

# SCIENTIFIC MISCONCEPTIONS AMONG *DAUBERT* GATEKEEPERS: THE NEED FOR REFORM OF EXPERT REVIEW PROCEDURES

JAN BEYEA\* AND DANIEL BERGER\*\*

## I

### INTRODUCTION

The Supreme Court's opinions in *Daubert v. Merrell Dow Pharmaceuticals, Inc.*<sup>1</sup>, *General Electric Company v. Joiner*<sup>2</sup>, and *Kumho Tire v. Carmichael*<sup>3</sup> contain two inconsistent views of science. In some places, the Court views science as an imperfect "process" for refining theories, whereas in other places, the Court views science as universal knowledge derived through "formal logic." The latter view, long out of favor with philosophers and historians of science, comports with the current cultural vision of science and is likely to be adopted by district and appeals court judges, without vigorous "education," or until such time as higher courts recognize that the two views need to be synthesized into a

---

Copyright © 2001 by Jan Beyea and Daniel Berger

This article is also available at <http://www.law.duke.edu/journals/64LCPBeyea>.

\* Ph.D., Senior Scientist, Consulting in the Public Interest, Lambertville, New Jersey.

Dr. Beyea is a regular panel member and reviewer of studies carried out by the National Research Council of the National Academy of Sciences. He can be contacted at [jbeyea@cipi.com](mailto:jbeyea@cipi.com).

\*\* Member, Berger & Montague, P.C., a firm that specializes in complex litigation and is a leader in the field of class actions, including toxic tort litigation.

We thank Jim Cook for helping us gather statistics on *Daubert* decisions and Debora Fliegelman for assistance with editing.

1. 509 U.S. 579 (1993). In *Daubert*, the Court displaced the so-called *Frye* test of general acceptability as the guideline for the admission of expert scientific witnesses. In its stead, the Court held that the trial judge "must ensure that any and all scientific testimony or evidence admitted is not only relevant, but reliable." *Id.* at 589. Under its new standard, the Court held that, "[t]he subject of an expert's testimony must be 'scientific . . . knowledge.'" *Id.* at 589-90. In the Court's view, the adjective "scientific" implies "a grounding in the methods and procedures of science." *Id.* at 590. As for "knowledge," the Court stated that it "connotes more than subjective belief or unsupported speculation." *Id.* However, the Court did not conclude, "that the subject of scientific testimony must be 'known' to a certainty." *Id.* Instead, the Court held that, to qualify as "scientific knowledge," an inference or assertion "must be derived by the scientific method." *Id.* Furthermore, "[p]roposed testimony must be supported by appropriate validation—i.e., 'good grounds,' based on what is known." *Id.* In short, the Court established a standard of evidentiary reliability that requires an expert's testimony to pertain to "scientific knowledge." *Id.* at 589-90.

2. 522 U.S. 106 (1997). In *Joiner*, the Court held that abuse of discretion was the proper standard for review of a Federal District Court's decision to admit or exclude expert scientific testimony. The Court also loosened its view stated in *Daubert* that "focus must be solely on principles and methodology, not on the conclusions they generate." 522 U.S. 136, 146 (1997) citing *Daubert*, 509 U.S. at 595.

3. 526 U.S. 137 (1999) (holding that its *Daubert* rulings applied not just to scientific expert testimony, but to all expert testimony).

consistent whole. The interpretation of science as seamless logic puts an undue burden on scientific experts who are challenged under *Daubert*. In fact, *Daubert*, as interpreted by “logician” judges, can amount to a super-Frye<sup>4</sup> test requiring universal acceptance of the reasoning in an expert’s testimony. It also can, in effect, raise the burden of proof in science-dominated cases from the acceptable “more likely than not” standard to the nearly impossible burden of “beyond a reasonable doubt.”

Instead of relying solely on the second view, courts should combine the two. A synthesis of the two views of science can be achieved by recognizing that subjective assumptions and inferences can never be completely eliminated from expert testimony. As a result, expert testimony always amounts, in effect, to conditional statements. An expert’s statements can be considered “reasonable”—or likely, or beyond a reasonable doubt—if, and only if, the assumptions and inferences made by the expert are considered reasonable, or likely, or beyond a reasonable doubt. This “Bayesian” approach<sup>5</sup> may offer a pragmatic synthesis of the tension in the *Daubert* decisions for science-dominated cases, provided that courts can find ways to judge fairly the reasonableness of underlying scientific assumptions. This article offers a number of proposals to meet this goal, along with other suggestions for reform of the *Daubert* review process.

The gatekeeping function assigned to judges by the *Daubert* decision and strengthened by the subsequent Supreme Court cases *Joiner* and *Kumho Tire*, has been both vilified and praised in the legal literature.<sup>6</sup> This article supports the critics of *Daubert*, who see science as a contentious process, rather than a catalog of truths, and who argue that courts are now demanding more of individual scientists and engineers than is expected of them in their own research and practice. Moreover this article follows cues about science from world-

---

4. *Frye v. United States*, 293 F. 1013 (1923). In *Frye*, the Supreme Court articulated the “general acceptance” test that has been the dominant standard for determining the admissibility of novel scientific evidence at trial. The test states that a scientific technique is inadmissible unless the technique is “generally accepted” as reliable in the relevant scientific community. *See id.* at 1014.

5. *See* KENNETH J. ROTHMAN & SANDER GREENLAND, *MODERN EPIDEMIOLOGY* 20 (2d ed. 1998). For a discussion of the historical origin of Bayesian methods and their growing importance in science, see David Malakoff, *Bayes Offers a ‘New’ Way to Make Sense of Numbers*, 286 *SCIENCE* 1460, 1460-64 (1999).

6. For a critique of this assigned gatekeeping function, see Margaret G. Farrell, *Daubert v. Merrell Dow Pharmaceuticals, Inc.: Epistemology and Legal Process*, 15 *CARDOZO L. REV.* 2183 (1994); Michael H. Gottesman, *From Barefoot to Daubert to Joiner: Triple Play or Double Error?* 40 *ARIZ L. REV.* 753 (1998); Frank M. McClellan, *Bendectin Revisited: Is There a Right to a Jury Trial in an Age of Judicial Gatekeeping?* 37 *WASHBURN L.J.* 261 (1998); Eileen A. Scallen, *Classical Rhetoric, Practical Reasoning, and the Law of Evidence*, 44 *AM. U. L. REV.* 1717 (1995); Eileen A. Scallen & William E. Wiethoff, *The Ethos of Expert Witnesses: Confusing the Admissibility, Sufficiency and Credibility of Expert Testimony*, 49 *HASTINGS L.J.* 1143 (1998); *see also* Bert Black et al., *Science and the Law in the Wake of Daubert: A New Search for Scientific Knowledge*, 72 *TEX. L. REV.* 715 (1994) [hereinafter Black et al., *Science and the Law*]; Bert Black, *The Law of Expert Testimony—A Post-Daubert Analysis*, in *EXPERT EVIDENCE: A PRACTITIONER’S GUIDE TO LAW, SCIENCE AND THE FJC MANUAL* (Bert Black & Patrick W. Lee eds., 1997); Bert Black, *The Supreme Court’s View of Science: Has Daubert Exorcised the Certainty Demon?*, 15 *CARDOZO L. REV.* 2129 (1994); Dick Thornburgh, *Junk Science—The Lawyer’s Ethical Responsibilities*, 25 *FORDHAM URB. L.J.* 449 (1998).

renowned scientists, including Neils Bohr<sup>7</sup> and Norbert Wiener,<sup>8</sup> and influential philosophers of science, such as Thomas Kuhn,<sup>9</sup> all of whom have written about the limitations of individual scientists and the subjective elements inherent in the scientific process. The competing school, who might be considered the intellectual descendants of John von Neumann, the inventor of the basic methodology used in the digital computer, would likely disagree.<sup>10</sup> Von Neumann saw science as a set of universal facts deduced by logic, thereby placing scientific knowledge within a formal logical structure that he viewed as primary and universal.<sup>11</sup>

This article proposes that the court interpretation of *Daubert* as requiring “seamless logic” fails to recognize that science is an imperfect process frequently built on assumptions and inferences that simply cannot be proved by the formal logic approach *Daubert* seems to require. Part I outlines the two competing schools of the philosophy of science and examines the roles these two schools have played in litigation. Part II distinguishes *Daubert*, which tends to follow the “formal logic” approach, from the “process” view of science, showing both the flaws in *Daubert’s* approach and the ways in which a shift to the “process” school would render *Daubert* more acceptable. In addition, Part II examines how court attempts to keep “junk science” out of the courtroom have resulted in creating a sort of judge-made “junk science.” Finally, Part III sets forth proposals for reconciling the two views of science and for creating a workable test for determining the admissibility of expert testimony.

---

7. Neils Bohr was never afraid to admit that scientists make mistakes. He is reported to have said, “if one is a physicist, one has necessarily gone through the experience of making a statement and being proved wrong. To a social scientist or a philosopher, this might never happen.” Rudolf Peierls, *Some Recollections of Bohr*, in NEILS BOHR, A CENTENARY VOLUME 229 (A.P. French & P.J. Kennedy eds., 1985). Bohr also always was willing to admit his own limitations. For instance, he was known to exclaim after a moment of clarity, “What fools we’ve been!” John A. Wheeler, *Physics in Copenhagen in 1934 and 1935*, in NEILS BOHR, A CENTENARY VOLUME, *supra*, at 223. Bohr also recognized the complexity of science in subtle ways: He defined a “great truth” as a truth “whose opposite is also a great truth.” *Id.* at 223. This view that science is difficult and at the limit of human abilities contrasts with what we shall call in this article the “von Neumann school of science,” which tended to believe in absolute truths, did not believe that good scientists made mistakes, and thought that scientists who disagreed with them were fools. Of course, there are representatives of both views among scientists today.

8. Norbert Wiener, the father of information theory, “placed scientific activity and natural events all within the fundamental and comprehensive notion of ‘process,’ in which science was a process for describing process.” STEVE J. HEIMS, JOHN VON NEUMANN AND NORBERT WIENER: FROM MATHEMATICS TO THE TECHNOLOGIES OF LIFE AND DEATH 402 (1980).

9. “There is no neutral algorithm for theory-choice, no systematic decision procedure which, properly applied, must lead each individual in the group to the same decision. In this sense it is the community of specialists rather than its individual members that makes the effective decision.” THOMAS KUHN, THE STRUCTURE OF SCIENTIFIC REVOLUTIONS 200 (2d ed. 1970).

10. Von Neumann was an exceptional genius who impressed those he met with the brilliance of his logic. It is said, half-jokingly, that von Neumann, unlike Bohr, never did make a mistake. See HEIMS, *supra* note 8, at 26.

11. *See id.* at 402.

## II

## THE TWO COMPETING SCHOOLS OF THE PHILOSOPHY OF SCIENCE

## A. Science as Logical Reasoning

The von Neumann view that scientific knowledge is embedded in a rational, universal, and logical structure<sup>12</sup> comports with the popular concept of a scientist doggedly collecting irrefutable facts, step by step, and placing them in logical order.<sup>13</sup> Lay judges are most likely to approach science with this ideal in mind,<sup>14</sup> requiring scientists to make generalizations from observations or data to general laws of nature. The von Neumann view further assumes that scientific knowledge exists as chunks of data bound together by logical propositions that can be identified objectively.<sup>15</sup> According to the von Neumann school, disputes

12. For a discussion of von Neumann's views, see HEIMS, *supra* note 8.

13. In a National Science Foundation poll, "only 27% of the sample gave passable answers [in defining an experiment or hypothesis]. In sum, Americans are overwhelmingly interested in science, but don't understand it and know even less about how it is done." Boyce Rensberger, *The Nature of Evidence*, 289 SCIENCE 61 (2000).

The fact that the public is not aware of the philosophical dissections of the scientific method is not surprising, because education in our society is dominated by a particular view of rationality: "Our Western image of rationality assumes a single individual who observes environmental and social realities, deduces universal truths, and then manages environmental and social systems to meet better human needs." Richard Norgaard, *Environmental Science as a Social Process*, 20 ENVTL. MONITORING & ASSESSMENT 103 (1992).

This simplistic view is enhanced by the failure of the popular press, in general, to portray accurately the subject of scientific methodology. Were scientists to talk about their methods with journalists, as well as their findings, there would be a better chance that the public would learn the basic process of science. See Rensberger, *supra*, at 61.

