Washington Law Review

Volume 69
Number 2 *Dedicated to Cornelius J. Peck*

4-1-1994

Out of the Twilight Zone: The Implications of *Daubert v. Merrill Dow Pharmaceuticals, Inc.*

Diana K. Sheiness

OUT OF THE TWILIGHT ZONE: THE IMPLICATIONS OF *DAUBERT v. MERRELL DOW PHARMACEUTICALS, INC.*

Diana K. Sheiness

Abstract: In *Daubert v. Merrell Dow Pharmaceuticals, Inc.*, the United States Supreme Court held that the Federal Rules of Evidence had not implicitly incorporated the general acceptance, or *Frye*, test for scientific evidence. Instead, the Court interpreted Rule 702 to mean that judges should admit challenged scientific testimony only after determining that its underlying method or theory is scientifically valid. This Note argues that the essence of the *Daubert* decision is that judges must ascertain whether or not the studies underlying proffered testimony have been performed in accordance with sound scientific principles. The Note analyzes several cases to illustrate appropriate and inappropriate applications of the new standard. Finally, the Note suggests a three-stage inquiry that will enable judges to elicit the information necessary for adequate evaluation of scientific studies.

That maggots arose from rotting meat and frogs from muddy fields was common knowledge a few hundred years ago. Simple living things were thought to spring from non-living matter through a process called spontaneous generation.[1] This belief was not easily dispelled, and spontaneous generation retained a large following as recently as the mid-1800s.[2]

Around 1859, the French microbiologist Louis Pasteur made observations causing him to suspect that fermentation was a biological process mediated by yeast.[3] He supposed that air currents must carry the yeast to the surface of the fruit, and from there into the fermentation vats. In the course of this work, he hypothesized that microorganisms such as yeast grew only from other yeast, and did not spring to life spontaneously.[4] This proposal roused the ire of proponents of spontaneous generation.[5] To support his hypothesis, Pasteur designed an experiment. First, he placed sterilized fermentable fluid into several

1. H. G. Wells, et al., *The Science of Life* 436–37 (1934).

2. *Id.*

3. Gordon Rattray Taylor, *The Science of Life* 122–24 (1963).

4. René Dubos, *Pasteur and Modern Science* 44–48 (1988). A "hypothesis" is a conjecture offered to explain a specified group of observations, and serves as guide and impetus for further investigation. "Hypothesis" and "theory" are often used interchangeably, but more accurately a theory is a verified or established explanation for the observations. *Random House College Dictionary* 654, 1362 (rev. ed. 1982). *See also* Kenneth R. Kreiling, *Scientific Evidence: Toward Providing the Lay Trier with the Comprehensible and Reliable Evidence Necessary to Meet the Goals of the Rules of Evidence*, 32 Ariz. L. Rev. 915, 965–67 (1990) (describing how scientists establish scientific knowledge).

5. Taylor, *supra* note 3, at 122–24.

flasks with necks that were curved like the neck of a swan. He then broke the necks off of some of the flasks. After the flasks were incubated to provide an opportunity for microbial reproduction, only the broken-necked flasks contained microbial growth. Microorganisms did not enter the swan-necked flasks because the walls of the curved necks had entrapped airborne microbes. This single carefully designed experiment inspired numerous other scientists to conduct studies that confirmed Pasteur's hypothesis. Eventually, these studies collectively dislodged the entrenched belief in spontaneous generation.

Under the *Frye* test, which predicates admissibility of scientific evidence on general acceptance within the relevant scientific community,[6] Pasteur could not have testified to his elegant experiment until other scientists had endorsed its credibility. Today, testimony based on recent experimentation may be admissible at trial. The U.S. Supreme Court in *Daubert v. Merrell Dow Pharmaceuticals, Inc.*[7] rejected the common law *Frye* test[8] and ruled that trial judges should admit challenged scientific testimony if it is scientifically valid and applicable to the facts at issue.[9]

This Note argues that the essence of the *Daubert* decision is that judges themselves must ascertain whether challenged scientific evidence is based on sound empirical observations. The Note contends that the *Daubert* approach ensures better than the *Frye* test that judges will recognize and exclude substandard scientific testimony, while admitting valid but novel evidence. The remainder of the Note focuses on effective implementation of the new standard. Part II distinguishes between "scientific" and other kinds of expert testimony to which judges should not apply *Daubert*. Part III analyzes post-*Daubert* cases, pointing out pitfalls that judges should avoid in employing the new standard. Finally, the Note provides a hypothetical evaluation of a scientific study, and suggests a method of inquiry that should provide the information necessary for a thorough evaluation of scientific testimony.

6. Frye v. United States, 293 F. 1013 (D.C. Cir. 1923).

7. 113 S. Ct. 2786 (1993).

8. *Id.* at 2794.

9. *Id.* at 2796.

I. DAUBERT BANISHES THE *FRYE* TEST FROM FEDERAL COURTROOMS

In *Daubert,* the Supreme Court resolved a disagreement among the circuits regarding the treatment of scientific testimony under the Federal Rules of Evidence, adopted in 1975. For years prior to 1975, the *Frye* test provided the common law standard for the admissibility of scientific evidence.[10] However, the Federal Rules of Evidence did not address the *Frye* test, leaving each federal circuit to decide for itself whether the Rules were meant to supplant or to perpetuate this familiar standard. Until *Daubert,* the issue of whether courts should still employ the *Frye* test sparked considerable commentary.[11]

A. Debate Over the Frye *Test*

An appeals court created the *Frye* test in 1923 when it reviewed a trial court decision to exclude evidence based on a newly-developed deception test that was a predecessor of today's polygraph.[12] With no reference to precedent, the court declared that the evidence was inadmissible because the technique lacked standing and recognition among psychologists and physiologists.[13] In an effort to distinguish between experimental and demonstrable stages of scientific discovery, the court stated:

> Somewhere in this twilight zone the evidential force of the principle must be recognized, and while courts will go a long way in admitting expert testimony deduced from a well-recognized scientific principle or discovery, the thing from which the deduction is made must be sufficiently established to have gained general acceptance in the particular field in which it belongs.[14]

Despite the subsequent widespread adoption of *Frye,* the test is not discussed in the original Advisory Committee's Notes on the Federal Rules of Evidence.[15] Rule 702, which addresses expert testimony based on scientific and other specialized knowledge, states:

---

10. *Daubert* noted that the *Frye* rule was still followed by a majority of courts. *Id.* at 2792–93.

11. Debates over *Frye* were a "well-established part of the academic landscape." *Id.* at 2793 n.4.

12. Frye v. United States, 293 F. 1013, 1013–14 (D.C. Cir. 1923).

13. *Id.* at 1014.

