1994

The Meaning of Daubert and What that Means for Forensic Science

Randolph N. Jonakait
New York Law School, farrah.nagrampa@nyls.edu

Recommended Citation
Cardozo Law Review, Vol. 15 , Issue 6 (April 1994), pp. 2103-2118

THE MEANING OF *DAUBERT* AND WHAT THAT MEANS FOR FORENSIC SCIENCE

*Randolph N. Jonakait**

I. THE FLEXIBLE INQUIRY

Daubert's effect on forensic science is unclear because the opinion is unclear.[1] The *Daubert* Court did start blazing a useful path by stating that before scientific testimony is admitted, the trial court must be convinced that "the reasoning or methodology underlying the testimony is scientifically valid and . . . that [the] reasoning or methodology properly can be applied to the facts in issue."[2] But the Court's trailblazing was incomplete and often misleading.

For example, *Daubert* is premised on unarticulated assumptions. The opinion commands trial courts to determine whether something is "scientific," not whether it is physics, chemistry, biology, epidemiology, psychology, accidentology, clinical ecology, or forensic science. This can be done only if there are general standards and methods applicable to all fields of science that distinguish genuine science from pseudoscience.[3] Furthermore, the Court's command can only be followed if trial courts can understand those standards and use them to identify real science.[4] These premises, however, were not stated. It

---

* Professor of Law, New York Law School.

[1] *See* Daubert v. Merrell Dow Pharmaceuticals, Inc., 113 S. Ct. 2786 (1993).

[2] *Id.* at 2796.

[3] *See* Lee Loevinger, *Science and Legal Rules of Evidence: A Review of* Galileo's Revenge: Junk Science in the Courtroom, 32 JURIMETRICS J. 487, 500 (1992) (book review) ("While there are innumerable specialized fields in science today, and while knowledge in one field does not necessarily transfer to another field, there are, nevertheless, general standards applicable to all fields of science that distinguish genuine science from pseudo-science and quack science."); *cf.* W.I.B. BEVERIDGE, THE ART OF SCIENTIFIC INVESTIGATION 19 (Vintage Books 1950):

> Science as we know it to-day [sic] may be said to date from the introduction of the experimental method during the Renaissance. Nevertheless, important as experimentation is in most branches of science, it is not appropriate to all types of research. It is not used, for instance, in descriptive biology, observational ecology or in most forms of clinical research in medicine. However, investigations of this latter type make use of many of the same principles. The main difference is that hypotheses are tested by the collection of information from phenomena which occur naturally instead of those that are made to take place under experimental conditions.

[4] *Cf.* Bert Black & John A. Singer, *From* Frye *to* Daubert: *A New Test for Scientific Evidence*, 1 SHEPARD'S EXPERT & SCI. EVIDENCE Q. 19, 27 (1993):

> If lawyers and judges hope to apply the new *Daubert* test rationally, they will thus have to learn more about what distinguishes science from other forms of knowledge—what it is that makes science scientific. Science defies precise defini-

would have been better if they had been to help insure that trial courts would begin their analyses at the proper starting point.

Even if, however, trial courts start at the same place, it is unlikely they will follow the same course. The opinion fails to provide meaningful guidance on how to follow the path; no firm method for making the determination was given. On one level, *Daubert* gave pronouncements, such as " 'scientific' implies a grounding in the methods and procedures of science,"[5] and the "overarching subject is the scientific validity—and thus the evidentiary relevance and reliability—of the principles that underlie a proposed submission."[6] These generalities were followed by a mandate to undertake a "flexible" inquiry.[7] This second level, which should have outlined a clear trail, however, was muddy. The result may be that *Daubert* has little effect on future litigation. That, of course, is often the result of such flexible tests.

The wonder of flexible inquiries is that they can so often be made to fit just about any predetermined conclusion—practically any two-month-old can be manipulated into an infant's clothing. The skeptic may feel that flexible standards seldom produce results; instead, they are devices merely used to justify conclusions reached on other grounds. Initial reactions to *Daubert* indicate that its flexible inquiry is meaningless. As this symposium reflects, all sides claim victory. Such universal celebration indicates that all believe they can favorably apply the announced standard to a given situation.

The real question, however, is not whether advocates and academics can interpret *Daubert* to justify the results they desire, but whether judges will do the same. Will the opinion truly drive decisions, or will judges continue to reach the same conclusions, but now simply justify them differently? While it is too early to reach a definitive conclusion,[8] the flexible exploration mandated by the Court is so vague and inadequate that it is difficult to envision how trial court discretion will be meaningfully cabined. Although *Daubert* suggested that trial courts examine four factors,[9] it was only a suggestion.[10]

---

tion, but even nonscientists can learn enough about it to distinguish valid from invalid claims.

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

6 *Id.* at 2797.