14. "Most judges lack the scientific training that might facilitate the evaluation of scientific claims or the evaluation of expert witnesses who make such claims." Stephen Breyer, *Introduction to REFERENCE MANUAL ON SCIENTIFIC EVIDENCE* 1, 4 (Federal Judicial Ctr. ed., 2d ed. 2000). As a result, most judges are likely to approach science from the popular notion of scientific "rationality," which posits analysts deducing universal truths from data. For a discussion of the popular view of science, see discussion *supra* note 13. The first codification of this "inductive" view of science was made by a lawyer, Sir Francis Bacon.

In Bacon's view, the scientist should be a disinterested observer of nature, collecting observations with a mind cleansed of harmful preconceptions that might cause error to creep into the scientific record. Once enough such observations have been gathered, patterns will emerge from them, giving rise to truths about nature. Bacon's theory has been remarkably influential down through the ages.

David Goodstein, *How Science Works*, in REFERENCE MANUAL ON SCIENTIFIC EVIDENCE, *supra*, at 67, 69.

15. For von Neumann lucid, rapid reasoning based on a thorough familiarity with the facts was by far his strongest suit for dealing with anything. . . . The very language used by von Neumann in formulating quantum theory, namely his description of the process as identifying the 'abstract essence' of empirical theories, has the flavor of philosophical idealism, of the Platonic ideal.

HEIMS, *supra* note 8, at 130-31. Von Neumann

seemed to regard the empirical world, probably even life and mind, as comprehensible in terms of abstract formal structure. . . . He seems to fall under that tradition of Western thought in which it is believed that only rigorous logic will ever succeed in containing the timeless, universal truths that govern everything.

over what is “real” scientific knowledge, as opposed to merely purported scientific knowledge, can be resolved through logical analysis of the data quality and the generalizations made from the data.<sup>16</sup>

## B. Science as Process

Ironically, the “process” school began when historians and philosophers identified *logical* problems with the “logic” school’s view of science as a rational journey that starts with observation, proceeds to logical reasoning, and ends with additions to the “encyclopedia” of scientific knowledge.<sup>17</sup> Logical reasoning simply does not work well without complete confidence in the initial premises. “In classic logic, a deductive argument can provide no information about the truth or falsity of a scientific hypothesis unless you can be 100% certain about the truth of the premises of the argument.”<sup>18</sup> This 100% certainty in the premises can never be realized in science because it is impossible to ensure the validity of all raw, observational data and their interpretation.

The philosopher Karl Popper argued that the use of classic logic by scientists is only one intermediate step in a much larger “process,” which includes intuition, conjecture, inference, professional judgment, and repeated testing:

According to Popper, science advances by a process of elimination that he called “conjecture and refutation.” Scientists form hypotheses based on intuition, conjecture, and previous experience. Good scientists use deductive logic to infer predictions from the hypotheses and then compare observation with the predictions. Hypotheses whose predictions agree with the observations are confirmed only in the sense that they can continue to be used as explanations of natural phenomena. At any time, however, they may be refuted by further observations and replaced by other hypothe-

---

*Id.* at 129-30. Von Neumann held that “logical systems have universality and consequently comprehensiveness: formal logic structure in some way captures the essence of things. This view gave von Neumann an affinity with the seventeenth-century philosophers, especially Leibnitz.” *Id.* at 143.

16. According to Heims, von Neumann shared Leibnitz’s vision of creating a logical calculus that could make even philosophical thinking mechanical. Leibnitz said, as quoted in Heims,

If controversies were to arise, there would be no more need of disputation between two philosophers than between two accountants. For it would suffice to take their pencils in their hands, to sit down to their slates, and to say to each other (with a friend as witness, if they liked): Let us calculate. . . . Von Neumann also surely had great sympathy for this rationalistic aspiration.

*Id.* at 137. This confidence in the power of logic . . . to resolve the problems of human life is merely another form of the optimism and faith in technology that von Neumann inherited from his early years.”

*Id.* at 126. Von Neumann had a “commitment to timeless truths.” *Id.* at 130. Von Neumann dealt with “the philosophical problems underlying mathematics, science, life and mind . . . using the tools of formal logic and abstract mathematics.” *Id.* at 136.

The guiding characteristic that informs so much of von Neumann’s work is the effort to devise as far as possible—and even further—a formal or mathematical structure within which to contain the mysteries of life. It is naïve and optimistic faith in mathematical machinery.

*Id.* at 136.

17. The most notable examples are criticisms by Popper and Kuhn, which are discussed *infra* text accompanying notes 19-25.

18. ROTHMAN & GREENLAND, *supra* note 5, at 20.

ses that better explain the observations. This view of scientific inference is sometimes call *refutationism* or *falsificationism*.<sup>19</sup>

Popper's view has had a "profound effect on the theories of scientific method."<sup>20</sup> Scientific inference under this view is a mixture of intuition, which is subjective, and logical reasoning. Deductive logic can be applied only to the hypothesis when the hypothesis is *assumed* to be 100% true.

Three types of inferences are important in this discussion: (1) *Logical inferences*: inferences made about the world that are derived logically from an assumed hypothesis, theory, or model, whose validity is assumed; (2) *Creative inferences*: inferences that produce new hypotheses to explain previously unexplained data; and (3) *Data inferences*: inferences that are used to fill in missing data or bridge logical gaps in theories. All three types of inferences involve subjective elements. What about the "refutation" and "falsification" parts of science? Are they free of subjectivity? Thomas Kuhn, argued they were not.<sup>21</sup> Instead, he believed that "the consensus of the scientific community determines what is considered accepted and what is considered refuted."<sup>22</sup> Because of the power of this consensus, it is not possible by logic alone to resolve differences either in the laboratory or in the courtroom on what constitutes scientific knowledge. The logicians, of course, disagreed with Kuhn that "mob psychology" determines scientific knowledge: "[T]hose who believe in a rational structure for science consider Kuhn's vision to be a regrettably real description of what passes for scientific activity, but not prescriptive for any good science."<sup>23</sup> Still, Kuhn's view that logic alone cannot determine scientific knowledge enjoys wide popularity among modern scientists.<sup>24</sup>

### C. The Competing Schools in the Courtroom

Representatives of both the "logical" and "process" camps can be found in the courtroom today, usually testifying on opposite sides of a case. The Supreme Court wrestled with the two approaches in *Daubert*, but could not suc-

19. *Id.* at 18.

20. Goodstein, *supra* note 14, at 70.

21. *See* KUHN, *supra* note 9.

22. ROTHMAN & GREENLAND, *supra* note 5 at 20.

But falsification is itself socially constructed—that is, whether a methodology can falsify the conclusion will be determined by the standards, equipment, measurement, and error rates agreed upon by those within the relevant scientific community. Thus, falsifiability is as socially dependent as peer review and 'generally accepted' factors on prevailing scientific paradigms and the norms of the specialized scientific community that they support.

Farrell, *supra* note 6, at 2205.

23. ROTHMAN & GREENLAND, *supra* note 5, at 20.

24. "Kuhn's contribution is important. It gives a new and useful structure (a paradigm one might say) for organizing the entire history of science." Goodstein, *supra* note 14, at 73. For further evidence of the popularity of Kuhn's view in the scientific community, see the quotation from the American Association of Science in the discussion of *Kumho Tire*, 526 U.S. 137 (1999) discussed *infra* text accompanying note 80.

cessfully resolve the inconsistencies between them.<sup>25</sup> As a result of the confusion, lay judges exercising their gate-keeping responsibilities are likely to fall back on the common cultural myth of science as logic, that is, free from subjectivity.

The fact that neither school of scientific thought is universally accepted supports the thesis that science is a contentious process.<sup>26</sup> What makes “science” “scientific,” however, is not consensus, but rather a preference for trying to resolve controversies through experiment and testing, rather than through rhetorical argument alone.<sup>27</sup>

### III

#### ANALYSIS OF DAUBERT FROM THE PROCESS PERSPECTIVE

##### A. Narrow-Minded Interpretations of *Daubert* Are Based on a Fundamental Misunderstanding of Science as Seamless Objective Logic

The tension between the two schools of scientific thought played out in *Daubert*. Ultimately, the Supreme Court set forth two tests that relate intimately to science. First, trial courts must now determine whether the expert is testifying to scientific knowledge,<sup>28</sup> then courts must assess whether the reasoning or methodology underlying the testimony is scientifically valid.<sup>29</sup> Many courts and commentators have interpreted these tests to require testimony that has no flaws in it at all, no leaps in logic, and no unprovable inferences or assumptions.<sup>30</sup> Only then could testimony be based on scientific knowledge and reasoning. This attitude is apparent whenever someone cites a judicial critique

25. See *infra* text accompanying notes 29-42.

26. The thesis that scientists are a contentious lot is contrary to the view expressed by Black. See Black et al., *Science and the Law*, *supra* note 6, at 715.

27. Although we agree with Scallen and Wiethoff that expert scientific testimony is both rhetoric and opinion, the appeal to experiment differentiates scientific rhetoric and opinion from other forms. See Scallen & Wiethoff, *supra* note 6, at 1144.

28. “[T]he trial judge must determine at the outset . . . whether the expert is proposing to testify to . . . scientific knowledge.” *Daubert v. Merrell Dow Pharms., Inc.*, 509 U.S. 579, 592 (1993). “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.” *Id.* at 593 (emphasis added).

29. “This [determination] entails a preliminary assessment of whether the reasoning or methodology underlying the testimony is scientifically valid.” *Id.* at 592-93. “The inquiry envisioned by Rule 702 is, we emphasize, a flexible one. Its overarching subject is the scientific validity . . . of the principles that underlie a proposed submission.” *Id.* at 594. Note that earlier the court used the term “validation,” which has a technical meaning (*partial* test of a theory), that may differ from the common usage of the word (confirmation or proof of validity). As a result, the connection between “validation” and “scientifically valid” as used by the Court in *Daubert* is not entirely clear to a scientific reader. “Proposed testimony must be supported by appropriate validation—i.e., ‘good grounds,’ based on what is known.” *Id.* at 590.

30. See, e.g., *National Football League Prop. v. Prostyle*, 57 F. Supp. 2d 665, 672 (E.D. Wisc. 1999) (requiring that even general statements be tested); *Valente v. Sofamor*, 48 F. Supp. 2d 862, 869-70 (E.D. Wisc. 1999) (“scientific method” requires testing of alternative hypotheses); *Saia v. Sears Roebuck & Co.*, 47 F. Supp. 2d 141, 148 (D. Mass. 1999) (court finds that criticism of methodology indicates unreliability).

of an expert as sufficient evidence that junk science is rampant in the courtroom. However, virtually any complex scientific argument can be criticized, and subjectivity can be found everywhere in science. This blind acceptance of scientific criticism proves that attorneys and judges are not free from their own cultural biases about the nature of science. Most judges, like most laymen, start from the premise that von Neumann's outdated view of science is correct: Scientific knowledge equals universal truths developed by formal logic.<sup>31</sup> They read *Daubert* to have institutionalized this view.

For example, consider the language of a group of frequent *Daubert* commentators, led by Bert Black<sup>32</sup> in an enthusiastic 1994 commentary on *Daubert*: "Scientific knowledge is also unique in the way it emerges by way of consensus. . . . Modern scientists would not challenge the existence of atoms or of a universe with myriad stars, and they would not deny that DNA determines heredity."<sup>33</sup>

The comment is revealing in several ways. First, like von Neumann, it takes the view that science is a collection of universal truths. Second, it actually appeals to scientists as the arbiters of scientific knowledge, as does *Frye*.<sup>34</sup> But here, Black et al. go beyond *Frye* to imply that the goal is consensus: They give examples where apparently all modern scientists agree. The authors, however, do not give examples where there might not be a consensus, where the claim perhaps should be softened to say, "most modern scientists generally accept" the premise

Surely there is no dispute that "consensus" scientific testimony, namely testimony that is accepted by virtually all modern scientists, should be introduced into a courtroom. But would this confine admissible evidence to science that was known in the eighteenth century or before and still holds up today, such as the examples given by Black et al.? Today, pictures of atoms can be seen in scanning electron microscopes.<sup>35</sup> On a clear night myriad stars are visible in the

---

31. For a discussion of judges and popular scientific theory, see *supra* note 14.

32. Bert Black is one of the most frequent commentators on *Daubert*. He was co-counsel on the amicus brief filed in support of Merrell Dow in *Daubert* by the National Academy of Sciences and the American Association for the Advancement of Science. He is one of the "fathers" of the doubling-of-risk standard. (His 1984 paper, written with David Lilienfeld, which proposed the doubling of risk standard, was cited in *DeLuca v. Merrell Dow*, 911 F. 2d 941 (3d. Cir. 1990)). This article agrees on the non-subjective parts of the process of science although it comes from a rather diametrically opposed school of thought.