14. *Id.*

15. Christopher B. Mueller & Laird C. Kirkpatrick, *1993 Federal Rules of Evidence with Advisory Committee Notes and Legislative Action* 156–57 (1993). However, in 1991, the Judicial Conference

> If 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.[16]

The *Frye* test's proponents doubted the competence of juries and judges to evaluate scientific methodology, and feared that the judiciary was too overburdened to act as arbiter of scientific relevancy.[17] Those favoring the test's abolition argued that it was illogical to reject sound scientific evidence solely because it was novel, and urged a more active judicial role in weighing scientific evidence.[18] Moreover, because scientists stress skepticism rather than consensus, proving general acceptance was sometimes indeed an excursion into the twilight zone.[19]

Although the commentary accompanying Rule 702 did not explicitly reject the *Frye* test, the Rules anticipated its demise by incorporating a generally permissive view of the admissibility of evidence.[20] After the adoption of the Rules, about a third of American courts abandoned *Frye* and adopted relevancy tests based on their individual interpretations of

---

Advisory Committee on Civil Rules proposed amending Rule 702 to require that experts substantiate the validity of their sources. The committee commented that the proposed amendment would not mandate a return to *Frye*. *Id.* at 157–58.

16. Fed. R. Evid. 702.

17. *E.g.*, N. Kathleen Strickland & Leah S. Elkins, *A Current Assessment of Frye in Toxic Tort Litigation*, in *Toxic Tort Case Essentials: Strategies, Experts, Motions, and ADR*, 446 PLI/Lit. 321 (1992) (describing both sides of the debate over the *Frye* test); Peter W. Huber, *Galileo's Revenge: Junk Science in the Courtroom* (1991) (expounding the theme that unfair toxic tort awards often result from overly liberal admission of scientific evidence).

18. *E.g.*, Edward J. Imwinkelried, *Abolish the* Frye *Test*, 12 Cal. Law., Apr., 1992, at 63 (arguing that the *Frye* test assesses scientific validity essentially by counting heads); James E. Starrs, Frye v. United States *Restructured and Revitalized: A Proposal to Amend Federal Evidence Rule 702*, 26 Jurimetrics J. 249 (1986); Strickland & Elkins, *supra* note 17 (summarizing case law to illustrate split in interpreting Rule 702, and recommending "trustworthiness" as touchstone for admissibility of scientific testimony).

19. *E.g.*, Shiela Jasanoff, *What Judges Should Know About the Sociology of Science*, 32 Jurimetrics J. 345 (1992) (explaining that scientists frequently disagree regarding extent to which various theories in their field are proven); United States v. Downing, 753 F.2d 1224, 1236 (3d Cir. 1985) (discussing how courts have placed different meanings on concept of general acceptance).

20. *See, e.g.*, Patricia A. Wrona, *Combating "Junk Science" in Breast Implant Litigation*, Prentice Hall Law and Business, Nov. 1992 (arguing that the liberal slant of the Federal Rules of Evidence contributed to development by some courts of a "let it all in" approach to expert testimony).

Rule 702.[21] By the time *Daubert* reached the Supreme Court, sharp differences had developed among the circuit courts regarding the standards for admissibility of scientific evidence.[22]

B. The Supreme Court Replaces the Frye *Test*

In *Daubert,* the plaintiffs alleged that two infants' birth defects resulted from their mothers' ingestion of the anti-nausea drug Bendectin during pregnancy.[23] The plaintiffs based causation on animal studies, chemical structure similarities between Bendectin and drugs known to be teratogenic,[24] and statistical reanalyses of the results of previously published epidemiologic studies.[25] The authors of the original epidemiologic studies had concluded that Bendectin caused no significant increase in birth defects.[26] The district judge dismissed the animal and chemical structure studies, stating that only epidemiologic studies were pertinent to the issue of causation.[27] The judge furthermore rejected the statistical reanalyses because they did not meet the *Frye* test, and granted summary judgment to the defendants.[28] The Ninth Circuit affirmed.[29]

The Supreme Court granted review to resolve the split among the circuits regarding the proper standard for the admission of expert scientific testimony.[30] The debate over *Frye* had generated pressure for a Supreme Court decision on this issue, and *Daubert* provided a good vehicle for resolving the matter. In numerous other Bendectin disputes, the same testimony rejected by the Ninth Circuit had been admitted by

---

21. Imwinkelried, *supra* note 18, at 63. Although state courts are not bound by the Federal Rules, states tend to be influenced by federal courts, and many states have adopted rules of evidence based on the federal model. *See, e.g.,* Wash. R. Evid. 702.

22. *E.g.,* Strickland, *supra* note 17 (discussing the split among the circuits); U.S. Dept. of the Army, Pamphlet 27-22, *Military Criminal Law Evidence,* ch. 6–3 (1987) (listing circuit court decisions that have supported or rejected *Frye*).

23. Daubert v. Merrell Dow Pharmaceuticals, Inc., 727 F. Supp. 570, 571 (S.D. Cal. 1989).

24. Teratogens are substances that induce abnormal embryologic development. *Stedman's Medical Dictionary* 1418 (24th ed. 1982).

25. Epidemiology is the study of the prevalence and spread of disease in a community. *Id.* at 474.

26. *Daubert,* 727 F. Supp. at 575.

27. *Id.*

28. *Id.* at 575–76.

29. Daubert v. Merrell Dow Pharmaceuticals, Inc., 951 F.2d 1128 (9th Cir. 1991).

30. Daubert v. Merrell Dow Pharmaceuticals, Inc., 113 S. Ct. 2786, 2792 (1993).

some courts and excluded by others.[31] Because large damage awards were involved, the Bendectin controversy had gained a high profile. The question *Daubert* presented to the Supreme Court was whether the *Frye* test was appropriate grounds for excluding the plaintiffs' reanalyses of epidemiologic studies.

The Supreme Court concluded that the "austere" *Frye* test was incompatible with the liberal thrust of the Federal Rules of Evidence.[32] On this issue, the Court held unanimously that a court may not disallow scientific evidence solely because it is not generally accepted. However, the Court emphasized that diminished reliance on *Frye* did not mean that the Federal Rules supported the indiscriminate admission of any evidence purporting to be scientific.[33] To the contrary, the Court reasoned that if scientific evidence is not based on sound scientific principles, then it is not relevant under Rule 402, and hence cannot satisfy the Rule 702 requirement that evidence be helpful to the trier of fact.[34] Thus, the Court designated trial judges as gatekeepers responsible for evaluating the admissibility of scientific evidence.[35] The new standard differs from the *Frye* test, but is no less exacting.[36]

The Court charged trial judges with determining whether proffered testimony constitutes "scientific knowledge" that will assist the factfinder in understanding or determining disputed facts.[37] According to

---

31. Michael D. Green, *Expert Witnesses and Sufficiency of Evidence in Toxic Substances Litigation: The Legacy of Agent Orange and Bendectin Litigation,* 86 NW. U. L. Rev. 643, 661–68 (1992) (reviewing the major Bendectin cases, and revealing that the inconsistencies among the Bendectin cases did not turn on the *Frye* test).

32. *Daubert,* 113 S. Ct. at 2794. *Frye* is "austere" because new or obscure theories that are otherwise sound may lack sufficient exposure to have gained a large following.