7 *Id.*

8 In the cases concerning scientific evidence since *Daubert* was decided, none that I have looked at indicate a different result was reached than would have been without *Daubert*. So far, at least, *Daubert* has compelled no new results. The records in these cases, however, were all apparently made before the decision. The true effect of *Daubert* cannot be assessed until more time passes.

9 *See Daubert*, 113 S. Ct. 2796-97 (listing factors including testability, peer review and publication, general acceptance, and error rate).

More important, *Daubert*'s guidance on how to use the listed factors is not only inadequate, it is sometimes downright misleading.

Daubert was most enlightening in its discussion regarding peer review and publication.[11] The Court suggested how this factor relates to a trial court's reliability analysis. "[S]ubmission to the scrutiny of the scientific community is a component of 'good science,' in part because it increases the likelihood that substantive flaws in methodology will be detected."[12] Peer review and publication, then, are not important in themselves, but in what they reveal about the likelihood of methodological flaws having been detected and, presumably, corrected. Consequently, even if the science has been peer reviewed and published, if these processes were unlikely to lead to the discovery and correction of problems, the reliability of the science is suspect. Conversely, if the trial court becomes convinced that such detection and alteration have occurred even without the scientific community's scrutiny, the testimony still could be admitted.[13]

While the Court gave some meaningful explication of the peer review and publication prong, its discussion of other aspects of the flexible inquiry was much less satisfactory. For example, *Daubert* suggests that trial courts use the *Frye*[14] test once again as part of its rubbery examination.[15] The Court states that general acceptance is not necessary for a reliability assessment, but that such acceptance "can be an important factor in ruling particular evidence admissible" and the lack of it "may properly be viewed with skepticism."[16] The Court's discussion, however, does not address problems that have bedeviled past applications of *Frye*. How widespread must the acceptance be for it to be "general"? How should a court go about defining the relevant scientific community? More important, what does it mean that a lack of general acceptance "may" allow a trial court to view the scientific evidence with skepticism? If lack of general accept-

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

11 *Id.* at 2797 (The trial court should consider "whether the theory or technique has been subjected to peer review and publication.").

12 *Id.*

13 While *Daubert* indicated that publication was not a necessary ingredient for a determination of reliability, the Court suggested limited circumstances for when it might be expected that the science would not be disseminated: "Some propositions . . . are too particular, too new, or of too limited interest to be published." *Id.*

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

15 *Daubert*, 113 S. Ct. at 2797 (" '[G]eneral acceptance' can yet have a bearing on the inquiry."); *see* Maiorana v. National Gypsum Co., 827 F. Supp. 1014, 1033 (S.D.N.Y. 1993) ("The decision in *Daubert* kills *Frye* and then resurrects its ghost.").

16 *Daubert*, 113 S. Ct. at 2797.

ance leads to skepticism in some instances, why not in all? If not in all, why in some and not others?

The opinion also listed an "error rate factor" as part of the flexible inquiry. "[I]n the case of a particular scientific technique, the court ordinarily should consider the known or potential rate of error, and the existence and maintenance of standards controlling the technique's operation."[17] Once again, however, crucial questions were not addressed. For example, what if the error rate is unknown? Does it matter if it is ascertainable, but no one has bothered to ascertain it? What does it mean for the reliability of a scientific technique if its error rate is not knowable? If the error rate is known, does it matter? If the error rate is less than fifty percent, does it satisfy a preponderance of the evidence notion of reliability? Or does the error rate have to be small enough to conform to "scientific" notions of confidence? Is there a connection between error rates and the statistical tests that normally require scientists to reach a ninety-five percent confidence level?[18]

Perhaps most disturbing, however, is the Court's treatment of the first factor discussed. "Ordinarily, a key question to be answered in determining whether a theory or technique is scientific knowledge that will assist the trier of fact will be whether it can be (and has been) tested."[19] Strikingly strange here is the introductory word, "ordinarily," which implies that in some extraordinary situation something can be "scientific" even though it cannot be (or has not been) tested. What those possible unusual circumstances might be are left undefined. The Court does cite three references,[20] but they do not clarify

---

[17] *Id.* (citation omitted). An inquiry about the existence and maintenance of standards is not likely to be meaningful if trial courts follow the Supreme Court's lead here. *See* Paul C. Giannelli, Daubert: *Interpreting the Federal Rules of Evidence*, 15 CARDOZO L. REV. 1999, 2022 (1994) (criticizing *Daubert*'s reliance on the voice print cases).

[18] The Court gives no indication how the normal statistical tests of science should affect admissibility. *Cf.* CARNEGIE COMM'N ON SCIENCE, TECHNOLOGY, AND GOV'T, SCIENCE AND TECHNOLOGY IN JUDICIAL DECISION MAKING: CREATING OPPORTUNITIES AND MEETING CHALLENGES 37 (1993) ("[M]ust all scientific studies on which an expert relies meet the 95 percent confidence level that is often used by scientists to reject the possibility that chance alone accounted for observed differences? Is a lower standard compatible with the objectives of the preponderance of proof standard in civil litigation?").

