presented by the parties. Still others have appointed a neutral panel of experts. All of these case-management techniques emphasize the continued importance of the opinions of those in the relevant sciences. The degree to which experts continue to be crucial to the screening function accentuates the urgency of this unaddressed problem of the level of reliability that must be demonstrated and the burden of persuasion that must be satisfied.

For two reasons, the *Daubert* decision makes this problem more urgent. First, as previously noted, the admissibility decision is being made for *all* scientific and technological evidence. Second, even if preponderance were generally acknowledged as the standard of reliability, when judges constantly have to rely upon experts, who employ a much higher standard of reliability, that higher standard is inevitably influencing, if not controlling, admissibility decisions.²⁶

25. Stephen S. Breyer, *Science in the Courtroom*, ISSUES IN SCI. & TECH., July 1, 2000.

26. The Federal Rules of Evidence Advisory Committee has summarily rejected a proposal by the Washington College of Law Evidence Project to incorporate an explicit level of reliability in the expert opinion rules. See THE EVIDENCE PROJECT (Thomas C. Goldstein ed.), at <http://www.wcl.american.edu/pub/journals/evidence/>. Advisory Committee members expressed the belief that Rule 104 establishes a preponderance of the evidence standard for admissibility that is applicable to all rules. See Rice & Delker, *supra* note 11, at 710-11.

The Committee's decision was unfortunate for several reasons. First, there is no guarantee that the Supreme Court will concur in the conclusion that Rule 104 establishes a preponderance standard of reliability for scientific evidence since the sciences employ a much higher level of reliability—something that is often more akin to “beyond a reasonable doubt.” Then we must confront the problem of how to coordinate a preponderance standard with a “beyond a reasonable doubt” standard.

A decade ago a similar problem arose with regard to the relationship of Rules 403 and 609. Rule 403 generally regulates the standard for the exclusion of relevant evidence because of potential unfair prejudice. Rule 609 controls the admissibility of prior convictions for impeachment purposes. The Advisory Committee withdrew proposed

Do trial judges understand that negative assessments of principles, methodologies, and applications by people in a particular science or technology need to be discounted because they are based on elevated expectations of reliability?²⁷ Do they have an appreciation of the degree to which those opinions must be discounted? We don't know the answers to either of these questions, because the Advisory Committee has conducted no studies of these issues.²⁸ If judges do not appreciate the limits of scientific and technological assessments, relevant scientific and technological testimony that could be helpful to the finder of fact is being excluded. This cannot be avoided if the standard by which validity and reliability is being measured is not explicitly delineated and the judges are instructed on how to integrate the lesser standard with the higher standards being employed by those advising them.²⁹

revisions to Rule 609 that explicitly referred to Rule 403, on the belief that the obvious relationship would be confirmed by the Supreme Court in *Green v. Bock Laundry Machine Co.*, 490 U.S. 504 (1989). To their surprise, however, the Court did not find what was so obvious to the Committee. Consequently, a specific reference to Rule 403 subsequently had to be incorporated in Rule 609.

Second, the standard by which scientific testimony is screened is influenced by the disproportionate impact that such evidence will likely have on lay jurors. As a consequence, it might be expected that many judges believe that a higher standard than preponderance must be employed when determining the reliability of such evidence. This, of course, is not known because no studies have been conducted by the Advisory Committee, and such information is generally not volunteered in the written opinions of judges—possibly because they understand that the less they say about subjects about which there is uncertainty, the less likely it is that they will be reversed on appeal. One of the early lessons that trial judges learn is that in silence, appellate courts assume that trial courts applied appropriate standards.

27. Clearly, those who conducted the study for the Rand Institute for Civil Justice, did not understand this distinction. In Section 8.2, where next steps were discussed, they approached the subject of assessing how well judges are performing the gatekeeping functions. They noted that:

One promising approach is to assemble panels of experts to evaluate expert evidence in a sample of cases. The experts would evaluate the reliability of both admitted evidence and excluded evidence so as to understand how well the screening process is working. The results of such an evaluation could be the basis of a report card on how well judges are performing the gatekeeping function.