33. *Id.* at 2794–95. The *Daubert* opinion was unanimous only with respect to the disposition of the *Frye* test. Justices Rehnquist and Stevens dissented from the second part of the opinion, in which the Court described how Rule 702 should be applied. *Id.* at 2799. The dissenters preferred the opinion to stop with dissociating the *Frye* test from Rule 702. *Id.* at 2795 n.7, 2799.

34. *Id.* at 2795.

35. *Id.* at 2795 n.7, 2798. The dissenters lamented the majority's admonition to judges to become "amateur scientists." *Id.* at 2800. However, the majority expressed confidence that judges were capable of dealing effectively with scientific evidence. *Id.* at 2796.

36. Had the Court dismantled *Frye* without providing a strong alternative standard, the *Daubert* decision might have suggested that the Court favored unbridled admission of all testimony offered by qualified experts. *See, e.g.,* Wrona, *supra* note 20.

37. *Daubert,* 113 S. Ct. at 2796. Federal Rule of Evidence 104, which addresses preliminary questions, authorizes judges to consider admissibility of scientific evidence outside the hearing of the jury. The relevant sections of the rule are:

(a) Questions of admissibility generally. Preliminary questions concerning . . . the admissibility of evidence shall be determined by the court, subject to the provisions of subdivision (b). In

the Court, "scientific" implies a grounding in the methods and procedures of science, and "knowledge" connotes more than a subjective belief or unsupported speculation.[38] To qualify as "scientific knowledge," an inference or assertion has to be based on the scientific method.[39]

The Court further instructed trial judges to ensure that any scientific testimony or evidence admitted is both relevant and reliable, and thus helpful to the trier of fact.[40] For these purposes, the Court equated evidentiary reliability with scientific validity.[41] In considering relevance, the Court instructed judges to ensure that the proffered testimony was sufficiently tied to the facts of the case to be indeed helpful.[42] This means that judges must not only assess whether the reasoning or methodology underlying challenged testimony is scientifically valid, but also whether one can properly apply the evidence to the facts in issue.[43] Moreover, the Court indicated that judges were to focus their analyses on the methodology underlying the testimony, and not on the expert's conclusions.[44]

In what may prove to be the most influential aspect of the opinion, the Court recommended that judges weigh four factors in evaluating the

---

making its determination it is not bound by the rules of evidence except those with respect to privileges.

....

(c) Hearing of jury. Hearings on the admissibility of confessions shall in all cases be conducted out of the hearing of the jury. Hearings on other preliminary matters shall be so conducted when the interests of justice require....

Fed. R. Evid. 104.

38. *Daubert*, 113 S. Ct. at 2795.

39. *Id.* The Court essentially equated the "scientific method" with the method of empirical observation practiced by modern scientists. This practice is embodied in the concept of "falsifiability." *See infra* note 46 and accompanying text.

40. *Daubert*, 113 S. Ct. at 2795.

41. "In a case involving scientific evidence, *evidentiary reliability* will be based upon *scientific validity*." *Id.* at 2795 n.9.

42. *Id.* at 2796.

43. The Court cautioned judges that what was scientifically "fit" for one purpose was not necessarily valid for other purposes. As an example, the Court explained that a study of the phases of the moon might demonstrate that a certain night was dark, but evidence that the moon was full on a certain night could not assist a trier of fact in determining that an individual was likely to have behaved irrationally on that night. *Id.*

44. *Id.* at 2797.

scientific validity of expert testimony.[45] These factors are general acceptance, publication in peer reviewed journals, the known or potential rate of error and standards controlling the technique's operation, and whether the theory or technique has been tested in accordance with scientific methods.[46] The Court stipulated that these four factors were not meant to be a definitive checklist or test, and did not indicate how they should be weighed or balanced.[47] Nonetheless, because the four factors provide specific guidance for assessing the admissibility of scientific evidence, they are likely to be influential in future cases. At the same time, their application could vary considerably, because the Court did not clarify how future judges should gauge the relative importance of each factor.

II. TESTING IS THE HEART OF THE *DAUBERT* STANDARD

Unless judges apply *Daubert* in a consistent fashion, confusion may develop over how to interpret the new standard. The testing requirement embodies the essence of the new standard, and is the primary feature that distinguishes the *Daubert* standard from the *Frye* test. Accordingly, judges should base admissibility decisions primarily on the soundness of the observations or tests that support the evidence, using the other three factors to help determine whether or not the proffered evidence is scientifically sound.

Courts should not misconstrue *Daubert's* inclusion of general acceptance and peer review among the four factors as somehow perpetuating the *Frye* test as a prerequisite for admissibility. General acceptance merits judicial attention because methodology or theories that have been generally accepted are likely to withstand scrutiny under the

45. *Id.* at 2796–97.

46. *Id.* The Court characterized "scientific" hypotheses or methods as those subject to "falsification," thereby endorsing the prevailing paradigm of "falsifiability" as defining what are sound scientific principles. *Id.* The concept of "falsification" is based on the work of philosopher Karl Popper, and concerns the modern scientific practice of inferring general conclusions based on a set of organized empirical observations. Green, *supra* note 31, at 645. The concept recognizes that while induction can never conclusively prove a hypothesis, it can falsify one. *Id.* This viewpoint emphasizes the fact that empirical observations tend to work better to disprove than to prove hypotheses. For example, a single striking exception to a broad and widely-held theory has great power to destroy the theory's credibility, but rarely can a single observation establish confidence in a new theory.

47. *Daubert,* 113 S. Ct. at 2796.

other considerations mandated by the new standard.[48] Similarly, peer review is helpful in determining admissibility only because of its relationship to the general acceptance and testing factors.[49] Although publication in peer-reviewed journals is an important avenue by which scientific theories eventually become generally accepted, it does not necessarily indicate that a theory has already achieved general acceptance. This is because such journals serve largely as a forum for scientific progress reports. These reports contribute to the scientific enterprise by inspiring other scientists to further test new theories. In this sense, peer review and publication have a direct bearing on whether or not theories and methods ever become generally accepted. However, because general acceptance is not determinative under the new standard, the Court would not have instructed judges to consider publication in peer-reviewed journals unless peer review had some significance other than its relationship to general acceptance.

Peer review does have a significant bearing on *Daubert's* testing requirement. Since peer review tends to validate a study's soundness, publication of the studies supporting an expert's opinion can add credence to claims that the opinion's theoretical basis has been adequately tested. Peer review tends to confirm that the testimony's supporting studies were proficiently performed, even though the publications may involve a novel hypothesis. Indeed, the only rationale the Court offered for including the publication factor was that peer review provides some quality control for studies submitted for publication.[50] But while publication in a peer-reviewed journal may attest to the validity of a scientific study, the converse is not true.

---

48. *Id.* at 2797. The Court noted further that firmly established theories were properly subject to judicial notice. *Id.* at 2796 n.11.