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

[20] *See id.* at 2796-97. The Court first cited 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, 645 (1992) ("Scientific methodology today is based on generating hypotheses and testing them to see if they can be falsified; indeed, this methodology is what distinguishes science from other fields of human inquiry."). The Court then suggested that the reader also see CARL G. HEMPEL, PHILOSOPHY OF NATURAL SCIENCE 49 (1966) ("[T]he statements constituting a scientific explanation must be capable of empirical test.") and KARL R. POPPER, CONJECTURES AND REFUTATIONS: THE GROWTH OF SCIENTIFIC

the Court's language. Indeed, they contradict it. The sources indicate that scientific methodology is always tested, not that it ordinarily is.[21] And that is what the Court should have recognized.

II. What *Daubert* Ought to Mean

A. *Testability Factor*

Daubert treats falsifiability or testability as a factor similar to the others in its flexible inquiry. The Court was wrong. The defining touchstone of science is a testable proposition that is tested. To be a scientist requires finding ways to test hypotheses that are generated; otherwise the "science" is just an exercise in fantasy.[22]

In other words, scientists seeking to advance knowledge about the empirical world do not rely on assertions or convictions or logic, but rather upon encounters with the empirical world.[23] Scientists do this not just by creating testable or falsifiable propositions, but also by analyzing and testing them.[24] Most scientific hypotheses are proven to be wrong.[25] Their creation is essential, but their rigorous testing is even more crucial.[26]

The analysis and testing is not limited to one form,[27] but the pro-

---

KNOWLEDGE 37 (5th ed. 1989) ("[T]he criterion of the scientific status of a theory is its falsifiability, or refutability, or testability.").

21 *See Daubert*, 113 S. Ct. at 2796-97.

22 *See* PHILIP KITCHER, THE ADVANCEMENT OF SCIENCE: SCIENCE WITHOUT LEGEND, OBJECTIVITY WITHOUT ILLUSIONS 33 (1993) ("In attempting to instantiate the Darwinian schemata, biologists are compelled to advance hypotheses about the historical development of life, and it is incumbent on them to specify ways of testing these hypotheses . . . if they are to avoid the charge that evolutionary biology is simply an exercise in fantasizing."); *see also* BEVERIDGE, *supra* note 3, at 63 ("Hypothesis is the most important mental technique of the investigator, and its main function is to suggest new experiments or new observations.").

23 *Cf.* KITCHER, *supra* note 22, at 306 ("[W]e work our way free of the mistakes of earlier [scientific] generations through further encounters with nature.").

24 Sir Karl Popper, distinguished philosopher of science, explained: "A scientist . . . puts forward statements, or systems of statements, and tests them step by step. In the field of the empirical sciences, . . . he constructs hypotheses, or systems of theories, and tests them against experience by observation and experiment." KARL POPPER, THE LOGIC OF SCIENTIFIC DISCOVERY 27 (Harper Torchbooks 1965); *see also* Bert Black, *A Unified Theory of Scientific Evidence*, 56 FORDHAM L. REV. 595, 622 (1988) (quoting Frederick Suppe, *Afterword* to THE STRUCTURE OF SCIENTIFIC THEORIES 650 (Frederick Suppe ed., 2d ed. 1977)) ("The goal for scientists is 'to obtain systematic knowledge that provides understanding of the world we live in.' Scientists evaluate their ideas against criteria that involve testability, objectivity, impartiality, and a belief in a deep and obvious connection between evidence and reason.").

25 BEVERIDGE, *supra* note 3, at 79-80 ("Most hypotheses prove to be wrong whatever their origin may be. . . . W. H. George points out that even with men of genius, with whom the birth rate of hypotheses is very high, it only just manages to exceed the death rate.").

26 *Id.* at 80 ("The productive research worker is usually one who is not afraid to venture and risk going astray, but who makes a rigorous test for error before reporting his findings.").

27 John Veilleux, Note, *The Scientific Model in Law*, 75 GEO. L.J. 1967, 1970 n.16 (1987)

cess is always dependent on skepticism and doubt—a skepticism that promotes inquiry, experimentation, and validation that can remove that doubt.[28] The scientist looks at hypotheses and determines how the ideas could be proven wrong by reproducible experiments or studies.[29] Only after the possible shortcomings of the assertion are tested and the hypothesis has not been proven false does a scientific idea emerge.[30]

("Of course, there is no *one* scientific method. The techniques used to develop and test hypotheses necessarily vary for different disciplines.").

[28] *Cf.* Dirk Eshleman, Note, *Different Standards and Conflicting Results: A Re-Evaluation of the* Frye *Test for Admitting Novel Scientific Evidence in Light of Decisions Involving Spectrographic Evidence Introduction*, 5 REV. LITIG. 327, 364 (1986) ("Science thrives on doubt. Scientific skepticism promotes inquiry and experimentation, and, hence, advances our understanding of the world around us.").