Dixon & Gill, *supra* note 24, § 8.2. This of course, is inaccurate. The result would only measure how well judges play scientist and make scientific decisions on reliability. Since that is not the issue, the proposed study would measure nothing particularly relevant to judges' proper function in judicial proceedings.

28. Apparently, the Advisory Committee has not been as enamored with quantitative theory as the Supreme Court, since it has seldom used its resources to survey judicial attitudes and practices.

29. Theoretically, scientific principles that might be considered "junk science" by

Under *Frye*, the question of validity and reliability, and therefore of admissibility, was officially delegated to the scientists. They defined what “general acceptance” meant and assessed whether the principle under examination had achieved it. Of course, each scientific discipline had a different reliability standard, and each expert, as a representative of a discipline, had a different perception of “general acceptance.” Nevertheless, their conclusions were accepted at face value by the courts. This, of course, was both the advantage and disadvantage of the *Frye* test. Primarily it was advantageous because it left technical decisions about science to those with scientific expertise. The disadvantage, of course, was that science in the courtroom was always behind the curve of scientific discovery, and relevant evidence was being excluded.

If judges, rather than the scientists, are now expected to make this determination, the nonscientifically trained judges must be given more direction on how reliability is to be measured. Until we clarify the gatekeeper’s screening mechanism, the secret password will remain “general acceptance in the relevant science,” and too many admissibility decisions will be as unpredictable and unfair as commentators claimed they were under *Frye*.³⁰ But again, the situation may be significantly worse, because this unacknowledged standard is now being used as the indirect screening standard for *all scientific evidence*, rather than just that which is novel.

the scientific community might be sufficiently probative to be helpful in a judicial proceeding if used with reservation. For example, a scientific expert might testify, “Scientifically speaking, this principle is junk, and would never be used by us, but it’s good enough for the limited purpose for which you are proposing to use it.”

I broached the possibility of this type of testimony with scientists while speaking on the evidentiary problems in introducing satellite images at a conference entitled “A View from Space: Digital Earth Applications in Environmental Resources Management,” sponsored by Environmental Legal Information Systems, Washington, D.C., Jan. 26, 2001. This suggestion received robust laughter because scientists perceive it as being beyond their expertise to venture an opinion on anything other than scientific uses.

The term “junk science” was made popular by Peter Huber in his book, PETER W. HUBER, GALILEO’S REVENGE: JUNK SCIENCE IN THE COURTROOM (1991). An example of “junk science” was provided by Justice Stevens in *General Electric Co. v. Joiner*, 522 U.S. 136 (1997) (Stevens, J., concurring in part and dissenting in part). In his concurring opinion, Justice Stevens stated that “the testimony of a phrenologist who would purport to prove a defendant’s future dangerousness based on the contours of the defendant’s skull” would be considered “junk science.” *Id.* at 153 n.6. It is not clear how the phrenological testimony in *Joiner* might have utility despite being rejected in the scientific community.

30. Strong, *supra* note 7, at 1. For an informative examination of the advantages and disadvantages of the *Frye* test, see *Symposium on Science and the Rules of Evidence*, 99 F.R.D. 187 (1983).

III. COMMON SENSE VERSUS SCIENTIFIC ACCURACY

What the exclusively scientific assessment ignores is the possibility that some uses of evidence can tolerate far less accuracy. Exploring disputed issues through science is a little like fishing, where the fisherman has the choice of using a range of different strength lines. When trying to catch small fish, the one pound test line should be quite adequate. If, however, catching large game fish is anticipated, usually a line with a much stronger pound test is recommended, although not essential if the equipment is used with some finesse. Similarly, at trials where a litigant has offered independent evidence to prove a particular proposition, and the scientific evidence is being offered to confirm or reinforce a conclusion that otherwise could have been reached based on the evidence heard or seen, the lighter pound test line should be adequate. When the scientific evidence is not the only compelling evidence on the point for which it is offered, there is little reason for judges to exclude it, even though it might be considered unreliable, or even “junk,”³¹ in the scientific community. Because the evidence only needs “to assist” the finder of fact in order to be admissible, every piece of scientific evidence should not have to be capable of landing the “big one” before it is admitted. This is particularly true if the trier of fact is the judge, who, like the fisherman, may be adept at using the more marginal scientific evidence with some finesse.