49. Scientists invite peer review by submitting manuscripts describing their studies to the editor of a scientific journal. To determine whether the manuscript merits publication, the editor will ask two or three qualified scientists (the author's "peers") to referee the paper for conformity to scientific standards. Based on the work's merits, each editor recommends whether the manuscript should be published, revised after additional experimentation, or rejected outright. Copies of the reviews, minus the referees' identities, are submitted to the author along with the editor's final decision for the disposition of the paper. Journals generally also notify the referees of the final disposition of the paper, a practice that discourages editors from arbitrarily overriding reviewers' recommendations. If not satisfied with the review process, authors can request that the editor solicit opinions from additional referees. As a consequence of these practices, the process provides a relatively objective evaluation of scientific validity. *See, e.g.*, Kreiling, *supra* note 4, at 970–71; Arnold S. Relman & Marcia Angell, *How Good is Peer Review?*, 321 New Eng. J. Med. 827 (1989).

50. The Court stated that "submission to the scrutiny of the scientific community is a component of 'good science' . . . because it increases the likelihood that substantive flaws in methodology will be detected." *Daubert*, 113 S. Ct. at 2797. Once peer reviewers detect flaws in a submitted report, publication becomes conditional on the correction of these flaws. *See supra* note 49.

Accordingly, lack of publication is no longer a sufficient reason for judges to label scientific evidence invalid.

Like publication, the rate of error is related to whether or not theories or methods have been adequately tested.[51] Although the Court did not articulate any rationale for including this factor, it is important because the very purpose of testing is to establish that a method is reliable.[52] Because a method's rate of error reflects the extent of its reliability, rate of error bears directly on relevance. Accordingly, the Court's emphasis on this factor is consistent with a judge's need to balance relevance against potential for prejudice under Rule 403.[53] The danger of juries placing undue weight on the evidence is greater for scientific than for other types of evidence.[54] For example, a judge might find otherwise relevant evidence inadmissible when it derives from a diagnostic test with a high rate of misidentification. Unless a court takes rate of error into account, it cannot accurately weigh a statement's probative value against its potential for prejudice.

In light of the foregoing, *Daubert's* primary message is that only adequately tested theories provide a valid basis for scientific evidence. This focus on testing will enable judges to determine whether challenged scientific evidence is sufficiently sound to warrant admission, even when they are dealing with novel methodologies.[55] *Daubert* reflects the fact

51. The rate of error for a technique or method is the frequency with which the method produces erroneous results. Error can result from flaws in the underlying hypothesis, equipment failure, arbitrary changes in established procedures, or human error. For discussions of this topic, *see, e.g.*, Kreiling, *supra* note 4, at 978–82 (enumerating sources of error in scientific procedures); Green, *supra* note 31, at 649–58 (reviewing common sources of error in epidemiologic studies).

52. Scientific studies often include statistical analyses or other indicia of reproducibility that provide a measure of the data's reliability and significance. Such analyses can help judges determine whether the studies are sufficiently reliable to support scientific testimony.

53. Fed. R. Evid. 403 states: "Although relevant, evidence may be excluded if its probative value is substantially outweighed by the danger of unfair prejudice, confusion of the issues, or misleading the jury, or by considerations of undue delay, waste of time, or needless presentation of cumulative evidence."

54. *See* Aronson, *supra* note 21, at 403-7 (explaining that juries can be "mesmerized" by scientific facts and theories, and hence are more susceptible to being confused, misled, or unduly prejudiced by scientific than by other types of evidence).

55. Ironically, the Court's requirement for empirical testing does not resolve the admissibility of the controverted evidence in the *Daubert* trial itself. The issue on appeal to the Supreme Court was not whether the evidence was admissible, but only whether the *Frye* test was appropriate grounds for its exclusion. The trial judge excluded the reanalysis of previous epidemiologic studies primarily because the plaintiffs had proposed an unorthodox method of assigning significance to the studies' results. Daubert v. Merrell Dow Pharmaceuticals, Inc., 727 F. Supp. 570, 575 (S.D. Cal. 1989). In claiming that the studies' results did indicate teratogenicity, the plaintiffs asserted that the studies' findings had surpassed a threshold beyond which Bendectin should be considered dangerous, and in essence argued for a lower threshold than the one generally used by epidemiogists. *Id; see also*

that scientists themselves derive new knowledge by drawing generalizations, or inferences, from concrete observations. To permit scientific testimony based on untested theories would exempt the testimony from the dictates of common sense. Unless somebody has systematically verified a given method or hypothesis, not even the testifying expert knows whether the method works as claimed or if the hypothesis is credible.

III. DETERMINING *DAUBERT'S* APPLICABILITY

The new standard offers specific guidance for judges faced with scientific evidence. Hence, before applying *Daubert,* judges evaluating proffered expert testimony must distinguish "scientific" from technical or other types of specialized knowledge.[56] However, the Court did not explain how judges could determine which areas of expertise fell into these presumably distinguishable categories. In light of this ambiguity, there is a risk that courts will develop different criteria for determining when *Daubert* is applicable. Some courts may confine the standard too narrowly, while others may extend the new standard beyond its intended reach. Defining the boundaries of the new standard will facilitate the development of consistent applications.

A. Daubert *Should Not Be Applied to All Expert Testimony*

While setting "scientific" apart from other types of expert testimony, the Court provided no basis for differentiating "scientific" from the "technical" evidence mentioned in Rule 702. According to one dictionary, "science" means pertaining to a branch of knowledge dealing with a body of facts systematically arranged and showing the operation of general laws, or pertaining to any skill that reflects a precise application of facts or knowledge.[57] By comparison, "technical" means pertaining to an art or science, being of significance primarily to persons

---

DeLuca v. Merrell Dow Pharmaceuticals, Inc. 911 F.2d 941, 946–49 (3d Cir. 1990), *cert. denied,* 114 S. Ct. 961 (1994) (describing in detail the opposing parties' views of how the statistical analyses should be done). Assignments of such thresholds are analogous to value judgments, hence the proper designation for such a threshold cannot be determined by tests or experiments.

56. The Court stated, "[o]ur discussion is limited to the scientific context because that is the nature of the expertise offered here." Daubert v. Merrell Dow Pharmaceuticals, Inc., 113 S. Ct. 2786, 2795 n.8 (1993).

57. *Random House College Dictionary* 1179 (rev. ed. 1982).

with specialized knowledge, or pertaining to the mechanical or industrial arts and the applied sciences.[58] These overlapping definitions provide little guidance for the correct application of *Daubert.*

But courts need not rely on semantic distinctions to determine whether or not *Daubert* should be applied. According to the Court, a key question in resolving whether evidence is "scientific" is whether the theory or technique is susceptible to testing, that is, to experimental confirmation.[59] Hence, judges should apply the new standard to all fields of specialized knowledge whose tenets ordinarily are based on experimental observations. Theories or methods susceptible to testing should not escape scrutiny under *Daubert* merely because they can be characterized as "technical." If an expert derives authority from a field susceptible to testing but offers an opinion that eludes empirical confirmation, the testimony should be excluded. However, a requirement of empirical testing for testimony whose sources are not susceptible to verification would be too restrictive.