[29] *See* POPPER, *supra* note 24, at 40-41.

> Theories are . . . *never* empirically verifiable. . . . [But a system is] empirical or scientific only if it is capable of being *tested* by experience. These considerations suggest that not the *verifiability* but the *falsifiability* of a system is to be taken as criterion of demarcation. . . . *[I]t must be possible for an empirical scientific system to be refuted by experience.*

Id. (emphasis in original) (footnotes omitted); *see also* H.J. Walls, *Whither Forensic Science?*, 6 MED. SCI. & L. 183, 187 (1966):

> Modern philosophers of science . . . have pointed out that a scientific hypothesis can, logically, never be proved, only disproved. However often the result of an experiment supports an hypothesis, we cannot be logically certain that some unknown and unexpected cause will not produce a contrary result the next time, and if that does happen, and if the unexpected result cannot be explained in some way consistent with the hypothesis, then the hypothesis is in strict logic disproved. If all the evidence supports it, it is provisionally accepted, but . . . it is never free from the risk of being upset by one reliable incompatible observation. . . . That . . . is why the scientific world picture . . . is always provisional, never final, always changing.

Cf. BEVERIDGE, *supra* note 3, at 118 ("Generalisations can never be *proved*.").

In addition, experiments must be reproducible to be part of scientific proof.

> Only when certain events recur in accordance with rules or regularities, as is the case with repeatable experiments, can our observations be tested—in principle—by anyone. We do not take even our own observations . . . until we have repeated and tested them. Only by such repetitions can we convince ourselves that we are not dealing with a mere isolated "coincidence" . . . .

POPPER, *supra* note 24, at 45; *see also* BEVERIDGE, *supra* note 3, at 23 ("The essence of any satisfactory experiment is that it should be reproducible."); *cf.* Lee Loevinger, *Standards of Proof in Science and Law*, 32 JURIMETRICS J. 323, 342 (1992):

> Science has the advantage in this respect of dealing generally with events that, being replicable, occur in series of like events, thus justifying the expression of judgments in statistical terms. Law, on the other hand, deals with unique cases that are not subject to frequency calculations.

[30] *See* Steven Goldberg, *The Reluctant Embrace: Law and Science in America*, 75 GEO. L.J. 1341, 1342-43 (1987).

> A scientist can come up with a hypothesis about the natural world through any process at all—systematic study, inspired speculation, or fevered dreams. But that hypothesis must ultimately be subject to controlled tests, reproducible by others. Only if the tests support the hypothesis can the hypothesis be accepted.

Daubert's flexible framework fails, then, by suggesting that testability or falsifiability is not an absolute. The Court should have made clear that an assertion is science if, and only if, it can be tested or falsified.[31] Furthermore, a testable proposition can be relied upon if, and only if, it has been rigorously tested.[32] Or, as the Nobel laureate Luis Alvarez cautioned fellow scientists, "[o]nly trust what you can prove."[33]

This rigorous testing has a basic rule. To be scientific, controls must be employed.

> This is the most fundamental commandment in the canon of experimental technique. To reach an unimpeachable conclusion

---

Id.; *see also* Black, *supra* note 24, at 623 (quoting Frederick Suppe, *Afterword* to THE STRUCTURE OF SCIENTIFIC THEORIES 706 (Frederick Suppe ed., 2d ed. 1977)):

> Observation and experimentation are used to find shortcomings, to determine how to make improvements, and "to discover how to eliminate known artificialities, distortions, oversimplifications, and errors in the descriptions, explanations, and predictions of reality that the theory affords." Only after a theory has survived a period of this kind of testing, review and refinement can it be used without significant questions, and even then, it remains open to renewed doubt.

Cf. POPPER, *supra* note 24, at 42:

> [W]hat characterizes the empirical method is its manner of exposing to falsification, in every conceivable way, the system to be tested. Its aim is not to save the lives of untenable systems but, on the contrary, to select the one which is by comparison the fittest, by exposing them all to the fiercest struggle for survival.

Cf. State v. Davis, 742 P.2d 1356, 1361 (Ariz. 1987) (rejecting the testimony of a graphologist who formed opinions about personality and mental states from handwriting and concluded that, "graphology is not a science because its results are neither verifiable nor repeatable").

31 Chief Justice Rehnquist expressed bewilderment as to the concept of "falsifiability." "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." *Daubert*, 113 S. Ct. at 2800 (Rehnquist, C.J., concurring in part and dissenting in part).