Perhaps the recent decisions on the admissibility of expert fingerprint comparisons is illustrative.³² Initially concerned about the unknown error rates and lack of accepted standards for making comparisons, Judge Pollak admitted the fingerprints (because all prints are stable and different) as well as the expert testimony on the various points of comparison. He excluded, however, all expert opinions on whether the latent print and the print taken from the defendant were the same. On reconsideration, Pollak reversed himself because the standards of comparison—the number of points of similarity necessary for an identification—are apparently far more flexible than he originally was led to believe. On the issue of the unknown error rate, Pollak was satisfied by the absence of negative evidence—the fact that “there is no evidence that certified FBI fingerprint examiners present erroneous identification testimony, and, as a corollary, that there is no evidence that the rate of error of certified FBI fingerprint examiners is unacceptably high.”³³

31. Even the Supreme Court in the *Daubert* opinion recognized that the jury should be capable of dealing with “shaky but admissible evidence” if it is properly attacked on cross-examination and challenged with opposing testimony. *Daubert v. Merrell Dow Pharm., Inc.*, 509 U.S. 579, 596 (1993).

32. *United States v. Llera Plaza*, 188 F. Supp. 2d 549 (E.D. Pa. 2002).

33. *Id.* at 572.

While the absence of negative evidence could be seen as a questionable way to qualify a science or technology, perhaps it was appropriate here, where the court was addressing a type of evidence that had been used worldwide for a lengthy period of time (thereby providing decades of experience to draw upon), was regarded as reliable by a diverse cross-section of the communities using it, was available for visual examination by lay jurors who could readily understand the evidence and critically evaluate the expert's assessments of the points of comparison, and the evidence was being used in a prosecution in which it simply complemented a larger factual pattern establishing the identity of the defendant (thereby, making a higher potential error rate more acceptable).

IV. THE CONCERN ABOUT OVERVALUATION BY JURORS

Of course, the excuse that has justified the exclusion of relevant and probative scientific evidence (albeit unaccepted in the scientific community) has been the fear that jurors will give it disproportionate value. While this is not an irrational concern, courts seem to have taken it to irrational extremes. While the directive in Rule 403 to balance probative value against unfair prejudice is often mentioned in judicial opinions as a final check on the admissibility of otherwise relevant scientific evidence,³⁴ implicitly this balance is already being made part of the admissibility decision before that point, because relevant and probative evidence is being excluded on the basis of elevated scientific standards of reliability. Therefore, we are left with two questions. The first is whether the Rule 403 standard is implicitly being factored into the initial screen to exclude relevant evidence only when the potential prejudice *substantially outweighs* the probative value. The second is how are judges concluding that jurors cannot understand, and therefore will be overwhelmed and led astray by, particular types of expert testimony?

The unstated potential danger implicitly being infused into the initial screen standard may be exaggerated. Because of this possibility, perhaps, as Professor Crump advocates, courts should be encouraged to make admissibility decisions of scientific evidence only on the basis of logical relevance (Rules 401 and 402) balanced against unfair prejudice that must substantially outweigh value to justify exclusion (Rule 403), leaving the technical *Daubert* factors to the end of

34. See, e.g., *Daubert*, 509 U.S. at 595 ("Finally, Rule 403 permits the exclusion of relevant evidence 'if its probative value is substantially outweighed by the danger of unfair prejudice, confusion of the issues, or misleading the jury' Judge Weinstein has explained: 'Expert evidence can be both powerful and quite misleading because of the difficulty in evaluating it. Because of this risk, the judge in weighing possible prejudice against probative force under Rule 403 of the present rules exercises more control over experts than over lay witnesses.'") (quoting Jack B. Weinstein, *Rule 702 of the Federal Rules of Evidence Is Sound; It Should Not Be Amended*, 138 F.R.D. 631, 632 (1991)).

trial, after all relevant evidence has been heard and the relative significance of the scientific evidence is more apparent. This could be accomplished by tentative rulings when the evidence is initially presented, with a final assessment when the trial has been completed, or, if tentative rulings are not palatable, through motions for judgment at the end of the trial (if there is not sufficient evidence to raise a genuine factual controversy) under Rule 50 of the Federal Rules of Civil Procedure.³⁵