In cases subsequent to *Daubert,* courts have correctly refused to extend *Daubert* to evidence that clearly was not scientific. In *United States v. Locascio,*[60] the defendants urged the Second Circuit to overrule a pre-*Daubert* admission of the testimony of an expert on organized crime families.[61] The expert had based his opinion primarily on hearsay and other ordinarily inadmissible sources.[62] The appeals court refused to interpret *Daubert* as imposing a mandatory requirement that only verifiable sources be permitted to form the basis for expert testimony.[63] In a similar dispute in *Hawthorne Partners v. AT & T Technologies, Inc.,*[64] a trial court refused to apply *Daubert* even though quantitative statistical methods theoretically could have been used by the real estate appraiser who testified as an expert witness.[65] As in *Locascio,* the *Hawthorne* witness's field of expertise could not rationally be classified as scientific. Had these two courts ruled differently, most or all expert testimony in those courts thereafter might have been restricted to knowledge based on verifiable or documented sources. Such an

58. *Id.* at 1349.

59. *Daubert,* 113 S. Ct. at 2796.

60. 6 F.3d 924 (2d Cir. 1993).

61. *Id.* at 936.

62. *Id.*

63. *Id.* at 938.

64. No. 91 C 7167, 1993 WL 311916 (N.D. Ill. Aug. 11, 1993).

65. *Id.* at *4.

extension of *Daubert* would have been inconsistent with Rule 702's liberal inclusion of testimony by those whose expertise comprises "knowledge, skill, experience, training, or education."[66] Nothing in the *Daubert* opinion indicates that the Supreme Court intended to constrict the scope of Rule 702 as suggested by the parties in *Locascio* or *Hawthorne.*

Dilemmas for judges may arise when they are confronted with testimony based on areas of expertise that could apply scientific techniques but typically do not. The *Hawthorne* decision nicely resolved such a dilemma by presuming that the testimony was valid when others in the same field routinely used similar methods. Application of *Daubert* should be restricted to instances when the expertise at issue derives from a field that normally utilizes scientific or quantitative techniques.

B. *Scientific Testimony Presented by Physicians Should Be Subject to* Daubert

Judges should base their applications of *Daubert* on the nature of the proffered testimony, rather than relying on the professional designation of the expert witness. Physicians who testify in court traditionally have been held to a subjective standard based on their willingness to assert their opinion to a "reasonable degree of medical certainty."[67] Although this standard is credible when physicians offer testimony that rests on their years of clinical experience, it becomes illogical when it authorizes physicians to ground their opinions in inadequate scientific studies.

The danger of misapplying the "reasonable degree of medical certainty" standard is illustrated by *Wells v. Ortho Pharmaceutical Corp.*[68] In *Wells,* the Eleventh Circuit affirmed an award based on the plaintiffs' contentions that a spermicide used during early pregnancy had caused a child's birth defects.[69] The plaintiffs' testimony rested on a few highly questionable epidemiologic studies, while numerous other studies

66. Fed. R. Evid. 702. When the Rules were adopted, "experience" was added to the traditional means by which persons could become qualified as experts. Under prior law, education and training were the only permissible means for becoming qualified. Aronson, *supra* note 21, at 702-2.

67. Bert Black, *Evolving Legal Standards for the Admissibility of Scientific Evidence,* 239 Science 1508, 1510 (1988) [hereinafter Black, *Evolving Legal Standards*]; Bert Black, *A Unified Theory of Scientific Evidence,* 56 Fordham L. Rev. 595, 659–61 (1988).

68. 615 F. Supp. 262 (N.D. Ga. 1985), *aff'd,* 788 F.2d 741 (11th Cir.), *cert. denied,* 479 U.S. 950 (1986).

69. Wells v. Ortho Pharmaceutical, 788 F.2d at 741–42.

overwhelmingly supported the product's safety.[70] Disregarding the relative merits of the conflicting epidemiologic studies, the district court painstakingly evaluated the demeanor, bias, and consistency of several physicians and other experts who had presented testimony.[71] Adhering to the "reasonable degree of medical certainty" standard, the judge delved deeply into the manner in which the opposing experts had expressed their beliefs, while placing little weight on the relative validity of their underlying theories.[72] However, the primary study upon which the plaintiffs had relied was a preliminary report that was subsequently repudiated by its authors.[73]

The *Wells* dispute fit squarely into the field of epidemiology, which has well-established principles for determining whether or not a substance causes deleterious effects in humans.[74] Under *Daubert,* an evaluation of this testimony would have focused on the quality of the opposing epidemiologic studies, regardless of the witnesses' demeanors. Scientific evidence should derive its aura of authority not from the eloquence of its proponents, but from a general belief that the scientific approach produces trustworthy results.[75] The relative degree of certainty expressed by several witnesses is less probative of a substance's toxicity than are the actual data underpinning the witnesses' opinions. Whenever testimony presented by physicians involves facts that derive from scientific studies, courts should evaluate the testimony according to *Daubert,* and not according to the "reasonable degree of medical certainty" standard.

---

70. Wells v. Ortho Pharmaceutical, 615 F. Supp. at 269.

71. *Id.* at 266–91.

72. The court explained that it did not intend to resolve the dispute about the safety of spermicides, but rather to assess the credibility of the opposing witnesses by paying "close attention to each expert's demeanor and tone." *Id.* at 267. Indeed, the testimony viewed most favorably by the court was that of a physician who the court found to be particularly effective in conveying the impression of credibility. *Id.* at 269–72. The inconclusive nature of the studies upon which the physician relied apparently did not detract from his credibility in the eyes of this court. *Id.*

73. Jeff L. Lewin, *Calebresi's Revenge? Junk Science in the Work of Peter Huber,* 21 Hofstra L. Rev. 183, 202 (1992). *See also* Black, *Evolving Legal Standards, supra* note 67, at 1511 (describing the outrage in the medical community at the irresponsibility of the *Wells* decision).

74. Cogent and succinct summaries of epidemiologic principles are given in Green, *supra* note 31, at 646–49, and in DeLuca v. Merrell Dow Pharmaceuticals, Inc., 911 F.2d 941, 946–49 (3d Cir. 1990).

75. *See, e.g.,* Porter v. Whitehall Laboratories, Inc., 791 F. Supp. 1335, 1343 (S.D. Ind. 1992) (stating that "an expert is a conduit of facts and not merely a subjective speculator relying on stature alone"), *aff'd,* 9 F.3d 607 (7th Cir. 1993).