Falsifiability, however, is not so alien to a trial attorney. It is akin to the kind of thinking in which many lawyers who deal with factual disputes routinely engage. For example, if a client tells his attorney that he could not have committed the crime because he was at his regular post office job, the good attorney immediately tries to determine what type of information might tend to corroborate this assertion. If what the client says is true, there ought to be employment records, pay stubs, recollections of coworkers and carpoolers, and the like. If such information does not support the alibi, the defense attorney knows that the prosecution will have a greater chance of convincing the jury that the alibi is false. In other words, the attorney looks for information that tends to corroborate or falsify the claimed defense. In a sense, trial attorneys regularly create falsifiable hypotheses (if the alibi is correct, then a time card should show the client at work), and then look for the information that might prove the hypothesis false.

32 Peter H. Schuck, *Multi-Culturalism Redux: Science, Law, and Politics*, 11 YALE L. & POL'Y REV. 1, 16 (1993) ("[S]cientists subscribe to and are actuated by rigorous standards of empirical investigation and proof; to deviate from these standards is to be deemed professionally incompetent, or worse.").

33 GARY TAUBES, BAD SCIENCE: THE SHORT LIFE AND WEIRD TIMES OF COLD FUSION 197 (1993) (quoting Luis Alvarez).

> establishing the cause of an effect, run controls. E. Bright Wilson, Jr., in his classic 1952 volume *An Introduction to Scientific Research*, described controls as "similar test specimens which are subjected to as nearly as possible the same treatment as the objects of the experiment, except for the change in the variable under study."[34]

Without controls, the cause of an outcome cannot be determined. As E. Bright Wilson also said: "If one doubts the necessity for controls, reflect on the statement: 'It has been conclusively demonstrated by hundreds of experiments that the beating of tom-toms will restore the sun after an eclipse.' "[35] While the proper methodology of controlled experiments, tests, or inquiries may vary from field to scientific field, such controlled experiments, tests, or inquiries are required if the endeavor is to be considered science.[36]

Since *Daubert* commands trial courts to determine that scientific testimony is indeed scientific and, therefore, based on a reliable methodology, then testability and testing with rigorous, controlled experiments are not merely factors in a balancing task. They are absolutes. Without them, it is not science. *Daubert*'s approach, however, to the other factors was basically correct. Those factors are not essential, but generally can shed light on whether the testimony is based on trustworthy methods.

B. *Peer Review and Publication Factor*

Peer review and publication can help determine whether the sci-

---

34 *Id.* at 120-21 (footnote omitted); *see also* BEVERIDGE, *supra* note 3, at 20:

> The "controlled experiment" is one of the most important concepts in biological experimentation. In this there are two or more similar groups[;] . . . one, the "control" group, is held as a standard for comparison, while the other, the "test" group, is subjected to some procedure whose effect one wishes to determine. The groups are usually formed by "randomisation," that is to say, by assigning individuals to one group or the other by drawing lots or by some other means that does not involve human discrimination. The traditional method of experimentation is to have the groups as similar as possible in all respects except in the one variable factor under investigation, and to keep the experiment simple.

Compare John Monahan & Laurens Walker, *Social Authority: Obtaining, Evaluating, and Establishing Social Science in Law*, 134 U. PA. L. REV. 477, 502 (1986):

> There is consensus among social scientists of all disciplines that research must possess "validity." That is, the methods used in research must be able to justify the conclusions drawn by the investigator. . . . To have "high" validity, a study must rule out, or "control for," competing hypotheses that may account for an observed state of affairs.

35 TAUBES, *supra* note 33, at 162.

36 *Cf.* BEVERIDGE, *supra* note 3, at 25 ("Unless the basic needs of the controlled experiment can be satisfied it is better to abandon the attempt. . . . Most of the experiments have proved nothing [when] the controls were not strictly comparable.").

ence has been adequately validated. As the *Daubert* Court recognized, "the scrutiny of the scientific community . . . increases the likelihood that substantive flaws in methodology will be detected."[37] The investigating scientist herself can seldom, if ever, be confident that she has thought of all the possible alternative causes for the observed phenomenon or properly tested for those possibilities.[38] Two heads are better than one; a collection of scientific heads are even better. The scientific community, with its greater collective wisdom, knowledge, and experience can often suggest alternative causes and experiments that have not occurred to the investigator doing the initial research.

Peer review and publication increase the likelihood that many knowledgeable, skeptical, and probing scientists have examined the data and conclusions. Without this scientific examination, it is almost impossible to conclude that a hypothesis has been adequately tested. A trial court should consider science that has not been peer reviewed and published as science whose methodology has probably not been scrutinized closely and, therefore, whose reliability has not been established. But rigorous, empirical validation is the touchstone—not peer review and publication. A court, then, can still find the science reliable if, by some unusual circumstance, the scientific community has still scrutinized the methodology or propositions even though the science has not entered the normal streams of scientific intercourse.