33. Black et al., *Science and the Law*, *supra* note 6, at 753.

34. In saying that "[m]odern scientists would not challenge the existence of atoms," Black relies on the views of modern scientists as his evidence that his list of items constitutes scientific knowledge. He does not appeal directly to the rationality of the reader, who, like a judge, under the logician's view should be able to judge for herself whether the statements about stars and DNA constitute valid scientific knowledge.

35. Although it is unlikely that any scientist doubts the existence of atoms *per se*, the view of these entities is still changing. As a result, there could be plenty of argument in courts about the very nature, and certainly the behavior, of atoms in modern electronic devices. There is plenty of controversy about quantum "teleportation." See Charles Seife, *Spooky Action Passes a Relativistic Test*, 287 SCIENCE 1909 (2000). There is also controversy over the behavior of atom condensates at low temperatures, which show other strange quantum properties. "In Bose-Einstein condensed systems, instead of each

sky. These are not likely to be contested issues in a courtroom. Black et al.'s third example, namely that DNA is the determiner of heredity, is a principle that might actually be an issue in a modern courtroom. The "consensus" idea that DNA determines all heredity was recently dethroned as a universal principle, albeit after the 1994 article by Black et al. was published, as a result of new studies on infectious proteins, called prions.<sup>36</sup> So much for consensus as the key to scientific knowledge under *Daubert*. A review of the literature often shows that novel, prion-like concepts are floating around as a minority opinion, challenging the notion of universal acceptance.<sup>37</sup> When agreement about what constitutes scientific knowledge can range so widely that even long-held ideas are challenged, it is not easy to come up with a workable alternative to the *Frye* test, which requires the judge to be an arbiter of the views of practicing scientists. Trying to decide which expert is reasoning properly seems a rather difficult task for a court, when even scientists often disagree on how to do it.

Yet the majority of the Supreme Court, like Black et al., are optimistic that the courts can deal with science adequately on their own.<sup>38</sup> In fact, according to Black and his colleagues, it is apparently quite easy: "Though the details of science may be complex, the characteristics of valid scientific knowledge and the

---

atom occupying its own quantum world, they have all entered a single macroscopic quantum state." Keith Burnett, *Go Forth and Multiply*, 282 *SCIENCE* 1657, 1657 (1998).

36. "[P]rions are implicated in several fatal neurodegenerative diseases, including human Creutzfeldt-Jakob disease and 'mad cow disease'." Michael Balter, *Generating New Yeast Prions*, 287 *SCIENCE* 562, 562-63 (2000) (stating that in certain instances, prions can act as genetic elements, changing proteins). Also, prions can reproduce themselves, "violating long-standing dogma that a DNA- or RNA-based genome is necessary for such autonomous behavior." *Id.*

37. Even such a hallowed concept as the gene has been challenged. Evelyn Fox Keller in her book, *The Century of the Gene*, "insists that the gene is neither the stable, self-replicating entity we thought it was, nor a repository of information about development." Jerry A. Coyne, *The Gene Is Dead; Long Live the Gene* 408 *NATURE* 26 (2000). Others have a similar, if more modest, critical view, holding simply that the definition of a gene is unclear. "Is it a heritable unit corresponding to an observable phenotype? Or is it a packet of genetic information that encodes a protein, or proteins? Or perhaps one that encodes RNA? Must it be translated? Are genes genes if they are not expressed?" Teresa K. Atwood, *The Babel of Bioinformatics*, 290 *SCIENCE* 471 (2000).

Even more heretical, some respectable scientists challenge, at least in part, the genetic laws of Gregor Mendel that underlie Darwinian natural selection. The arguments of these revisionist scientists have been

creeping back into the scientific literature. . . . The reason: an explosion in the field of epigenetics, the study of changes in genetic expression that are not linked to alterations in DNA sequences. Some of these epigenetic changes can be passed on to offspring in ways that appear to violate Mendelian genetics . . . [and] could play a role in evolution.

Michael Balter, *Was Lamarck Just a Little Bit Right?*, 288 *SCIENCE* 38 (2000).

38. "We are confident that federal judges possess the capacity to undertake this review." *Daubert v. Merrell Dow Pharms., Inc.*, 509 U.S. 579, 593 (1993).

kind of reasoning that produce it are not difficult to grasp.”<sup>39</sup> If this is so easy, one wonders why the recognition of invalid science cannot be left up to juries.<sup>40</sup>

Although Black et al.’s article gives a lot of useful generalizations about the process of science, the authors never come up with operational methods for recognizing the universal scientific knowledge and universally valid science that they and other supporters of *Daubert* seek, other than the special case when there is actually universal agreement. The authors concede that they have only guideposts—“[t]hese points do not by themselves constitute an adequate basis for resolving scientific issues”—but never offer the missing steps that would provide an adequate basis.<sup>41</sup> The Supreme Court also fails to provide those steps. The expectation is that lay judges will be able to spot false steps in scientific reasoning, and therein lies a most fundamental problem. One can almost always find a false step or two within any scientific theory; few theories are perfect, even the most useful ones.

Just as jury members may leap to the conclusion that the proximity of exposure and disease most likely means causation, so too may judges leap to the conclusion that a gap in a scientific chain of logic makes an expert’s testimony faulty science. Both are incorrect views. *Daubert*, as interpreted by many courts, is not a satisfactory solution to the jury “problem.” Instead, the Supreme Court has taken us from the frying pan of the supposedly causation-challenged jury into the fire of judicial misconceptions about science.

## B. Defining the “Scientific Method”

The *Daubert* Court’s insistence on defining “scientific method” as a logical pursuit exhibits the common misconception that scientists “prove” something. Proof is the province of mathematicians, not scientists, and even mathematicians start from unprovable assumptions.<sup>42</sup>

---

39. Black et al., *Science and the Law*, *supra* note 6, at 753. In contrast, Bauer states that “a major long-standing effort of the philosophy of science was to establish unambiguous criteria by which to distinguish science—the activity that produces reliable knowledge about the world—from pseudoscience, which misleads and inevitably falls into error. The scientific method does not provide such criteria.” HENRY BAUER, SCIENTIFIC LITERACY AND THE MYTH OF THE SCIENTIFIC METHOD 57 (1994).

40. “There is no hard empirical evidence supporting the argument that a lay jury cannot critically evaluate scientific evidence.” Marilee M. Kapsa & Carl B. Meyer, *Scientific Experts: Making Their Testimony More Reliable*, 35 CAL. W. L. REV. 327 (1999). See Luke M. Froeb & Bruce H. Kobayashi, *Naïve, Biased, yet Bayesian: Can Juries Interpret Selectively Produced Evidence?*, 12 J.L. ECON. & ORG. 257 (1996); Robert D. Myers et al., *Complex Scientific Evidence and the Jury*, 83 JUDICATURE 150, 155 (1999). The authors propose Arizona as a model for providing special resources to help juries in this regard.

41. Black et al., *Science and the Law*, *supra* note 6, at 782. Black et al. never come to terms with those who hold contrary views about resolving scientific issues, such as those expressed by Kuhn and others: “There is no neutral algorithm for theory-choice, no systematic decision procedure which, properly applied, must lead each individual in the group to the same decision.” KUHN, *supra* note 9, at 200. In fact, “science is, above all, an adversary process. It is an arena in which ideas do battle, with observations and data the tools of combat.” Goodstein, *supra* note 14, at 74.

42. Imagine Euclid testifying in a modern day *Daubert* proceeding: “Professor Euclid, I understand that one of your postulates is that parallel lines do not meet at infinity. Can you prove this to be true? Have you ever tested this? Isn’t it also true that Professor Einstein has proven that your geometry doesn’t work in the presence of gravity?”

What then do scientists do, if they don't prove things? The most important action scientists take is to find errors.<sup>43</sup> Philosopher David Miller notes that “[f]rom the point of view of rationality, science is above all its method—essentially the critical method of searching for errors.”<sup>44</sup> Scientists *disprove* things.<sup>45</sup> In the process, they filter error from theories and methodology, but they do not prove that the surviving methodologies—those that are left standing or those that are changed to correct errors—are valid.

What a semantic nightmare for plaintiffs under *Daubert*, who generally have the burden of proof. How can plaintiffs succeed under *Daubert* with a discipline that eschews proof, relying only on its negative? The Supreme Court has never directly addressed this fundamental contradiction between the absence of proof in the “scientific method” and its requirement in the legal system. The Court has never faced up explicitly to the fact that the testing of theories and methodologies is almost always partial. Those theories that survive testing still have components that have never been tested,<sup>46</sup> contain subjective elements, and require that reasonable inferences be made if they are to be used in real world examples. Scientific theories, including those that are extremely useful, are imperfect.<sup>47</sup> It follows that scientific knowledge itself must be imperfect. As a result, the underlying theories and “knowledge” on which experts rely in testimony almost always have imperfections. The *Daubert* Court glossed over these philosophical problems, leaving lower courts to deal with an internally inconsistent decision that can be interpreted to support the process view of science in some parts and the formal logic view in others.

1. “*Scientific Knowledge*” and the *Scientific Network*. “Disproof” through testing is powerful, playing a role analogous to that of natural selection in removing species from the evolutionary timeline. The method of disproof works without requiring scientists to agree with each other on theories or to accept every supposed discovery (although judges and juries might wish

---

Even the field of mathematics is not free from inherent contradictions, however, as discovered by the mathematician Godel in the first half of the twentieth century. “The result is remarkable in its quasi-paradoxical ‘self-denial’: It will never be possible to acquire with mathematical means the certainty that mathematics does not contain contradictions.” HEIMS, *supra* note 8, at 1436 (quoting John von Neumann). It is worth noting that Godel’s work was in part a response to the efforts of Bertrand Russell and other logicians to reconstruct mathematics from fundamental principles of logic. *See id.*

43. Scientists obviously do much more than just find errors. They do exploratory research. They collect and analyze data. They make up hypotheses in attempt to explain previously murky phenomena. However, in the context of scientific proof and disproof, the identification of errors is the most important scientific activity.

44. David Miller, *Being an Absolute Skeptic*, 284 SCIENCE 1625 (1999).

45. *But see* Goodstein, *supra* note 14, at 71 (arguing that it is not necessarily theories that get disproved, since it may be the assumptions connecting the theory with the test that are wrong). However, proving an assumption wrong in the application of a theory also advances science.

46. When a theory survives a test, the test is called a “validation.” A validation is partial. It does not make the whole theory “valid.” There is a possible confusion in the law over the scientific use of the word “validate.” To a lay person, “validate” suggests proving that the theory is totally correct; to a scientist, “validate” implies the process of undertaking a series of (necessarily) incomplete tests.

47. “[N]ot all science is based entirely on empirically verified statements,” David W. Hogg, Letters, *A Barely Tested Hypothesis*, 286 SCIENCE 1679 (2000).

otherwise). Controversies do not need to be resolved by submission to orthodoxy. Experiments serve as the judge and jury, deciding on particular components of a theory—at least until the next set of experiments is undertaken.

The scientific enterprise is a complex network of hundreds of thousands of interacting scientists, each communicating with overlapping subsets of colleagues, all of whom are tied to a body of constantly growing scientific literature. Different scientists can have vastly different opinions, even contradictory ones, and still be legitimate parts of the network. In fact, it can be argued that controversy among scientists is the driving force in the system.<sup>48</sup> Regardless, the system as a whole is greater, more productive, and more robust than its parts. It is constantly weeding out error—not all error at any one time, but significant amounts of it—from the inferences, theories, and methodologies that are generated through subjective, creative energy. Still, it should never be forgotten that plenty of error is left in the system and therefore in the minds of testifying experts. It would be convenient for the courts if scientists agreed where the remaining errors are hidden, but they do not.

Even within specific fields, the scientific network is so complex and constantly changing that it defies comprehension and definition by an individual scientist, let alone a lay judge. What we loosely call “scientific knowledge” is embedded in that network. Scientific knowledge is an abstraction that cannot be caged with language. Agreement on what constitutes scientific knowledge ranges from the near-universal to the hotly contested.<sup>49</sup> Statements about scientific knowledge can be honed by committees and eventually worded with sufficient caveats to gain fairly wide-spread agreement at any one time. However, use of scientific committees is not an avenue open to the courts in individual cases. Furthermore, widespread agreement does not satisfy the goal of fitting scientific knowledge into formal and universal logic. “Widespread opinion” is simply a fraction of the whole; it does not represent unanimity.

Yet, in the *Daubert* decision, the Supreme Court demands that an expert testify to “scientific knowledge,”<sup>50</sup> when neither it nor anyone else can possibly

---

48. “Science is, above all, an adversary process. It is an arena in which ideas do battle, with observations and data the tools of combat.” Goodstein, *supra* note 14, at 74.