IV. POST-*DAUBERT* DECISIONS: THE GOOD, THE BAD, AND THE UGLY

In addition to determining whether *Daubert* is applicable to proffered expert testimony, courts encountering scientific evidence will face numerous decisions concerning how the new standard should be applied. Inconsistent applications could result from variations in how judges evaluate scientific validity, or from differences in their willingness to re-evaluate pre-*Daubert* rulings. Courts can avoid these inconsistencies if they carefully limit admission of scientific testimony to that based on adequately tested theories or methodologies, and if they ascertain that *Daubert* has been satisfied before adhering to earlier admissibility rulings.[76]

A. The Good: Chikovsky *Correctly Applied* Daubert

Within three months of *Daubert,* a federal district court in Florida refused to admit evidence supporting allegations that a pregnant woman's topical application of Retin-A, an acne medication,[77] had caused her newborn's birth defects.[78] In *Chikovsky v. Ortho Pharmaceutical, Corp.,* the defendants countered that the infant's defects were consistent with a known genetic condition, and that absorption through the skin of Retin-A was minimal.[79] An obstetrician/gynecologist testified on behalf of the plaintiff, and the court's analysis of this doctor's testimony provides a model for effective application of the new standard.[80]

The court weighed the doctor's opinion on its scientific merits despite its lack of published support, noting that *Daubert* had precluded using

---

76. Appellate courts will be required to re-examine cases that were tried prior to *Daubert. See, e.g., Porter,* 9 F.3d at 607 (concluding that criteria used by the trial court to evaluate the testimony would have satisfied *Daubert*); Thomas v. American Cyanamid Co., No. 91-5326, 1993 WL 369129 (6th Cir. Sept. 21, 1993) (requiring both parties to submit supplemental briefs where the trial records were insufficient for *Daubert* analysis).

77. Retin-A is applied topically, that is, directly to the skin, rather than being ingested.

78. Chikovsky v. Ortho Pharmaceutical Corp., 832 F. Supp. 341 (S.D. Fla. 1993).

79. *Id.* at 343.

80. Additional cases exemplify the appropriate analysis under *Daubert. See, e.g.,* Government of the Virgin Islands v. Penn, No. CRIM. 91-38, 1993 WL 388146 (V.I. Aug. 26, 1993) (evaluating DNA fingerprinting under *Daubert*); Scales v. George Washington University, Civ. A. No. 89-0796-LFO, 1993 WL 304016 (D.C. Cir. July 27, 1993) (revealing the flaws and inconsistencies in the statistical analysis presented by a social scientist to prove racial discrimination in hiring).

general acceptance as a litmus test for admissibility.[81] The doctor admitted that Retin-A had never been tested directly for its teratogenicity,[82] but testified that he had based his opinion in large part on studies suggesting that high oral doses of vitamin A caused fetal harm in humans. Retin-A is a derivative of vitamin A. The doctor attributed the properties of vitamin A to Retin-A, a related but different compound, without justifying this extrapolation.[83] Ironically, the doctor acknowledged prescribing vitamin A to his own pregnant patients.[84] Moreover, the doctor presented no evidence that Retin-A was capable of penetrating the skin and entering the bloodstream. In short, there was no experimental basis for the notion that Retin-A applied to the mother's face could have reached the fetus.

In addition to the vitamin A studies, the doctor relied on published indications that Accutane, another vitamin A derivative, was a potential teratogen.[85] However, Accutane is taken orally and in much higher doses than Retin-A. Moreover, the birth defects with which it is associated are different from those suffered by the infant in *Chikovsky*. Recognizing these differences, the court held that the Accutane studies were not relevant to the plaintiff's claims.[86]

The *Chikovsky* court focused on the elements of sound scientific study when it noted that the witness's hypothesis had not been directly tested, that he had not ruled out other causes for the observed effect, and that he had relied upon studies that were insufficiently related to the disputed issue.[87] In essence, the court discerned that the physician's opinions were not based on scientifically valid tests, and hence did not meet the *Daubert* standard.

81. *Chikovsky*, 832 F. Supp. at 344, 345 note 5 (noting that under *Daubert* publication was not an indispensable prerequisite for admissibility).

82. *See* definition *supra* note 24.

83. Toxicologists have established that even a single change in a chemical's molecular structure can vastly alter its effects in humans. Green, *supra* note 31, at 658.

84. *Id.* at 345–46. Furthermore, since the doctor was not trained in genetics, by his own admission he could not rule out a genetic cause for the infant's defects. *Id.*

85. *Id.* at 346.

86. *Id.*

87. *Id.* at 345–46.

B. *The Bad:* Bynum *Fails to Ensure that a Generally Reliable Method Has Been Applied Appropriately*

Despite a good faith effort to follow *Daubert,* the Fourth Circuit confirmed the admission of evidence that had not met *Daubert's* testing requirement. In *United States v. Bynum,*[88] the prosecution implicated the defendant in a conspiracy to sell drugs in part by using gas chromatography[89] to demonstrate that three crack cocaine samples seized from different locations had derived from a single source. After the trial, the defendant discovered that gas chromatography had not been used before for this purpose. He argued on appeal that the *Frye* test should have precluded the testimony. *Daubert* was decided while the appeal was pending, and the *Bynum* court concluded that the "government's proffer of evidence could hardly have better anticipated *Daubert.*"[90]

The prosecution had explained the hypotheses underlying gas chromatography, had listed publications for which the technique had been subjected to peer review, and had stated that the technique enjoyed general acceptance among forensic chemists.[91] However, this testimony demonstrated merely that gas chromatography was a valid and generally accepted technique. Even though the expert established that gas chromatography could distinguish mixtures of chemicals that differed in their composition, he did not establish that batches of cocaine invariably differed in their chemical composition.[92] Unless batches of cocaine so differed, individual batches could not be distinguished by gas chromatography. Under *Daubert,* the court should have required the

88. 3 F.3d 769 (4th Cir. 1993).

89. Chromatography is any physical process whereby the individual components of a chemical mixture are separated by differential movement through a two-phase system. One phase of the system contains the sample being analyzed, and the other phase is an adsorbent to which the individual components of the sample become differentially attached according to size, electrical charge, or other physical properties. In gas chromatography, the sample to be analyzed is vaporized, then separated into its individual chemical components on a stationary phase. The resulting pattern on the stationary phase reflects the kinds and amounts of the components in the original mixture. *Stedman's Medical Dictionary* 275 (24th ed. 1982).

90. *Bynum,* 3 F.3d at 773.

91. *Id.* at 773.

92. The expert's theory was that naturally derived products like crack cocaine will contain numerous impurities and byproducts of decomposition. The expert proposed further that the precise mathematical proportions of the various impurities would inevitably differ among samples of cocaine unless the samples were manufactured together. *Id.* at 772 n.3. While plausible, the theory contained two untested assumptions. The first of these was that crack cocaine actually contains detectable amounts of several different impurities, and second, that different batches of crack vary significantly with regard to their content of these putative impurities.

prosecution to present studies showing that cocaine samples from several different sources did not give similar results when analyzed by gas chromatography.