On the other hand, a court should not leap to the conclusion that because the science has been peer reviewed or published, it has been rigorously scrutinized. Peer review is hardly a perfect system—it is often less than demanding because scientists are busy or because of conflicts of interest.[39]

Furthermore, strict scrutiny does not necessarily follow from the fact of publication. Too much scientific literature is published for scientists to scrutinize most of it. As the philosopher of science David Hull noted:

> [S]cientists spend very little time in haphazard reading of the literature, even the literature in their own area of expertise. There is much too much published for that to be a productive activity. In-

[37] *Daubert*, 113 S. Ct. at 2797.

[38] *Cf.* David E. Bernstein & Peter W. Huber, *Defense Perspective*, 1 SHEPARD'S EXPERT & SCI. EVIDENCE Q. 59, 61 (1993) ("[A]n expert should only be permitted to identify an agent as the cause of an injury or illness under Rule 403 if he or she has studied and discounted alternative causes.").

[39] *Cf.* TAUBES, *supra* note 33, at 14 ("It is considered in poor taste, to point out that the experimental techniques of a competitor leave something to be desired. That would also be inviting the competitor to return the favor someday."); *see also id.* at 35.

stead they scan the literature looking for findings that bear on their own research.[40]

For these and other reasons, the work of others is often not checked, and some suggest that only a fraction of verifications or refutations are meaningful.

> Scientists cannot spend very much time checking the work of other scientists if they themselves are to make contributions. They reserve checking for those findings that bear most closely on their own research, chiefly those that threaten it. Because different scientists are committed to different views, the checking that goes on in science rarely degenerates into empty show. When scientists refute their own favorite hypotheses or their opponents confirm them, one can place considerable confidence in the results.[41]

Peer review and publication, then, can bear on the crucial question of whether the opinion's foundation is based on testable propositions that have been adequately tested. Only in extraordinary situations is it possible to conclude that sufficient testing has occurred in the absence of peer review and publication. The existence of peer review and publication, however, do not necessarily mean that the technique or theory has been rigorously scrutinized by the scientific community.

C. *General Acceptance Factor*

Daubert's general acceptance factor, which overlaps significantly

---

[40] DAVID L. HULL, SCIENCE AS A PROCESS: AN EVOLUTIONARY ACCOUNT OF THE SOCIAL AND CONCEPTUAL DEVELOPMENT OF SCIENCE 348 (1988). The inability to stay abreast of the literature is not merely a recent scientific phenomenon. Four decades ago, W.I.B. Beveridge calculated that scientific periodicals were publishing "two million articles a year, or 40,000 a week, which reveals the utter impossibility of keeping abreast of more than the small fraction of the literature which is most pertinent to one's interest." BEVERIDGE, *supra* note 3, at 3.

[41] HULL, *supra* note 40, at 394. *See also* TAUBES, *supra* note 33, at 426:

> What cold fusion had proven, nonetheless, was that the nonexistence of a phenomenon is by no means a fatal impediment to continued research. As long as financial support could be found, the research would continue. And that support might always be found so long as the researchers could obtain positive results. In fact, the few researchers still working in the field would have little incentive to acknowledge negative results as valid, because such recognition would only cut off their funds.

Cf. BEVERIDGE, *supra* note 3, at 65:

> It is the originator who gets both the personal satisfaction and most of the credit if his idea is proved correct, even if he does not do the work himself. A man working on an hypothesis which is not his own often abandons it after one or two unsuccessful attempts because he lacks the strong desire to confirm it which is necessary to drive him to give it a thorough trial and think out all the possible ways of varying the conditions of the experiment.

with peer review and publication, can play a similar role. The extent of a scientific consensus can help determine how thoroughly the testable propositions have been tested.

Science may become generally accepted because it has been published, scrutinized by scientists, and found trustworthy. If so, general acceptance speaks to a reliable methodology. But not all generally accepted science falls into this category. Some science becomes accepted without publication and without significant scrutiny by the scientific community. Information is passed from mentor to mentee or between collaborators or through electronic mail. Methodologies distributed in these ways may become accepted without real communal scrutiny. For example, a scientist may perform a particular technique in a particular manner because she learned to do it that way while a postdoctoral fellow. Another scientist may conduct an experiment in a certain manner because a coauthor does it that way. The scientist may be accepting the methodology not because she has closely examined and tested it, but instead because of the perceived authoritative nature of its source.[42]

This kind of general acceptance, however, can also be an indicator that a methodology is reliable, but only if the inherited technique is actually used by many scientists, and it is the kind of technique whose flaws would be revealed by widespread use.[43] If such a technique, although seemingly accepted by many, is used by only a few or its use is not likely to reveal its defects, then general acceptance of

---

[42] *Cf.* KITCHER, *supra* note 22, at 306:

> [W]e can distinguish three ways in which [reliance on authority] permeates the cognitive lives of scientists. First, there is the general epistemic dependence on the past that figures in everyone's early intellectual ontogeny. . . . Second, at the time of entry into the scientific community, novices endorse a communitywide conception of legitimate epistemic authority. Certain people are to be trusted to decide on certain issues, and the novice must accept whatever agreements they reach on those issues. Third, during the course of individual research, scientists interact with one another, adopting the claims made by *some* of their colleagues, investigating the proposals of others, ignoring the suggestions of yet others, when the claims, proposals, and suggestions in question go beyond what is agreed upon by the pertinent community.