49. *See id.*; *see also* discussion *supra* note 38. In Kuhn’s historical view, the extent of disagreement about what constitutes scientific knowledge is affected by the stage to which a particular paradigm has evolved.

If the paradigm is one destined to win its fight, the number and the strength of the persuasive arguments in its favor will increase. More scientists will then be covered, and the exploration of the new paradigm will go on. Gradually the number of experiments, instruments, articles, and books based upon the paradigm will multiply. Still more men, convinced of the new view’s fruitfulness, will adopt the new mode of practicing normal science, until at last only a few elderly hold-outs remain. And even they, we cannot say, are wrong.

KUHN, *supra* note 9, at 159.

50. *See Daubert v. Merrell Dow Pharms., Inc.*, 509 U.S. 579 (1993). “[T]he trial judge must determine at the outset . . . whether the expert is proposing to testify to . . . scientific knowledge.” *Id.* at 592.

recognize “scientific knowledge” due to its abstract and amorphous nature. The *ipse dixit* of the Supreme Court that “federal judges possess the capacity” to undertake a proper review is not an answer.<sup>51</sup> No scientific instrument or meter can be held over expert testimony to identify scientific knowledge; no algorithm can tell us which testimony is valid and which is not. By failing to resolve the semantic problems that engulf the concept of scientific knowledge, the Supreme Court has encouraged lower courts to undertake hopeless searches for scientific Holy Grails. The resulting confusion will surely lead to judicial error, primarily because the courts have no institutional mechanisms for peer review by *scientists* to minimize the amount of scientific error that will creep into legal decisions and precedents.<sup>52</sup> The courts will end up enforcing a kind of quasi-scientific orthodoxy that, in the absence of reform, could well institutionalize unfairness and bring disrespect to the judiciary.

2. *Scientific Knowledge and Disagreement Among Scientists.* Even the most revered scientists have their detractors. For example, physicist Paul Dirac, remains a hero to many physicists. His exposition of quantum mechanics is perceived as a wonderful example of elegance in science.<sup>53</sup> Almost every physicist would qualify Dirac’s book on quantum mechanics as an example of scientific knowledge, certainly suitable for use in a court of law, if relevant.<sup>54</sup> Yet, here is what the eminent mathematician, Jean-Alexandre-Eugene Dieudonne, had to say about Dirac’s book:

When one arrives at the mathematical theories on which quantum mechanics is based, one realizes that the attitude of certain physicists in the handling of these theories truly borders on delirium. . . . One has to ask oneself what remains in the mind of a student who has absorbed this unbelievable accumulation of nonsense, real hogwash! It would appear that today’s physicists are only at ease in the vague, the obscure, and the contradictory.<sup>55</sup>

---

51. “We are confident that federal judges possess the capacity to undertake this review.” *Daubert*, 509 U.S. at 593. The court is confident despite the fact that “it is difficult to find evidence standards that are applicable to all sciences because each science is at a different stage of sophistication; each field uses different tools, operates on different assumptions, employs different methodologies, and has different goals, resulting in large cultural and language gaps.” Kapsa & Meyer, *supra* note 41, at 327. Also, “our U.S. common law system is not conducive to comprehensive scientific contemplation.” *Id.* Note that one of the authors, Carl Meyer, is a chemist as well as an attorney.

52. The courts, of course, have *legal* peer-review mechanisms institutionalized through the appeals process. For instance, an appeals court reviewing a decision of a lower court for errors in law is analogous to a group of scientists reviewing the science in a manuscript submitted for publication. A published opinion of an appeals court, if extensive, is analogous to a scientific paper criticizing or supporting previous work. Future courts can take advantage of the wisdom therein expressed.

53. For instance, the physicist, Subrahmanyan Chandrasekhar cites Dirac’s book as one of five examples of “harmoniously organizing a domain of science with order, pattern, and coherence.” N. David Mermin, *What’s Wrong with this Elegance?*, PHYSICS TODAY, Mar. 2000, at 11 (2000).

54. The extraordinary achievements that Dirac made in physics and the respect with which he is held can be summed up by an exchange in a set of letters in *Physics Today*: “P.A.M. Dirac’s goal was to find a description of the electron consistent with both relativity and quantum mechanics . . . . No theorist has since found anything comparable to the Dirac equation. Remarkable and even great theoretical achievements cited in this set of letters are simply not in the same league.” Harry J. Lipkin, *Letter to the Editor*, PHYSICS TODAY, Jan. 2001, at 68. See also Mermin, *supra* note 53, at 16.

55. See Mermin, *supra* note 53, at 16.

Physicists would be shocked if a lay judge pronounced Dirac's treatise on quantum mechanics as invalid and lacking evidentiary reliability. Yet, a lay judge could not be blamed for doubting the validity of Dirac's work, if the judge had an affidavit containing language like this from someone of Dieudonne's stature. Such an affidavit would serve as reasonable grounds under *Daubert* to exclude hypothetical expert testimony based on Dirac's book.<sup>56</sup> A skeptical court might also add the weight of Einstein's name, by quoting the criticisms of quantum mechanics that Einstein spent half his life developing.<sup>57</sup>

Criticism of scientist by scientist is not uncommon. It may not be polite, but harsh criticism is healthy science. When two scientists criticize each other in the courtroom, it does not mean that one must be wrong. They are both part of a process that converges upon, but never reaches, truth. The challenge of resolving scientific conflict in the courtroom is much more complicated than the *Daubert* methodology indicates. It would probably be easier on everyone if scientists were simply viewed as imperfect eyewitnesses to the scientific enterprise and handled as the legal system handles other sets of conflicting eye-witnesses.

"Valid" science is thus a relative and subjective term, especially if no one bothers to inquire about the purpose for which the science is needed.<sup>58</sup> Opposing counsel, if they look hard enough, can always find some expert with sincere convictions to contradict another expert. As stated earlier, science thrives on disagreement and controversy. A legal system that takes no cognizance of these process realities is operating under a substantial misconception. Doubt can always be cast on an expert, even as great a scientist as Paul Dirac, let alone an ordinary university professor who has been drafted in support of a David taking on a Goliath.<sup>59</sup> If judges are supposed to rule an expert's testimony invalid based on the existence of some gaps and inferences in the scientific argument, then the *Daubert* system of review is nothing more than a method of unclogging the legal calendar, with no pretense of connection to the real scientific enterprise.

Scientists should be held to a standard in expert testimony that is no higher than the standard by which higher courts judge the *de facto* scientific opinions espoused by district judges. Although district judges have been given discretion under *Joiner* for legal purposes, the *Joiner* opinion includes discretion for judges even when they reason as amateur scientists.<sup>60</sup> Thus, so long as a district

---

56. See *Daubert*, 509 U.S. at 593 (on the grounds that it was vague and contradictory).

57. See Seife, *supra* note 35, at 1909. Despite Einstein's brilliance in so many subfields of physics, all experiments to date have supported quantum mechanics, not Einstein's criticisms. See Anton Zellingner, *The Quantum Centennial*, 408 NATURE 639, 639-41 (2000).

58. For instance, a methodology of isotope analysis might be valid for identifying the country from which a sample of coal was dug but totally invalid in determining from which mine in the country the sample came from.

59. "Absolute certainty is never likely to be present. Thus, it is always possible to cast some doubt on any scientific testimony, even that which is based on well-established principles." Lawrence S. Pinsky, *The Use of Scientific Peer Review and Colloquia to Assist Judges in the Admissibility Gatekeeping Mandated by Daubert*, 34 HOUS. L. REV. 529 (1997).

60. *General Electric Co. v. Joiner*, 522 U.S. 136, 143 (1997).

court's scientific conclusions about an expert are "reasonable," according to the Supreme Court, the judge's scientific opinion cannot be challenged.<sup>61</sup> The same degree of discretion should be afforded to experts, who, after all, actually have some training and experience in the field at issue, in contrast to judges, who generally do not. At a minimum, an expert's testimony should be ruled invalid only upon a finding of abuse of scientific discretion, namely reasoning that is held to be scientifically unreasonable.<sup>62</sup> We applaud proposals to reform expert review along these lines,<sup>63</sup> which we applaud so long as the phrase "scientifically reasonable" leaves room for the subjective assumptions and inferences that scientists generally make.

In addition to making experts fully state the bases for their opinions,<sup>64</sup> judges should concentrate on explaining the limitations of experts to juries. They should not try to protect jurors from being exposed to imperfect testimony, a hopeless endeavor unless all scientists who go beyond stating that atoms exist and that there are myriad stars are banned from the courtroom.

3. *Limitations Imposed by the Methodology of Disproof.* As we have argued, the underlying methodology of science is disproof—not precise deductive reasoning from postulates like mathematics. Therefore, scientific theories, and hence expert testimony, must contain subjective elements.<sup>65</sup> The number of subjective elements in a theory will gradually be reduced over time, but never to zero. Thus, when an expert testifies in the courtroom, some degree of subjectivity is inevitable.

---

61. "In our view, the doubts that triggered the District Court's initial inquiry here were *reasonable*, as was the court's ultimate conclusion." *Kumho Tire Co. v. Carmichael*, 526 U.S. 137, 153 (1999) (emphasis added). "Thus, whether *Daubert's* specific factors are, or are not, *reasonable* measures of reliability in a particular case is a matter that the law grants the trial judge broad latitude to determine." *Id.* (emphasis added).

62. As argued previously in this article, there is no precise definition of "scientifically valid," as applied in the scientific community, and no precise scientific definition of other similar terms used by the Supreme Court. Similarly, there can be no precise scientific definition of "scientifically unreasonable." Thus, an expert's testimony should be considered scientifically reasonable in the legal context, when an expert demonstrates reasoning based on informed scientific judgment. Failure to meet such a standard renders an expert's opinion scientifically unreasonable. Our goal here is to endorse a definition that makes it clear to all parties that "proof" of an expert's opinion is not required before it can legitimately reach a jury. Branding an expert's report reasonable does not require a judgment that it is either correct or "proven."

63. For example, the National Council on Uniform State Laws has submitted to the Federal Judicial Council a proposed draft of a revised FED. R. EVID. 702 that provides for testimony based upon principles or methodology that is "reasonably reliable." See Kapsa & Meyer, *supra* note 40, at 330.

64. "But scientists may always be asked to explain their choices, to exhibit the bases for their judgments. Such judgments are eminently discussable, and the man who refuses to discuss his own cannot expect to be taken seriously." THOMAS KUHN, *THE ESSENTIAL TENSION, SELECTED STUDIES IN SCIENTIFIC TRADITION AND CHANGE* 337 (1977).

65. "Even in the most rigorous and quantitative fields, such as theoretical physics, scientific reasoning is not purely mathematical, but depends upon simplification by creative use of approximations." *Id.* at 324.

In fact, science itself is a set of beliefs largely passed on from professor to student and colleague to colleague.<sup>66</sup> These beliefs, for the most part, must be taken on faith by an individual expert, or else she would never have time to undertake new research.<sup>67</sup> Although there is overlap in beliefs that are constantly changing due to the results of experiments and theoretical analyses, scientific literature is a cacophony of voices. It is a hopeless endeavor to try to find a universal belief set that amounts to anything more than a list of platitudes—platitudes that are of little use in a particular case. And they would still be unproven, as we learn every time a cherished theory is overturned. For instance, until recently, it was accepted that the brain never produced new neurons in adults. That belief is now demolished.<sup>68</sup>

Any legal opinion that acknowledges the process view of science and uses the terms “scientific belief” or “scientific knowledge” implicitly incorporates the idea of subjective, professional judgement, as the following hypothetical case study illustrates.

*a. The hypothetical scientific paper at issue in a Daubert challenge.* This hypothetical manuscript starts out, as do many, with a “methods” section that attempts to lay out the techniques used, as well as the assumptions, approximations, interpolations, and extrapolations that were required during the course of the study. For example, a number of subjective data decisions may have been required to correct data, combine it into meaningful categories, eliminate outliers, choose statistical methods, and to determine how to handle missing data. Approximate values for quantities are often used when a researcher infers that a more precise treatment would have a negligible impact on the analysis. As support for their chosen methods, the authors may cite opinions from other papers. Not every scientific reader of the paper will agree with the decisions that were made. Some of the reviewers may have doubts about some of the choices made and would have decided otherwise. However, legitimate doubts can be raised about virtually all exercises of judgment in any research paper.

---

66. “Scientific theories are *beliefs* upon which explanations are based, and changes in theory are changes in beliefs, and such changes demonstrate that earlier beliefs were inadequate, or unsatisfactory, or, not to mince words, *wrong*.” BAUER, *supra* note 39, at 64.