Nor did the *Bynum* court consider the gas chromatography technique's potential rate of error. Two different but equally qualified laboratories had independently analyzed the three samples and had differed in their conclusions regarding one of the samples.[93] Despite this compelling suggestion that the technique produced inconsistent results for cocaine analysis, the Fourth Circuit affirmed the admission.[94]

By contrast, a Washington state case illustrates a proper evaluation of gas chromatography evidence.[95] In *State v. Huynh,* the appellate court's standard of review was the *Frye* test combined with a threshold requirement of scientific reliability of the underlying technique.[96] The court's inquiry elicited the same type of information that would have been drawn out in an analysis under *Daubert.* The prosecution's expert had used gas chromatography to compare partially burned gasoline from the scene of a suspicious fire with unburned gas from a can found in the defendant's car. The court acknowledged the general utility of gas chromatography for comparing gasoline samples. The court went on to list several variables that may have affected the particular results in the case and held that the variables had not been sufficiently tested.[97] Accordingly, the court held that the testimony should have been excluded. In contrast to the *Bynum* court, which considered only the general acceptability of gas chromatography, the *Huynh* court evaluated the reliability of the specific application of the technology.

The *Bynum* court's error was not benign. Because of this ruling, other courts within the Second Circuit could consider gas chromatography admissible without considering its potential unreliability in particular applications. In *Daubert,* the Supreme Court explicitly cautioned judges to ascertain that an expert's methods "properly can be applied to the facts

93. *Id.* at 772–73. For a sample-testing procedure to be considered reliable, the same results should be obtained by any qualified laboratory that performs the test correctly. In *Bynum,* there was no suggestion that either laboratory had performed the procedure incorrectly.

94. The court remarked that admission of the gas chromatography results was not central to the case because so much other evidence had supported the conviction. Thus, the court could have rejected the gas chromatography evidence without overturning the conviction. *Id.* at 773–74.

95. State v. Huynh, 49 Wash. App. 192, 742 P.2d 160 (1987), *review denied,* 109 Wash. 2d 1024 (1988).

96. *Id.* at 195–96, 742 P.2d at 163–64.

97. *Huynh,* 49 Wash. App. at 196–98, 742 P.2d at 164–65.

in issue."[98] Thus, when evaluating testimony based on well-established technologies, courts should verify that the technique produces accurate results when used for the purpose at stake. Unless a method's suitability for a new use is verified, judges should not assume that the method is relevant to the instant controversy.

C. *The Ugly: Inappropriate Reliance on Past Decisions*

During the next several years, much expertise ruled admissible under other standards will need to be reevaluated under *Daubert.* A judge's failure to carefully perform this duty can perpetuate past errors, or can result in the unfair exclusion of evidence based on technology that may have been improved since its prior court review. Even well-established forensic techniques should be re-examined to ascertain that they have indeed satisfied *Daubert's* testing requirement.

Testimony should not be admitted merely because prior courts have found it admissible. The paraffin glove test fiasco illustrates the folly of disregarding this admonition. In 1936, *Commonwealth v. Westwood*[99] was the first appellate decision to approve the results of this new test as evidence that a suspect had recently fired a gun. To perform the test, the suspect's hand was coated with hot wax, and the cooled and hardened wax was then peeled off along with any particles that were on the hand. The particles were treated with a chemical that makes them turn blue if they contain nitrates. A chemist testifying for the defendant pointed out that not only gunpowder, but also common substances such as toothpaste, cigar ashes, and matches contain nitrates.[100] Moreover, the F.B.I. the previous year had published a bulletin reporting that smokers, miners, chemists, and pharmacologists were likely to have substances on their hands that would give a positive result with this test.[101] Hence, a basis for rejecting this test existed even in 1936.

Despite being aware of the paraffin test's innate unreliability, the *Westwood* court admitted the testimony. The defendant claimed that the test did not prove that he had fired a gun, but did not challenge the

98. Daubert v. Merrell Dow Pharmaceuticals, Inc., 113 S. Ct. 2786, 2796 (1993).

99. Commonwealth v. Westwood, 188 A. 304 (Pa. 1936).

100. *Id.* at 308–09.

101. *The Diphenylamine Test for Gunpowder,* 4 F.B.I. L. Enforcement Bull. 5 (1935). With no apparent justification, the F.B.I. concluded in this same bulletin that the test was of value in determining whether subjects had recently fired a gun. Perhaps this F.B.I. endorsement of the test was responsible for the test's widespread acceptance among forensic scientists in the years that followed.

testimony under *Frye*. Thereafter, trial courts treated the technique as if it were generally accepted.[102] No report of the test's rejection by an appellate court appeared until 1959,[103] and the test was not thoroughly discredited until 1967 when a comprehensive scientific evaluation of the test was published.[104] The *Daubert* standard can ensure that mistakes like that in *Westwood* are not perpetuated.[105] Despite the paraffin test's general acceptance, a proponent of this test would have been unable to demonstrate the test's scientific validity as *Daubert* requires.

Post-*Daubert* courts have erred by excluding previously disreputable evidence without a fair hearing,[106] by admitting evidence based on unvalidated methodology,[107] and by neglecting indications that a method was error-prone when applied to the issues in controversy.[108] Judges should take care to avoid these traps, and also to avoid perpetuating prior endorsements of evidence without first ascertaining that earlier courts had probed the methodology in a manner consistent with *Daubert*.[109]

V. EVALUATING THE VALIDITY OF SCIENTIFIC EVIDENCE

Application of the new standard will be fraught with pitfalls, especially for judges with little understanding of science. Such judges should take full advantage of training and resources in the area of science and law.[110] In so doing, they will discover that the basic elements of a

---

102. Randolph N. Jonakait, *Will Blood Tell? Genetic Markers in Criminal Cases*, 31 Emory L. J. 833, 854–57 (1982) (stating that forensic scientists embraced the paraffin test despite lack of evidence for the test's reliability). As described by one commentator, "[t]he number of convictions tainted by the paraffin test cannot be estimated." Alfred H. Novotne, *Scientific Evidence: Challenging Admissibility*, Dept. Army Pamphlet, 27-50-190, 23 Army Law (1988).

103. Brooke v. People, 339 P.2d 993 (Colo. 1959) (rejecting the paraffin test based on a 1956 report noting the test's tendency to give false positive results).

104. Mary E. Cowan & Patricia L. Purdon, *A Study of the "Paraffin Test,"* 12 J. Forensic Sci. 19 (1967).

105. *See* Starrs, *supra* note 18, at 257–58 (discussing instances when judicial notice has served to perpetuate erroneous decisions to admit scientific evidence).

106. United States v. Black, 831 F. Supp. 120 (E.D. N.Y. 1993) (excluding polygraph evidence on the basis of earlier decisions that did not discuss the scientific validity of the method).

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

108. *See supra* note 93 and accompanying text.

109. Once the scientific validity of a method or theory has been ascertained under *Daubert*, judges thereafter can invoke judicial notice with respect to that same method or theory.