[43] *See* Randolph N. Jonakait, *Will Blood Tell? Genetic Markers in Criminal Cases*, 31 EMORY L.J. 833, 846-52 (1982).

> [A] new scientific procedure will normally go through a three-step process. . . . Suppose, for example, that some types of blindness are caused by tiny blood vessels in the eyes wearing out and bursting. . . . A scientist . . . [w]orking on animals . . . develops a laser surgery technique that he believes can stop the bleeding without doing damage to the rest of the eyeball. He publishes his results.
>
> Other eye specialists reading this report may be impressed by it, agree with the theory behind it, and understand how the new procedure should alleviate the condition. They would not yet, however, accept the procedure as reliable. Instead, verification through the use of controlled studies would next be required . . .

such a technique says nothing significant about the methodology's reliability.

While general acceptance may indicate that the science has been tested, lack of general acceptance has no singular meaning. The absence of acceptance may mean the scientific community has examined the science and found it wanting. If so, opinions based on the science should not be admitted. On the other hand, consensus may be absent because the science has not been disseminated to the broader commu-

---

> to see if the procedure worked the cure as claimed; and . . . to see if any limitations, such as side-effects, could be discovered. . . .
>
> If the new procedure passed the controlled studies, . . . [t]he ultimate test would be the actual use of the technique in general practice. Any limitations that were not unearthed in the studies would surface as the procedure was used under various circumstances. It might eventually be learned that while the surgery at first seem to be successful with diabetics, after a year or so the blood vessels burst again causing greater problems than before. Certainly, only when a procedure has gone through these three steps—development of the methodology, verification of the claimed results, and actual employment of the new technique—can the community of concerned scientists know both the procedure's reliability and its limitations.

Id. at 846-47; *see also* MICHAEL J. SAKS & RICHARD VAN DUIZEND, THE USE OF SCIENTIFIC EVIDENCE IN LITIGATION 74 (1983) ("A scientific or professional field has a powerful defense against unsubstantiated or exaggerated conclusions: replication. Professional/technical fields that apply such knowledge have a 'defense': erroneous principles will be found out because patients will not get well, planes will not fly, or chemicals will not be synthesized."); Patricia K. Woolf, *Deception in Scientific Research*, 29 JURIMETRICS J. 67, 82 (1988):

> Replication of specific and detailed experiments is unlikely to repeat exactly what was done before for well-known reasons: there is no professional incentive for a "me-too" experiment. However, if replication is, as it should be, broadly conceived to include the comparability of data from different laboratories, and building one experiment on the results and implications of a previous one, then "replication" too has played a role in detecting apparent research fraud.

See also Thomas S. Burack, Note, *Of Reliable Science: Scientific Peer Review, Federal Regulatory Agencies, and the Courts*, 7 VA. J. NAT. RESOURCES L. 27, 31-32 (1987) (quoting Stevan Harnad, *Creative Disagreement*, 19 SCIENCES 18, 18 (1979)):

> Peer interaction, in the form of repeating and building upon one another's experiments, testing and elaborating one another's theories, and evaluating and criticizing one another's research, is the real medium for the self-corrective aspect of science.

Cf. Peter J. Neufeld, *Admissibility of New or Novel Scientific Evidence in Criminal Cases, in* BANBURY REPORT 32: DNA TECHNOLOGY AND FORENSIC SCIENCE 80 (J. Ballantyne et al. eds., 1989):

> When a new nonforensic scientific hypothesis, instrument, or procedure is developed, it is first internally examined by blinded protocol. Once the creators are satisfied with their data, the data are submitted for publication in peer review journals. . . . Once published, the concept is debated by others in the relevant fields. Although other laboratories may attempt to replicate the results, this is not always the case. Often the initial premise is relied upon to move into related areas of experimentation. However, if there are any flaws or limitations in the initial findings, they will inevitably be detected as the technique's use becomes widespread in other controlled experiments.

nity or because it is too new or too specific to have spread widely. If so, the testimony can conceivably be admitted under *Daubert*, but only if the trial court is convinced that any flaws in the methodology have been exposed and corrected.

While general acceptance and peer review and publication may help determine whether testimony is based on reliable science, another *Daubert* factor hardly does. That "the court ordinarily should consider . . . the existence and maintenance of standards controlling the technique's operation"[44] makes little sense, at least in the opaque way presented by the Supreme Court. Abstract "standards" reveal nothing about whether the science has been adequately tested. Instead, they are only of importance if they have a source in rigorous testing. Standards must derive from testing to be meaningful; their mere existence does not indicate anything about the adequacy of the testing.