67. As an example of learning on faith, see Bauer discussing the textbook science that “we learn through typically dogmatic teaching in school and undergraduate college and that we tend to believe in unreservedly.” *Id.*

The scale of modern science is so large that no individual scientist can know the basis and limitations of everything taught in class, or presented in a textbook or review article. To do so, a scientist would have to be familiar with the entire literature in a discipline, which is enormous. For instance, consider the biosciences alone. PubMed, which is the National Library of Medicine’s search service, provides access to over 11 million citations in the medical literature. To get an idea of the size of the literature base for sub-fields and specialties, we ran a number of searches for various key words and phrases. “Cancer” turned up 1.3 million articles and books. “Science, medicine, and measurement,” turned up 800,000, 500,000, and 300,000 citations, respectively. A specialty like “breast cancer” turned up 55,000 citations. “Toxicology” turned up 43,000 citations to the literature. “Virus” turned up 300,000. See PubMed (visited Feb. 8, 2001) <<http://www.ncbi.nlm.nih.gov/entrez>>.

68. See Fred H. Gage, *Mammalian Neural Stem Cells*, 287 SCIENCE 1433, 1433-38 (2000).

Let us assume that this hypothetical paper was part of a major research project, of many year's duration, with sufficient funding to allow the appointment of an advisory committee stocked with major figures in the field. Such an advisory committee represents the gold standard for outside advice, since not every study is sufficiently well-funded. In this hypothetical case, the study director and team meet with the advisory committee every six months to review various decisions that have been proposed by the project staff, often after long internal debate. The outside reviewers are likely to respond in different ways to each proposed methodological choice. If the advisors generally assent to a particular step as being reasonable, the meeting moves forward. If there is considerable opinion against a particular proposal, then major changes are made, consistent with budgetary realities. Sometimes, the study director has no choice but to go with an unpopular decision, because of time and budget restraints. In contrast to the opinion of one district judge, time and money are not irrelevant to the scientific method.<sup>69</sup> Finally, there will be some cases in this hypothetical example in which the outside advisors are all totally divided in their views, and the study director again proceeds with the internal staff plan.

Many papers will contain a "results" section, where the authors try to present the fruits of their labor. Presentation decisions must be made as to which data to present, what to emphasize, and how to present it. For example, the total amount of data collected in a study is usually vastly more copious than can be presented. Data must be combined into categories, whose definitions are subjective. There is no generally accepted method to decide where to draw the boundary between categories.<sup>70</sup>

Finally, many papers have a "discussion" section, which often contains quite speculative comments on the meaning, significance, or general applicability of the data and why it may differ from earlier results. Again, reference will be made to other papers, including opinions that agree or disagree with the authors.

In a complex research project, some disagreement is likely during all phases of the study before, during, and after publication. Sometimes even co-authors would have preferred alternate approaches to parts of the study.<sup>71</sup> Not surprisingly, some variation of these critical views about a particular study will likely show up in court, should the hypothetical study ever play a major role in expert testimony in a major case. This will even be true for the finest studies possible in the health and epidemiologic fields.<sup>72</sup>

---

69. See *In re Hanford Nuclear Reservation Litigation*, 1998 U.S. Dist. LEXIS 15028, at \*69 (E.D. Wash. Aug. 21, 1998) ("Lack of time and resources is not a valid excuse . . . It is irrelevant to the scientific method.")

70. See ROTHMAN & GREENLAND, *supra* note 5, at 206.

71. The fact that even scientific co-authors can differ, as we have learned through the process of writing scientific papers, demonstrates the correctness of the Supreme Court's statement that there is a range "where experts might reasonably differ." *Kumho Tire Co. v. Carmichael*, 526 U.S. 137, 153 (1999).

72. Secondary research papers are even more subjective, which is why there is concern that authors of review papers be free from conflicts of interest. Consider a scientist preparing a review paper who

*b. Applying the research paper to a particular case.* What happens to a scientist called as an expert who tries to apply the hypothetical study to a particular case? Rarely will a study apply to the exact facts of a case. In a toxic torts case, for example, the population in the study may be different from plaintiffs, and there will almost always be questions raised in a *Daubert* challenge about the applicability of the study results to the population seeking redress. For example, despite the overwhelming increase in thyroid disease that has followed the Chernobyl accident in the former Soviet Union, many analysts have held the results questionable to western nations, because the exposed population contained subgroups of people who were iodine deficient, a condition largely absent in the West.<sup>73</sup>

The duration of exposure may differ for study populations and plaintiffs, as may the age distribution of the two populations. Experts can disagree as to the significance of these differences, and the matter cannot be proven one way or the other without further research—research that probably cannot take place until after future accidents.

Faced with a study that shows a health effect, defendants challenging an expert under *Daubert* will try to split the data as finely as possible to distinguish the situation from the actual plaintiffs, while the plaintiffs will attempt the opposite. For a study that shows no health effect, the parties will likely reverse their approach.

Much of the process of applying a research paper to the facts of a case can be labeled “extrapolation,”<sup>74</sup> a tried and true process in science. Although fraught with potential dangers, extrapolation is essential. For example, we extrapolate that the sun will rise next month based on past experience with celes-

---

tries to synthesize data from a host of authors. The choice of studies reviewed and emphasized is subjective; the choice of language used to describe them is subjective; the weight given to different studies is subjective. Despite their subjectivity, review papers are in great demand by scientists. Over 64,000 reviews or review-like articles are published each year. Over 200 serial journals are devoted to reviews. See ISI Index to Scientific Reviews (visited Jan. 30, 2001) <<http://www.isinet.com/isi/products/indexproduct/scientificreviews>>. Yet, review papers are particularly vulnerable to a *Daubert* challenge because of their obvious subjectivity. Ironically, the best reviews of the state of knowledge in a field in our opinion are often prepared by science writers for magazines such as *Science* and *Nature*. The ability of a science writer to interview representatives of different schools of thought, and to include memorable quotes, often gives their articles a clarity and succinctness that more scholarly writers cannot seem to match. Articles by science writers are unlikely to be admitted as evidence in a courtroom, but perhaps courts and counsel could benefit sometimes from the talents of a science writer, more than they could from the opinion of a scientific expert, who will necessarily have one set of intellectual biases.

73. See David V. Becker et al., *Childhood Thyroid Cancer Following the Chernobyl Accident*, 25 ENDOCRINOLOGY & METABOLISM CLINICS N. AM. 206 (1996); E.P. Kasatkina et al., *Effects of Low Level Radiation from the Chernobyl Accident in a Population with Iodine Deficiency*, 156 EUR. J. PEDIATRICS 916-20 (1997); S. Davis et al., Fred Hutchinson Cancer Research Center, Draft Final Report of the Hanford Thyroid Disease Study II-8 (1999) (visited Feb. 8, 2001) <<http://www.fhcr.org/science/phs/htds>>.

74. Extrapolation is a process of inferring values for data points that lie outside the region containing the measured values. Often, a smooth functional form will fit to the existing data and the functional form will then be used to infer the actual imputed value. See Jan Beyea & M. Hatch, *Geographic Exposure Modeling: A Valuable Extension of GIS for Use in Environmental Epidemiology*, 107 ENV. HEALTH PERSPECTIVES 181 (1999).

tial mechanics, even though it cannot be proven with absolute certainty that the sun will not blow up as a supernova in the meantime. Almost every application of science requires extrapolation or interpolation.<sup>75</sup> Whether that leap of faith is reasonable or reliable is a subjective judgment. Some extrapolations engender more confidence than others, but professional judgment cannot be separated from the process. True, in quantitative exercises, one can try to assess the increased uncertainty thereby introduced, but that, too, is subject to professional judgment.

Had the Supreme Court clearly established a standard of reasonableness for expert testimony, rather than leaving room for interpretations by the lower courts that require an absolute standard of scientific validity, it would have grounded *Daubert* more firmly within the actual practices of science.

### C. Despite Firm Footing in the “Logic” School, the Court Recognizes the “Process” Aspect of Science

The *Daubert* Court acknowledged that in principle science cannot be free from all doubt, but the Court never indicated in the original decision how much doubt and uncertainty would be allowed.<sup>76</sup> Without some guidance to lower courts, the statement that some doubt would be allowed amounts to mere lip service. Only later, in *Kumho Tire*,<sup>77</sup> do we get a hint of how a lower court might safely decide on what amounts to an acceptable uncertainty: Presumably, if testimony falls within the “range where experts might reasonably differ,”<sup>78</sup> the uncertainty thereby introduced would be acceptable.

The Court also acknowledged the process view concerning testing. For instance, quoting a statement from the amicus brief of the American Association of Science, the Court stated, “[s]cience is not an encyclopedic body of knowledge about the universe. Instead, it represents a process for proposing and refining theoretical explanations about the world that are subject to further testing and refinement.”<sup>79</sup> The majority went on, however, to add a very significant caveat, the meaning of which is ambiguous, but which could be interpreted as putting the court in the camp of the “logicians”: “But, in order to qualify as ‘scientific knowledge,’ an inference or assertion must be derived by the scientific method.”<sup>80</sup>

---

75. Interpolation refers to the process of inferring data values at locations that lie between points where measurements have actually been made. Often, a smooth functional form will be fit to the existing data and the functional form used to infer the actual imputed value. *See id.* at 181-90.

76. “Of course, it would be unreasonable to conclude that the subject of a scientific testimony must be ‘known’ to a certainty; arguably, there are no certainties in science.” *Daubert v. Merrell Dow Pharms., Inc.*, 509 U.S. 579, 589 (1993).

77. “[T]he District Court determined that Carlson’s testimony was not reliable. It fell outside the range where experts might reasonably differ, and where the jury must decide among the conflicting views of different experts, even though the evidence is ‘shaky.’” *Kumho Tire Co. v. Carmichael*, 526 U.S. 137, 153 (1999).

78. *Id.*

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

80. *Id.*

The word “derived” comes from the language of mathematics and formal logic, not science.<sup>81</sup> If the court meant that an inference must be derived logically from the hypothesis, theory, or model on which the expert is relying, then the Court’s view is consistent with the process view of science because the hypothesis may still contain subjective elements. On the assumption that the hypothesis, theory, or model is 100% correct, it is possible to derive a number of consequences that should meet a test of formal logic. If the expert could not meet such a test, he would be inconsistent with his assumptions. Few philosophers, historians of science, or working scientists would disagree that such inconsistency, if it were significant, should brand testimony as invalid science.

On the other hand, other inferences do not fit so easily within a framework of derivation. Data inferences, such as extrapolations, for example, cannot be derived. They are “best judgments,” which are inherently subjective. How then can they be derived by the scientific method? If the Court meant to exclude testimony that depends on data inferences, it has ruled out virtually all of modern science.

In *Joiner*, the Court indicated that it did not intend to rule out all data inferences when it recognized that “trained experts commonly extrapolate from existing data.”<sup>82</sup> The Supreme Court went on to say, however, that a lower court may conclude that there is simply too great an analytical gap between the data and the opinion.<sup>83</sup> This puts an incredible amount of scientific discretion in the hands of a lay district court, which is one of the reasons that *Daubert* judicial rulings can be so subjective. At a minimum, such discretion poses a risk of unequal justice. As Justice Joseph T. Walsh of the Delaware Supreme Court put it, “the presence of a significant subjective factor in the gatekeeper calculus poses a real risk of differing results depending on the idiosyncrasies or predisposition of the trial judge.”<sup>84</sup>

Another problem with the Supreme Court’s caveat that inferences must be derived by the scientific method is that the definition of the scientific method

---

81. A search of the National Library of Medicine’s PubMed database turned up but one article in 11 million citations (which include abstracts) that discussed anything in the abstract (or title) that was “derived” by the “scientific method.” Specifically, we queried the National Library of Medicine’s 11 million record database (PubMed) for the words, “Derived,” and “Scientific,” and “Method.” Of the 89 hits, only one paper used the three words in the abstract or other search field in one sentence. Here is the unique sentence: “The standards presented in this paper are derived from the tenets of the scientific method, measurement theory, statistical principles, and research ethics.” N.M. Ryan-Wenger, *Guidelines for Critique of a Research Report*, 61 HEART LUNG 394, 394 (1992). See PubMed (visited Feb. 8, 2001) <<http://www.ncbi.nlm.nih.gov/entrez>>.

A similar search of PubSCIENCE, the physical science counterpart to PubMed, turned up no such articles. PubSCIENCE is a World Wide Web service developed by the Department of Energy’s (DOE) Office of Scientific and Technical Information (OSTI) to facilitate searching and accessing peer reviewed journal literature in the physical sciences and other energy-related disciplines. See PubSCIENCE (visited Feb. 8, 2001) <<http://pubsci.osti.gov>>.

82. *General Electric v. Joiner*, 522 U.S. 136, 146 (1997). The distinction between “trained” and “untrained” experts is not clear.