Moreover, rather than assuming that jurors are going to misuse certain types of scientific evidence, and keeping so much from them through highly restrictive screening standards (that are neither well understood nor consistently applied by the judges who use them), it may make more sense to permit all relevant scientific evidence to be heard by the jurors that will not obviously distort their deliberations.³⁶ If *Daubert* was intended to liberalize the admissibility of scientific evidence, in the same way the admissibility of all other evidence has been liberalized under the Federal Rules of Evidence, this may be the only effective way that that can be accomplished, given the restrictive attitudes that have evolved.³⁷

Post-trial procedural mechanisms have a number of advantages. First, they could address the problem of overvaluation by jurors as effectively as pre-

35. If made before the case is submitted to the jury, it would be through the equivalent of the old motion for a directed verdict. If after the jury has returned a verdict, it would be through the old motion for judgment notwithstanding the verdict.

36. Of course, this approach would confront special problems in criminal cases where there must be greater sensitivity to the special vulnerability of the accused, and there is no possibility of a directed verdict or judgment notwithstanding the verdict for the prosecution.

37. In the *Daubert* decision, the majority recognized how post trial motions could serve as an effective check on the admission of marginal scientific evidence:

Respondent expresses apprehension that abandonment of "general acceptance" as the exclusive requirement for admission will result in a "free-for-all" in which befuddled juries are confounded by absurd and irrational pseudoscientific assertions. In this regard respondent seems to us to be overly pessimistic about the capabilities of the jury and of the adversary system generally. Vigorous cross-examination, presentation of contrary evidence, and careful instruction on the burden of proof are the traditional and appropriate means of attacking shaky but admissible evidence. Additionally, in the event the trial court concludes that the scintilla of evidence presented supporting a position is insufficient to allow a reasonable juror to conclude that the position more likely than not is true, the court remains free to direct a judgment . . . and likewise to grant summary judgment. . . . These conventional devices, rather than wholesale exclusion under an uncompromising "general acceptance" test, are the appropriate safeguards where the basis of scientific testimony meets the standards of Rule 702.

Daubert, 509 U.S. at 595-96 (citations omitted).

admissibility screening. Indeed, they may be more effective, because they will be decided with a more complete understanding of the value of the science or technology in the panoply of the evidence presented. A concern that the jury may give too much weight to the evidence is really a *sufficiency of the evidence* issue that should be addressed at the end of the trial, not at the point when the evidence is offered. When dealt with as an admissibility issue, both the nature of the remainder of the evidence that will be offered and the mind-set and abilities of the jurors who will weigh it are far too speculative. Second, post-trial motions permit the trial to incorporate a broader range of relevant, and therefore helpful, evidence. Indeed, by encouraging the decision to be made later, judges may intuitively understand that their decision about scientific evidence is not purely a scientific one. Third, a jury's willingness to give controlling weight to a marginally probative science or technology may give important messages to the affected businesses, even though the jury's actions will not have legal consequence because the judge overturns it.³⁸ If, when the jury's verdict is returned, the presiding judge concludes that too much weight must have been given to the marginal evidence, the judge can either grant judgment notwithstanding the verdict or order that a new trial be held.

With luck, Professor Crump's contribution may be the catalyst for a complete reassessment of judicial approaches that have evolved over the past century to the use of expert scientific testimony in the courtroom. I, however, would not recommend that Professor Crump hold his breath in anticipation of this. Dislodging entrenched doctrine is a lot like renovating structures in which people happily reside. Regardless of its obvious need and intrinsic merit, change is often expensive, always disruptive, and, therefore, usually summarily dismissed along with its proponents, who are seen as trespassers, or worse yet, officious interlopers.

38. Of course, as with the admission of all other evidence, there is a possibility that the presiding judge (based on the remainder of the evidence heard) may conclude that the scientific evidence heard by the jury should not have been admitted. If it had been excluded and the remainder of the evidence would not permit a reasonable jury to conclude that a particular element of the cause of action were true, a verdict could be directed. If, despite the elimination of the scientific evidence, a directed verdict for the opposing side is still not warranted because sufficient other evidence is in the record to support the claim, the presiding judge would have to decide if it would suffice to give a limiting instruction to the jury to ignore the evidence, or whether fairness demands that a new trial be granted.