Scientifically sound standards, however, can help determine whether the particular use of a technique was done in a scientific manner. If the standards were derived from good science, then the court should ask whether the standards have been followed. If not, then the expert is unlikely to have based his testimony on a reliable methodology.

D. *Error Rate Factor*

Finally, *Daubert*'s error rate factor should be a crucial component in the analysis of scientific evidence. No scientific test is perfect. The potential for error exists in any measurement. Meaningful scrutiny of a technique should normally result in information about its error rate. A technique with an unknown error rate in all likelihood is a technique that has not been adequately tested.[45] And if not adequately tested, testimony based on it should not be admitted. Consequently, the lack of error rate information indicates that the scientific evidence should not be admitted.

Not all error rate data, however, should carry the same weight. The information should be collected from the technique in actual use.

---

44 *Daubert*, 113 S. Ct. at 2797.

45 *Cf.* TAUBES, *supra* note 33, at 321:

> [I]n science . . . the experimental measurement has not been invented that doesn't come with some measurement error. Calculating and presenting the limits of error in an experiment is usually taught in high school science classes. As E. Bright Wilson put it bluntly in *An Introduction to Scientific Research*, "A measurement whose accuracy is completely unknown has no use whatever. It is therefore necessary to know how to estimate the reliability of experimental data and how to convey this information to others."

A theoretical or potential error rate has little meaning. Unfortunately, the *Daubert* Court did not seem to understand this fact since it referred to examination of "potential rate[s] of error . . . ."[46] A reliable scientific methodology is not one that theoretically works well; it is one that actually works well. Knowledge of error rates empirically derived from a technique's practical use is crucial. Repeatedly, laboratory testing of a technique has indicated a flawless procedure while real world uses have shown significant defects.[47] The meaningful error rate is the one that has been derived from a rigorous study of the technique's actual use, and the *Daubert* Court's indication that a trial court should rely on theoretical rates is not helpful.

In sum, since the proponent of scientific testimony must establish that the testimony is based on a reliable scientific methodology, a court must determine whether that testimony is based on testable propositions that have been rigorously tested. Tests with controls are the norm. Peer review, publication, general acceptance, and error rates can help a court determine whether the science has a solid empirical base, but such factors should be examined only for what light they shed on the controlling issue—whether the testable propositions have been rigorously tested.

III. RELIABLE SCIENCE AND FORENSIC SCIENCE

I have examined elsewhere problems of quality in forensic sci-

---

46 *Daubert*, 113 S. Ct. at 2797.

47 The fight to conquer tuberculosis demonstrated that treatments showing promise in the laboratory did not always realize their potential in the real world. *See* FRANK RYAN, THE FORGOTTEN PLAGUE: HOW THE BATTLE AGAINST TUBERCULOSIS WAS WON—AND LOST (1993). For example, the initial trials for streptomycin showed a death rate one-third of that resulting from a sanitorium regime.

> But when another group of doctors analysed those same patients some five years later, [they found that t]hirty-five of the original fifty-two patients denied the drug were dead . . . . More disturbing was the fact that thirty-two of the original fifty-five patients treated with streptomycin were also dead.
>
> The conclusion was inescapable: few of the patients treated with streptomycin had been cured by the drug. What it had achieved was a temporary reduction in the severity of the disease.

Id. at 327 (footnote omitted). *Compare Symposium on Science and the Rules of Evidence*, 99 F.R.D. 187, 228 (1983) (comments of Renee Fox):

> [A wonder drug's] initial clinical trials and use typically receive great attention. Its therapeutic efficacy and promise are stressed, with side effects and limitations receiving scant attention. The next stage is characterized by increased awareness of the agent's side effects and limitations, perhaps in a larger number of patients with a wider range of attributes than initially suspected. The sequence often advances to the stage where more controlled doses of the drug are prescribed, for particular patients under specified circumstances, to limit the possible side effects.

ence.[48] That examination demonstrated that little or no meaningful testing has ever been performed on many forensic science procedures.[49] Little is also known about the true error rates for almost all forensic science techniques. The few disclosed error rates, however, are shockingly high.[50] Most of forensic science operates outside of the peer review systems, and forensic science is seldom published.[51] While forensic science techniques are accepted in forensic science, many are not accepted by a broader scientific community.[52] Furthermore, the techniques accepted in forensic science are not used in such a way that would reveal their methodological flaws, if any.[53]

In other words, if *Daubert* is taken seriously, then much of forensic science is in serious trouble.

---

48 *See* Randolph N. Jonakait, *Forensic Science: The Need for Regulation*, 4 HARV. J.L. & TECH. 109 (1991).

49 *Id.* at 137-48.

50 *Id.* at 109-24.

51 *Id.* at 133-34.

52 *Id.* at 148-50.

53 *Id.* at 151-54.