83. See *id.*

84. Joseph T. Walsh, *Keeping the Gate: The Evolving Role of the Judiciary in Admitting Scientific Evidence*, 83 JUDICATURE 140, 142 (1999).

almost always includes testing.<sup>85</sup> Did the Court expect that all inferences made by a scientific expert must be tested for validity, rather than merely the underlying hypothesis, theory or model?<sup>86</sup> If so, this demands too much of a testifying expert.<sup>87</sup> It is impractical to test every inference that can be drawn from a theory, especially those that are drawn about a particular legal case. Immediately after the sentence on the scientific method, the Court brings in language about validation: “[P]roposed testimony must be supported by appropriate validation—i.e., ‘good grounds,’ based on what is known.”<sup>88</sup> Perhaps, this was meant as an explication of the Court’s phrase “derivation by the scientific method.” Unfortunately, the meaning of “validation” is up for grabs here. If the Court were using “validation” as scientists use it, as a partial test of a theory, this sentence would be consistent with what scientists actually do. It would be consistent with the fact that theories are never fully and completely tested. However, the word “validation” is likely to be interpreted by most readers based on its common English meaning of confirmation or proof of validity, which is philosophically impossible.

---

85. There are exceptions. For instance, some scientists have argued that pure exploratory research is more significant than testing hypotheses and theories—at least for a time. See Harry J. Lipkin, *Who Ordered Theorists*, PHYSICS TODAY, July 2000, at 15. In a more philosophical vein, Paul Feyerband has argued that it is impossible to constrain the scientific method: “The only rule is that anything goes.” See A.F. CHALMERS, *WHAT IS THIS THING CALLED SCIENCE?* 135 (1982). Although these two views surely have merit for intervals of time and for particular individuals, they do not provide guidance to the most common situations in science. To handle the more typical research scenarios, most commentators on the philosophy of science—at least those with whom we are familiar—acknowledge that the testing of models and hypotheses play a significant role in science. For instance, Chalmers, who critically reviews all of the major figures in the philosophy of science discussed by us in this article, and many others we have not had space to discuss, still retains the idea of testing in his synthesis:

The results must be able to stand up to further testing procedures conducted, first, perhaps, by the experimenter’s colleagues and later, if the social structure of science happens to be similar to that of our own, by the referees of journals. If the results survive such tests and become published, their adequacy will be liable to be tested on a wider front.

CHALMERS, *supra*, at 119. Henry Bauer, who considers the scientific method to be a myth—only an ideal that is never reached—also sees testing as a major activity of scientists, although he maintains that competing scientists test each others theories, not their own. See BAUER, *supra* note 39, at 52-55.

86. Perhaps, the Court did not really intend to require an expert to have tested every inference. Perhaps, the court was using the phrase scientific method in a special way. For example, the next sentence speaks of “validation:” “[P]roposed testimony must be supported by appropriate validation—i.e. ‘good grounds,’ based on what is known.” *Daubert v. Merrell Dow Pharms., Inc.*, 509 U.S. 579, 590 (1993). Perhaps, this language is a guide to the Court’s definition of the scientific method.

Unfortunately, the meaning of “validation” is up for grabs here. If the Court were using “validation” as scientists use it, as a partial test of a theory, this sentence would be consistent with what scientists actually do. It would be consistent with the idea of providing ‘good grounds.’ However, without further clarification by the Court, the word “validation” is likely to be interpreted by most readers based on its common English meaning of confirmation or proof of validity, which is philosophically impossible.

87. David L. Faigman proposes to make testing the heart of a revised Rule 702 incorporating the “lessons” of *Daubert*. Although supposedly focused on testing the underlying methodology, it appears that his proposal would likely doom most toxic tort actions. There will always be aspects of a particular case that will require extrapolation and historical reconstruction, which for practical reasons cannot be tested. See David L. Faigman, *Making the Law Safe for Science: A Proposed Rule for the Admission of Expert Testimony*, 35 WASHBURN L.J. 401 (1996).

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

The final problem with the “derivation by the scientific method” requirement is that lay judges may substitute their own understanding of the scientific method language, which unknowingly may be rooted in a logician’s view. A legal logician would read the Court’s caveat to imply that any and all inferences must be derived seamlessly by classic logic, that is, be free from logical attack and significant doubt. By its choice of language, the Court has given comfort to an old-fashioned view of the philosophy of science, which tries to place scientific knowledge within a formal logical structure. Although many practicing scientists today still adhere to this view—usually those who do not pay much attention to the philosophy of science—it is an outdated view.<sup>89</sup> It does not help the legal system confront the basic contradiction between legal proof and scientific disproof.

D. Although Meant to Eliminate “Junk Science” from the Courtroom,  
*Daubert* Has Instead Created a New Form of Junk Science—Judicial Junk

Because the testing methodology is not available for resolving controversies between scientists disagreeing in the courtroom,<sup>90</sup> it is a logical impossibility for judges (or juries) to think that they can resolve disagreements about what constitutes either “scientific knowledge” or “valid” science—ambiguous terms at best—in an absolute and scientific way.<sup>91</sup> Of course, ambiguity in language is unavoidable. In the *Daubert* context, however, lay judges have been asked to define and operationalize terms, with which they have negligible experience. Although district judges represent a select group of the population and some have undergraduate degrees in science, the vast majority of judges are poorly trained in science.<sup>92</sup> Asking them, without expert assistance, to operationalize concepts like “scientific knowledge” and “validity” is a risky prescription for error and embarrassment.

The Ninth Circuit’s decision in *Daubert II*, expressed understated amazement at what the Supreme Court was asking it to do:

---

89. Such a view is outdated because it ignores Popper, Kuhn, and the Bayesians, discussed earlier. Scientists who spend their careers in the laboratory do not need to know the intricacies of the philosophy of science. Some, like von Neumann, may be skeptical of philosophical limits: “By contrast, von Neumann, skeptical of limitations imposed by *a priori* philosophical considerations, sought to deal with the most general and abstract in terms of formal logic.” HEIMS, *supra* note 8, at 140. Von Neumann “avoided philosophizing, but loved formal mathematical and logical structure which transcends time and place.” *Id.* at 145.

90. One exception is the “unique court-ordered simulation conducted . . . [to help] . . . determine whether federal agents fired gunshots into the Branch Davidian compound shortly before it burned to the ground on April 19, 1993.” Jim Yardley, *Warring Conclusions on Waco Simulation*, N.Y. TIMES, Mar. 21, 2000, at A14.

91. The scientific methodology of testing does not always resolve controversies quickly because individual scientists have to make their own judgments about how studies should be interpreted. The immediate impact of scientific experiments can usually be seen in the evolution that takes place in conflicting theories to account for the test results. Remaining disagreements lead to a new round of testing. Controversies die when scientists get bored with arguing and move on, or can no longer get funding to pursue a line of research, or in Kuhn’s view, when partisans of the old view finally die off. See KUHN, *supra* note 9, at 151.

92. See Breyer, *supra* note 14, at 5.

Our responsibility, then, unless we badly misread the Supreme Court's opinion, is to resolve disputes among respected, well-credentialed scientists about matters squarely within their expertise, in areas where there is no scientific consensus as to what is and what is not "good science," and occasionally to reject such expert testimony because it was not "derived by the scientific method." Mindful of our position in the hierarchy of the federal judiciary, we take a deep breath and proceed with this heady task.<sup>93</sup>

And why is all this necessary? Some argue that it is to keep "junk science" out of the courtroom, thereby preventing future breast implant, asbestos, and agent orange litigation.<sup>94</sup> One ironic consequence of the attempt to have amateur scientists keep "junk science"<sup>95</sup> out of litigation is that some courts have ended up enshrining their own "junk science" into the body of the law.

An example of invalid judicial science is the standard for causation-in-fact that the Ninth Circuit tried to set for personal injuries in toxic tort cases under California law. After determining that proof of causation under California law requires a "more likely than not" standard,<sup>96</sup> the court considered how epidemiologic data might be used to prove causation-in-fact in a particular case. Faced with epidemiologic data alone, it equated a calculation of probability

93. *Daubert v. Merrell Dow Pharms., Inc.*, 43 F.3d 1311, 1316 (9th Cir. 1995) (hereinafter *Daubert II*). Whether the courts will only "occasionally" reject challenged expert testimony, as the Ninth Circuit guesses, or whether they will do so more frequently, is open to dispute. See *infra* text accompanying notes 131-136 (discussing the data on expert challenges).

94. Kapsa and Meyer offer the following alternate explanation for the change in expert standards:

A dominant force for change of expert standards has nothing to do with science, but is motivated by a shift in economic concerns. Thus, scientific and professional trade organizations tend to encourage testimony that is favorable to the industries on whose funding they depend, but usually discourage scientific testimony on behalf of consumers or environmental groups. Also, during the past two decades, the public policy concerning product liability and environmental issues has shifted due to economic and political pressures. Defendants in product liability cases have claimed that unbridled testimony by scientific experts and junk science have misguided juries to come out with run-away verdicts that damage our economy.

Kapsa & Meyer, *supra* note 40, at 317.

Farrell has another view: "Seen in that light, *Daubert* is as much a case of class struggle between common jurists and elite judges as a struggle between the scientific and legal communities." Farrell, *supra* note 6, at 2207 n.113.

95. This pejorative term, which is rarely defined, is used in this article to hoist its users on their own petard. There are two definitions for "junk science." The first is an interesting one by former Attorney General Richard Thornburgh, who seems to define junk science as plaintiffs' science: "Lawyers casting about for new theories to use to sue manufacturers of drugs, medical devices and other products create a limitless demand for junk science." Thornburgh, *supra* note 6, at 450 (quoting Joseph M. Price & Gretchen Gates Kelly, *Junk Science in the Courtroom: Causes Effects and Controls*, 19 HAMLIN L. REV. 395, 396 (1996)).

More even-handedly, Eileen Scallen and William Weithoff define junk science as "expert testimony based on natural or social science theories that are not accepted as reliable." Scallen & Weithoff, *supra* note 6, at 1145 n.6. The authors do not specify who decides on the reliability. If it is scientists who are supposed to be judging reliability, then there is a problem with defining the group of scientists who are to make the judgment. If it is the judge, then Scallen & Weithoff have defined "junk science" as science that a lay judge thinks is unreliable. Such a definition is rather unsatisfactory. Nevertheless, for the purposes of this article "junk science" is defined by a similar circular definition, namely, "junk science equals invalid science." It would have been preferable to avoid the term completely.

96. "California tort law requires plaintiffs to show not merely that Bendectin increased the likelihood of injury, but that it more likely than not caused *their* injuries. *Daubert II*, 43 F.3d at 1320 (citing *Jones v. Ortho Pharm. Corp.*, 163 Cal. App. 3d 396, 403 (Cal. Ct. App. 1985)).

(doubling-of-risk) with the legal requirement of “more likely than not.”<sup>97</sup> This so-called doubling-of-risk standard has been adopted in toxic tort cases by a number of other courts as well,<sup>98</sup> although not always with the Ninth Circuit’s caveat that evidence beyond epidemiologic data alone could reduce the plaintiffs’ burden of proof.<sup>99</sup> Caveat aside, the doubling-of-risk methodology for assigning individual causation is based on false science, as explained below.

1. *The Doubling-of-risk Methodology.* The doubling-of-risk methodology is applied to data taken from epidemiologic studies in which the suspected causal factor increases the relative risk for the study population as a whole by some factor.<sup>100</sup> For example, consider the side effects of women taking a certain medicine. Exposure to the medicine by a sample of pregnant women might increase the risk of birth defects by 30% (a factor of 1.3) compared to women who never took the medicine. Or, it might increase the relative risk by 100% (factor of 2), or by 200% (factor of 3). (See Table 1.)

TABLE 1:  
THREE HYPOTHETICAL EXAMPLES OF EPIDEMIOLOGICAL DATA

|                                   | Sample I<br>(no exposure)<br>Number of birth<br>defects per million | Sample II<br>(with exposure)<br>Number of birth<br>defects per million | Risk ratio<br>(or relative risk) |
|-----------------------------------|---|--|----------------------------------|
| Example I<br>30% increase in risk | 1000  | 1300   | 1.3 = 1300/1000                  |
| Example II<br>Doubling of risk    | 1000  | 2000   | 2.0 = 2000/1000                  |
| Example III<br>Tripling of risk   | 1000  | 3000   | 3.0 = 3000/1000                  |

Only in Example III does the relative risk more than double. Example II sets the boundary between less than a doubling of the risk and more than a

97. *Daubert II*, 43 F.3d at 1316.