110. If judges assimilate basic scientific principles, these will be applicable to a wide range of scientific issues. Resources soon available to judges will include a set of protocols for forensic techniques, and a reference manual jointly sponsored by the Federal Judicial Center and a task force of the Carnegie Commission on Science, Technology and Government. The manual will detail the range of techniques that judges have used to manage science and technology issues in litigation.

sound empirical study are similar across a broad range of scientific disciplines. Hence, judges who become familiar with one or more fields of science will be better equipped to deal with the diverse categories of scientific testimony presented in their courtrooms. Although each type of testimony will require fact-specific determinations, each evaluation should focus on the basic design of the underlying studies.

A. A Hypothetical Daubert *Analysis of Pasteur's Experiment*

Analysis under *Daubert* can be illustrated by asking whether Pasteur's compelling refutation of spontaneous generation would be admissible today. In this hypothetical, spontaneous generation is still widely believed to occur, but Pasteur is asked to testify in court that airborne germs probably account for the microbial growth seen in flasks of incubated fermentation broth. Pasteur would describe the experiment with the swan-necked flasks as support for his opinion. Since the work is too recent to have been published, peer review cannot lend credence to its soundness.

Thus, the judge should ask whether Pasteur had adequately ruled out other explanations for his observations. If Pasteur's control flasks were identical to the experimental ones save for their broken necks, the judge should conclude that alternative explanations have been adequately ruled out. A possible reason to exclude the evidence would be if the two groups of flasks were made of different kinds of glass, and growth-inhibiting chemicals had leached out of the swan-necked flasks, hence preventing the replication of any spontaneously-generated organisms that may have appeared in those flasks. But if all of the flasks were made from the same glass and no other flaws in experimental design become

---

Other task force projects include assembling rosters of experts for use by the judiciary, and generally fostering institutional ties between the judiciary and the scientific community. Moreover, the task force has recommended that scientific methodology be integrated into traditional training programs for judges, and that a national Science and Justice Council be established to facilitate developments in this area of law. In addition, the Federal Courts Study Committee has issued a set of recommendations for how judges could handle scientific issues in a report published in 1990, and the Judicial Conference of the United States is actively researching similar issues. Carnegie Commission on Science, Technology, and Government, *Science and Technology in Judicial Decision Making* (1993). In another joint project, the task force and the Federal Judicial Center have devised and tested a training program to be made available to all of the judiciary, and that focuses on problems arising from computer-generated evidence, recent developments regarding DNA evidence, questions regarding causation in toxic tort litigation, and case management techniques. Rorie Sherman, *"Junk Science" Rule Used Broadly,* Nat'l L.J., Oct. 4, 1993, at 3. In addition to these resources, the journal *Jurimetrics* deals routinely with the various intersections of law and science.

apparent, Pasteur's experiment should be accepted as having satisfied *Daubert's* testing requirement.

Pasteur's findings will not satisfy a general acceptance requirement because the hypothetical specifies that he has had no opportunity to publish. Even if he had, his skeptical colleagues would withhold judgment on the new theory until other scientists had verified its predictions. Under *Frye,* this lack of general acceptance would be fatal to Pasteur's testimony. While *Daubert* permits consideration of general acceptance, it explicitly forbids excluding testimony solely because general acceptance is lacking.

Under these circumstances, Pasteur's testimony should be viewed favorably provided the experimental method he used had a low rate of error. The judge should ask the frequency with which the expected result was obtained in several trials of the experiment. For example, if microbial growth appeared in the swan-necked flask in 15 out of 30 trials, the judge probably would conclude that Pasteur's sterilization procedure was so unreliable that the demonstration was not relevant to the issue of spontaneous generation. But if microbial growth were consistently and reproducibly absent from the swan-necked flasks, the judge should proceed to weigh the testimony's probative value against the likelihood of the jury being prejudiced by Pasteur's presentation.

As this hypothetical illustrates, the new standard corrects several deficits of the *Frye* test. Critics of *Frye* have long protested that automatic exclusion of novel testimony was unfair.[111] The Pasteur hypothetical illustrates how even novel testimony can receive a fair hearing under the *Daubert* standard. Moreover, the *Daubert* standard, unlike the *Frye* test, acknowledges that the courtroom is the judge's, not the scientist's, bailiwick. Communities of scientists rarely display the unified consensus envisioned by the *Frye* test, and scientists' reluctance to draw unqualified conclusions is at odds with the legal necessity to do just that within the short time span of a trial.

B. Scientific Evidence Should Be Evaluated at Three Levels

Although the range of scientific knowledge is broad, the following three-stage inquiry should suffice in most cases to elicit the information necessary for analyzing testimony under *Daubert.* First, is the technique or method itself scientifically valid? Second, is it valid to apply the technique to the specific factual matter at issue? Third, was the

111. See *supra* notes 17–18.

technique competently applied in this particular instance? For methodologies previously ascertained in court to be scientifically sound, judicial notice may suffice to answer the first two questions, but the third can be addressed only on a case-by-case basis. For novel evidence, judges should require that all three questions be answered satisfactorily before admitting the testimony.

In the first stage of this inquiry, the judge should appraise the validity of the basic methods or theories supporting the testimony. A sound method or theory is one that is based on scientific studies that addressed sources and rates of error, and in which the investigators ruled out alternative explanations for their observations.[112] The judge's inquiry should include ascertaining that the methodology did not employ unwarranted analogies.[113]

In the second part of the inquiry, the judge should establish that the studies underlying the testimony actually encompassed the specific application of the technique at issue. For example, the judge in *Bynum*[114] should have required the prosecution to demonstrate that gas chromatography was capable of reliably identifying and differentiating samples of cocaine, even though the reliability of gas chromatography had been established for general purposes.

For the third part of the inquiry, the judge should require experts to demonstrate that the instant application was proficiently performed, and that testing was conducted by qualified personnel. For example, judges could examine laboratory records of specific applications in order to ascertain that tests were performed in accordance with the proper procedures.

VI. CONCLUSION

The blossoming of modern science has increased the need for the judiciary to ensure the credibility of scientific evidence. The *Daubert* decision provides a requirement for empirical testing that will bolster judicial efforts to screen out unreliable scientific testimony while admitting all testimony that is relevant and helpful to the trier of fact. However, judges should apply *Daubert* only to fields of expertise that

112. *See, e.g.*, Kreiling, *supra* note 4, at 965–67 (discussing in general the format for scientific studies).

113. *See supra* note 83 and accompanying text (describing the court's refusal to ascribe the properties of vitamin A to Retin-A).

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

customarily are considered scientific. When the new standard is applicable, judges should ascertain not only that the expert's opinion is founded on adequately tested methods or theories, but also that those methods produce meaningful results when applied to the matter at issue. Judges should base admission decisions on precedents only when the earlier evaluations used criteria compatible with *Daubert*. By thus limiting the admission of scientific testimony, courts can effectively bar irrelevant scientific evidence without repressing demonstrably valid evidence even when it is novel and unpublished.