98. See *In re Joint E. & S. Dists. Asbestos Litig.*, 758 F. Supp. 199, 203 (S.D.N.Y. 1991); *Manko v. United States*, 636 F. Supp. 1419, 1434 (W.D. Mo. 1986)); LINDA A. BAILEY ET AL., FEDERAL JUDICIAL CENTER, REFERENCE GUIDE ON EPIDEMIOLOGY 169 & n.128 (1994) (citing *Deluca v. Merrill Dow Pharms.*, 911 F.2d 941, 958-59 (3rd Cir. 1990)). However, at least one court has correctly characterized the doubling-of-risk methodology as an *inference* about individual causation, rather than as a statistical proof one way or the other. See *In re Joint E. & S. Dists. Asbestos Litig.*, 758 F. Supp. at 203 (citing *In re Fibreboard Corp.*, 893 F.2d 706, 711-12 (5th Cir. 1990)).

99. “A statistical study showing a relative risk of less than two could be combined with other evidence to show it is more likely than not that the accused cause is responsible for a particular plaintiff’s injury.” *Daubert II*, 43 F.3d at 1321 n.16.

100. Relative risk is a ratio. It equals the risk of getting a particular illness *with exposure* divided by the risk of illness *without the exposure*.

doubling of the risk. The doubling-of-risk standard for generic causation requires the relative risk to have exceeded two in a study population before the study can be said to provide evidence that the toxic substance causes the injury in question.<sup>101</sup> For individual causation, the doubling-of-risk standard requires the plaintiffs to have been exposed to enough of the toxic substance to place them in a population whose relative risk also more than doubled.

The mathematical idea behind the doubling-of-risk standard is that when the ratio is two, half of the birth defects are supposed to be caused by the medicine and half by other causes (background causes). This suggests a ratio. Assuming the validity of the statistical analysis, it is tempting to equate a doubling of the risk with the legal standard “more likely than not.” Then, the only factual situation in Table 1 that would sustain a verdict for the plaintiff would be Example III.

Consider a plaintiff with a birth defect whose mother had taken the medicine. Can anything be inferred—scientifically and mathematically, using the epidemiologic data in Table 1—about the odds that the birth defect of a particular plaintiff had been caused by the medicine? The Ninth Circuit asserted that this inference was possible as a matter of statistics, referring the reader only to a prior Third Circuit decision, *DeLuca v. Merrill Dow Pharmaceuticals*<sup>102</sup> and an unpublished report by the Washington Legal Foundation,<sup>103</sup> a conservative think tank whose mission is “to defend and promote the principles of free enterprise and individual rights.”<sup>104</sup> The *DeLuca* decision cites a 1984 law review article<sup>105</sup> that was written before the critiques of the doubling-of-risk methodol-

---

101. “[T]he relative risk of limb reduction defects arising from the epidemiological data . . . will, at a minimum, have to exceed ‘2.’” *Daubert II*, 43 F.3d at 1321 (quoting *DeLuca*, 911 F.2d at 958). See also Michael D. Green et al., *Reference Guide on Epidemiology in REFERENCE MANUAL ON SCIENTIFIC EVIDENCE*, *supra* note 14, at 333, 383-84.

102. 911 F.2d 941 (3d Cir. 1990).

103. The Court refers to an unpublished report by the Washington Legal Foundation to provide “a more complete explanation of the relationship between the burden of proof and relative risk.” *Daubert II*, 43 F.3d at 1321 (citing ROBERT P. CHARROW & DAVID E. BERNSTEIN, *SCIENTIFIC EVIDENCE IN THE COURTROOM: ADMISSIBILITY AND STATISTICAL SIGNIFICANCE AFTER DAUBERT* 28-33 (Washington Legal Found. ed., 1994)).

104. *WLF's Mission and Goals* (visited Feb. 16, 2001) <<http://www.wlf.org/mission.htm>>. Although scientists routinely cite data from unpublished reports, they do so only when they have the expertise necessary to make an independent assessment of the quality of the material. The Ninth Circuit did not possess the expertise to assess the statistical arguments in any report and might have been more cautious in its language about the doubling-of-risk statistics, given that the report was written by an attorney and a law professor, not by mathematicians. On the other hand, the court may not have been given counter-documents from opposing counsel, which may explain its belief that the statistics were not disputed.

105. See *Deluca*, 911 F.2d at 945-47, 956, 959. One of the authors of the cited law review article, David Lilienfeld, is an eminent epidemiologist. However, the authors simply assert that the doubling-of-risk methodology is true without statistical proof: “If, in an exposed population, more than half the cases of a disease can be attributed to the exposure . . . then absent other information about a diseased individual, it is more likely than not that his or her illness was caused by the exposure.” Bert Black & David Lilienfeld, *Epidemiologic Proof In Toxic Tort Litigation*, 52 *FORDHAM L. REV.* 732, 767 (1984).

The authors here use the word “attributed” in the epidemiologic sense of “attributable risk.” For critiques of using attributable risk to assign individual probabilities, see Jan Beyea & Sander Greenland, *The Importance of Specifying the Underlying Biologic Model in Estimating the Probability of Cau-*

ogy appeared in the biostatistical literature. In any case, the *DeLuca* court failed to note the caveats<sup>106</sup> stated by the authors.<sup>107</sup>

It is not surprising that the Ninth Circuit felt that its statistical theorem was so obvious that it needed only a citation to another court and to an unpublished report. More technically sophisticated analysts have leapt to the same conclusion.<sup>108</sup> However, in the future, when a court makes a pronouncement that “as a matter of statistics” a certain statement is valid, it may want to be sure it is thoroughly briefed by statisticians from the relevant field. Had the Ninth Circuit been so briefed, it would have learned that its math was only valid in a very limited case, as will be demonstrated below.<sup>109</sup>

2. *The Doubling-of-risk Method in the Courtroom.* Not only did the Ninth Circuit believe that it was possible to compute individual risks from group epidemiologic data, it also maintained that it could quantify the relationship:

In terms of statistical proof, this means that plaintiffs must establish not just that their mothers' ingestion of Bendectin increased somewhat the likelihood of birth defects, but that it more than doubled it—only then can it be said that Bendectin is more likely than not the source of their injury. Because the background rate of limb reduction defects is one per thousand births, plaintiffs must show that among children of mothers who took Bendectin the incidence of such defects was more than two per thousand.<sup>110</sup>

---

sation, 76 HEALTH PHYSICS 269 (1999); Sander Greenland & James M. Robins, *Conceptual Problems in the Definition and Interpretation of Attributable Fractions*, 128 AMER. J. EPIDEMIOLOGY 1185 (1988) [hereinafter Greenland & Robins, *Conceptual Problems*]; Sander Greenland & James M. Robins *Epidemiology, Justice, and the Probability of Causation*, 40 JURIMETRICS 321 (2000) (discussing biologic arguments) [hereinafter Greenland & Robins, *Epidemiology*]; James M. Robins & Sander Greenland, *Estimability and Estimation of Excess and Etiologic Fractions*, 8 STAT. MED. 845 (1989) [hereinafter Robins & Greenland, *Estimability*].

106. The authors of the article cited by *Deluca* restricted their analysis to substances that produce “additive” risks: “These attributable risks are either additive or multiplicative (in lay terms ‘synergistic’). If the former, the analysis of this Article can be applied with little further elaboration; if the latter, other rules of attribution are required.” Black & Lilienfeld, *supra* note 105, at 734 n.6. The *Deluca* court appears to have made no such distinction. Finally, the authors of the cited article defined scientific causation in a limited way, not accounting for substances that may simply speed up the onset of disease. *See id.* at 762.

107. For a history of the scientific criticisms of the doubling-of-risk methodology, see Beyea & Greenland, *supra* note 105. The problems with the doubling-of-risk methodology were demonstrated many years ago by two biostatisticians, Sander Greenland and James M. Robins. *See* Robins & Greenland, *Estimability*, *supra* note 105. Unfortunately, their highly mathematical work was published in statistical journals and never penetrated the legal consciousness. Recently, Greenland and Beyea have reviewed the situation in less mathematical language, along with simple examples, that one hopes are more comprehensible to non-scientists. *See* Beyea & Greenland, *supra* note 105. In addition, a new paper by Greenland and Robins has now been published that is written for the non-technical reader. *See* Greenland & Robins, *Epidemiology*, *supra* note 105.

108. *See* Beyea & Greenland, *supra* note 105; Greenland & Robins, *Epidemiology*, *supra* note 105, at 329 (“[I]n sworn statements, and declarations by epidemiologists and statisticians enrolled as expert witnesses, we found only one case where the expert was aware that biologic assumptions were required.”).

109. *See infra* text accompanying notes 119-122.

110. *Daubert II*, 43 F.3d 1311, 1320 (9th Cir. 1995).

And so a major precedent was set in the Ninth Circuit on the heels of the *DeLuca* decision, which no doubt helped to encourage other jurisdictions to follow the doubling-of-risk trend.<sup>111</sup>

This decision was somewhat ironic: Had an expert filed a report that included the court's math, it could never have survived a serious *Daubert* challenge. Although true for the identical red and white balls that are used as specialized examples in basic statistics courses, the theorem is false when applied to toxic substances that may interact in complex fashion in the body and may cause different impacts on individuals depending on their susceptibilities.<sup>112</sup>

The doubling-of-risk methodology is valid in only a few very specialized cases. In the example previously discussed of a hypothetical medicine that allegedly causes birth defects, the methodology would be correct only if (1) all the women and fetuses react the same biologically to the medicine, and (2) the toxic agent does not affect how the other causes of the disease behave.<sup>113</sup> When courts adopt the doubling-of-risk methodology, they are not usually aware that they are making these kinds of assumptions and thus do not consider whether the assumptions are reasonable in a particular case.

Some courts have cited the *Reference Guide on Epidemiology*, published by the Federal Judicial Center,<sup>114</sup> as a basis for adopting the doubling-of-risk standard. However, the *Reference Guide* does not present the relative risk standard as a fact. What the authors actually state is that “[a] relative risk greater than 2.0 would permit an *inference* that an individual plaintiff's disease was more likely than not caused by the implicated agent.”<sup>115</sup> This inference falls far short of a fact. In fact, the guide points out in a footnote that one school of statisticians and many epidemiologists resist this inference.<sup>116</sup>

Nevertheless, as inferences go, it is not an unreasonable one to make in the absence of any other information. Scientists make exactly this kind of data inference all the time in the absence of sufficient information. Other scientists may disagree with their choice, but neither group would be practicing invalid science by making inferences about the proper methodology to use in calculating individual risks, provided they did not claim more than they could prove, that is, provided that they listed appropriate caveats.

Consider a hypothetical damage case involving the responsibility of a gun manufacturer, the Hotgun company, for causing deaths in a certain community, even when the gun is not found. Hotgun, Inc., has developed a new, inexpensive, but elegant, lightweight handgun. Sales are brisk. Within six months of introduction of the new weapon, the murder rate involving handguns increases by

---

111. A substantial number of courts in a variety of toxic substance cases have accepted the doubling-of-risk standard. For a listing, see Green, *supra* note 101, at 140.

112. See Beyea & Greenland, *supra* note 105, at 269.

113. See *id.*

114. BAILEY ET AL., *supra* note 98, at 169.

115. *Id.*

116. See *id.* at 168 n.127.

30%. In other words, the relative risk of being killed by a handgun rose to 1.3 after the introduction of Hotgun.<sup>117</sup> The total murder rate went up because new buyers who had never carried guns before entered the market, and some of them ended up using those guns.

The plaintiff, Mary Smith, is suing Hotgun because her husband was killed in an incident involving a handgun. As is common in such incidents, the handgun was never recovered. Mary Smith brought her case against Hotgun in a state where the legal standard for causation is “more likely than not.” According to the Ninth Circuit’s methodology, the relative risk of dying from handguns has only increased by a factor of 1.3, which is less than the threshold relative risk of 2.0. This failure of the epidemiologic data to show a doubling of the risk supposedly makes the odds less than 50% that a Hotgun was involved in the murder of Ms. Smith’s husband. Summary judgment would be granted to defendant Hotgun in the Ninth Circuit and in any other jurisdiction that has adopted the doubling-of-risk standard.

Probing a little deeper into the science, however, produces a different result in this hypothetical case. Hotgun did more than attract new buyers: It was so attractive to the criminal element that it displaced older guns and captured 90% of the total market. Ninety percent of people carrying handguns were carrying Hotguns at the time of Mr. Smith’s murder, and 90% of the murders using handguns involved Hotguns. The *percentage* of murders involving Hotguns was higher than the 30% one might infer from the murder rate because criminals already carrying guns switched models. As a result, the chance that Mrs. Smith’s husband was killed by a Hotgun was actually 90%. Defendant should not be granted summary judgment on the basis that the plaintiff had to prove a doubling of the risk in the population as a whole.

It can be argued that this reasoning is unfair, because it requires the introduction of new information that is speculative and can never be known in a real case. Actually, this reasoning is no more speculative than the Ninth Circuit’s implicit inference that toxic torts do not interfere with other causes of disease. Furthermore, the Ninth Circuit’s implicit inference is not the most reasonable inference to make. Consider the Hotgun situation: Economists know that introduction of cheaper products have at least two effects on the market. First, total sales increase by a factor related to the elasticity of demand. Second, there is a substitution of products, cross-elasticity, that can often be the largest effect.<sup>118</sup> The expectation that the liability of Hotgun should be judged by total sales is ridiculous from an economist’s point of view.

---

117. This is our epidemiologic data, analogous to Example I in Table 1.

118. “When the price of Toyota Tercels rises, all else being equal, demand for Tercels falls and demand for Nissan Sentras, a substitute, rises.” *THE FORTUNE ENCYCLOPEDIA OF ECONOMICS* 8 (David R. Henderson ed., 1993). More precisely, the cross elasticity of demand refers to “the responsiveness of quantity demanded of one good to a change in the price of another good.” *THE M.I.T. DICTIONARY OF MODERN ECONOMICS* 91 (David Pearce ed., 4th ed. 1992). When the price of Toyota Tercels rises, people don’t stop buying cars; instead, those who are particularly price sensitive are more likely to purchase a different model, assuming that its price has remained constant.

In toxic tort cases, a toxicologist might find equally ridiculous the idea of estimating causation solely from group level increases in disease. For example, the phenomenon of biologic repair, which is quite common, can lead to toxic agents interacting with background causes, thereby destroying one of the assumptions on which the doubling-of-risk standard is based.<sup>119</sup> In general, epidemiologic data cannot determine an individual's probability of causation without the specification or assumption of an underlying model of how the toxic substance actually causes the disease (a biologic model).<sup>120</sup> The Ninth Circuit did not make an unreasonable *inference*, in the absence of further information about Bendectin, but it should not have presented its logic as a matter of mathematics. Testifying scientists could reasonably have drawn different inferences based on their knowledge of biology and toxicology.<sup>121</sup>

This example of courts enshrining invalid science into precedent should serve as a warning to all those who would have judges play the role of amateur scientists. What happens when courts make mistakes due to their ignorance of the correct scientific methodology and/or understanding? Are these errors of fact or errors of law? Can cases be re-opened if the scientific bases on which they were decided are shown to be false. And what happens in general when new discoveries overturn the scientific dicta that courts routinely espouse as a basis for their decisions, such as the inadequacy of animal testing? Can these cases also be re-opened?

3. *The Doubling-of-risk Methodology Has Falsely Made Other Methods of Assigning Causation Look Unscientific.* Unless one is aware of its scientific flaws, the statistical aura surrounding the doubling-of-risk methodology makes other methods of assigning causation in toxic tort cases seem primitive, speculative, and unscientific. Other methods such as "weight of the evidence"<sup>122</sup> and reliance on general toxicological experience<sup>123</sup> clearly involve subjective judgments and assumptions by the expert.<sup>124</sup> However, the doubling-of-risk standard is not based on assumption-free statistics. The doubling-of-risk methodology is just as subjective as other methods of assigning causation. To use it correctly, a scientist has to look at all available information about the disease and the plaintiff and reach, through some unspecifiable and hidden mental process, a decision about what underlying biologic model is best to assume. That process is inherently subjective. Imagine how difficult it would be under a *Daubert* challenge to justify the choice of an underlying biologic model, when the disease mechanisms are not understood. In the Hotgun

---

119. See Beyea & Greenland, *supra* note 105, at 272.

120. See *id.* at 270.

121. For example, competition between toxic chemicals and dietary substances for receptor sites on cells that permit disease initiation can be compared to gun manufacturers competing for the criminal market.

122. Gottesman, *supra* note 6, at 770.

123. Reliance on general toxicological experience is discussed *infra*.

124. See Phillip Harbor & Dennis Shusterman, *Medical Causation Analysis*, 38 J. OCCUPATIONAL & ENV'T MED. 577 (1996).

example, it is possible to think of hundreds of alternative scenarios that might be taking place. For instance, the introduction of Hotgun might stimulate one desperate competitor to lower its prices drastically, practically giving away its “old-fashioned gun” to preserve market share for as long as possible. This could change the cross-elasticity effect, causing its products to become the main handgun used in killings for months. To come up with a best estimate for the underlying economic model, an economist would have to rely on experience and professional judgment. The economist could not give a seamless, scientific explanation for making the choice that would satisfy a *Daubert* logician. Yet if a court were unaware of the need to specify an underlying economic model, it might well conclude that the doubling-of-risk methodology was more scientific than the testimony of a hypothetical economist, who might simply make a professional judgment about the likely percentage change to be expected in gun types, based on Hotgun’s potential market share.

To bring the analogy closer to toxic tort litigation, consider the approach used by toxicologist Dr. Daniel Teitelbaum, in *General Electric v. Joiner* as an alternative to the doubling-of-risk methodology:

As a toxicologist when I look at a study, I am going to require that the study meet the general criteria for methodology and statistical analysis, but that when all of the data is collected and you ask me as a patient, “Doctor have I got a risk of getting cancer from this?” That those studies don’t answer the question, that I have to put them together in my mind and look at them in relation to everything I know about the substance and everything I know about the exposure and come to a conclusion.<sup>125</sup>

This is a description of a subjective process that will apply at some level to any determination of causation in toxic tort cases, until such time when the molecular biology of disease is so well understood that the causal chain can be traced at the individual’s molecular and cell level all the way from exposure to final disease state, fully explaining the role of other causal agents. Given the pace of change in molecular biology, this cannot be ruled out. However, at the present time, subjectivity rules individual causation in toxic tort cases.<sup>126</sup>

What is so harmful about the doubling-of-risk methodology is not its subjectivity, but the deceptive impression it gives of being embedded in a scientific methodology that has escaped subjectivity. Scientific causation itself has definitional problems, particularly when no theory can explain all stages of the causal chain, as is the case for cancer, birth defects, and virtually all diseases that come up in toxic tort cases.<sup>127</sup> In the absence of such a theory, a doctor or scientist

---

125. *General Electric v. Joiner*, 522 U.S. 136, 153 n.4 (1997) (Stevens, J., dissenting).

126. Subjectivity also rules general causation. The so-called “Hill’s postulates,” used by epidemiologists are guidelines, not prescriptions in determining scientific causation. They help order the hidden mental process, but do not eliminate the subjective element in reaching a judgment on whether a substance causes disease. See Austin Bradford Hill, *The Environment and Disease: Association or Causation?* 58 PROC. ROY. SOC. MED. 295 (1965).

127. “Though it is basic to human thought, causality is a notion shrouded in mystery, controversy, and caution, because scientists and philosophers have had difficulties defining when one event *truly causes* another.” JUDEA PEARL, CAUSALITY: MODELS, REASONING AND INFERENCE 331 (2000). In his book, Pearl discusses how causal models, which consist of functional relationships among variables of interest, can be used to compute probabilities of causation. See *id.* at 202; see also Beyea & Green-

must fill in the gaps with subjective theory elements based on past experience and best professional judgment. As Teitelbaum describes, the analyst must take everything he or she knows about the substance or exposure, put the information into the mind—this “black box”—and see what comes out.<sup>128</sup> Is Teitelbaum’s process subject to abuse? Is the doubling-of-risk methodology subject to abuse? Yes, but so are all subjective decisions in which the true mental process is hidden from view, even to the thinker. Juries must wrestle with the credibility of the expert to decide whether the expert truly reached the decision as claimed, or “cheated” to favor one side. Judges are not in a better position to make such a determination,<sup>129</sup> although judges need to be involved to see that the opposing parties are given a fair chance to critically dissect for the jury the process, or lack of process, followed by experts.

E. If *Daubert* Is Interpreted as Requiring a Seamless Chain of Logic, It Sets an Unreasonable and Unfair Standard of Legal Proof for Science-Dominated Civil Cases

The difficulty scientists have had in defending evolution against the onslaught of creationists is an example of the widespread perception that scientists are supposed to prove things beyond any doubt.<sup>130</sup> Creationists have been successful in removing the teaching of evolution from school district after school district, and even entire states.<sup>131</sup> Creationists succeed because they focus on the fact that evolution, like all of science, is only a theory,<sup>132</sup> and then present scientific evidence against evolution.<sup>133</sup>

---

land, *supra* note 105. Causal models are rare in disease biology: “Biologic knowledge about epidemiologic hypotheses is often scant.” ROTHMAN & GREENLAND, *supra* note 5, at 23. “[L]ittle is known about the causes of most defects.” C.R. Weinberg & Allen J. Wilcox, *Reproductive Epidemiology*, in ROTHMAN & GREENLAND, *supra* note 6, at 606. See also Greenland & Robins, *Epidemiology supra* note 105, at 330 (discussing biologic arguments).

128. See *Joiner*, 522 U.S. at 153 n.4.

129. Some have argued that judges have more time to consider the issues. However; “[t]here is no hard empirical evidence supporting the argument that a lay jury cannot critically evaluate scientific evidence.” Kapsa & Meyer, *supra* note 40, at 327.

130. See Constance Holden, *Creationists Win in Kansas*, 286 SCIENCE 2242 (1999).

131. See Eugenie C. Scott, *Not (Just) in Kansas Anymore*, 288 SCIENCE 813 (2000).

132. Two of the biggest weaknesses of evolutionary theory are: 1. There is no adequate explanation for the *origin of life* from dead chemicals. Even the simplest life form is tremendously complex. 2. The *fossil record*, our only documentation of whether evolution actually occurred in the past, lacks any transitional forms, and all types appear fully-formed when first present. The evidence that ‘pre-men’ (ape-men) existed is dubious at best. So called pre-man fossils turn out to be those of apes, extinct apes, fully man, or historical frauds.

Creation Science, *The Big Issues* (visited Feb. 16, 2001) <<http://emporium.turnpike.net/cs/cartoon.htm>>.

133. As an example of how deeply creationists get into the details of scientific papers, much like attorneys do in *Daubert* challenges, consider this response by scientists who found their work used by creationists to attack evolution:

A recent study coauthored by several of us . . . examined inaccuracies in embryo drawings published last century by Ernst Haeckel. Our work has been used in a nationally televised debate to attack evolutionary theory, and to suggest that evolution cannot explain embryology. We strongly disagree with this viewpoint. Data from embryology are fully consistent with Darwinian evolution. Haeckel’s famous

The existence of gaps in logical proof has been the creationists' most powerful argument, and there are entire websites devoted to detailed critiques of scientific arguments concerning evolutionary theory.<sup>134</sup> Attackers of scientific experts under *Daubert* can use the same tactics as the creationists, creating doubt in the minds of lay judges, hoping that this will be perceived as fatal to a scientific theory. Yet, unlike mathematical theorems, theories are adopted by individual scientists because they are useful over time, not because they are true like mathematical theorems. As noted, no one can even prove that the sun will rise next month, though few doubt that it will do so.

From the logician's perspective, strict adherence to *Daubert* will doom most challenged experts. As a result, in front of a "logician" judge, the effective criterion for judgment in a science-dominated civil suit appears to have been raised from "more likely than not" to "beyond a reasonable doubt," and possibly to "beyond all doubt." Consider a toxic tort case where the contested issue is the amount of a plaintiff's exposure. If the exposure reconstruction performed by an expert is judged under *Daubert* from a logician's perspective, the expert testimony will survive only if it is beyond a reasonable doubt, which means in this case that the overall burden of proof for the plaintiff has been raised to beyond a reasonable doubt. In fact, the burden is so raised in any case in which a verdict for the plaintiff depends primarily on an expert scientific opinion.

Thus, although the Supreme Court's decisions on expert review contain elements of both the process and logician's view, lower courts are expected to interpret the decisions based on the cultural norm, which means scientists will be held to unrealistic expectations. To test whether this pessimism about the *Daubert* standard has any merit, it would be useful to determine the rate at which challenged experts are being excluded under *Daubert*. Unfortunately, sufficient data to answer the question exactly is not readily available, because many federal district court decisions on this subject are not published and because the *Daubert* database in Westlaw contains only a limited set of unpublished opinions. It is possible to come to some conclusions, however, based on data currently available by looking at the exclusion rate of challenged experts in published cases. First, we reviewed a set of federal district civil court cases through a search for opinions containing both the word "Daubert" and "expert"

---

drawings are a Creationist cause célèbre. Early versions show young embryos looking virtually identical in different vertebrate species. On a fundamental level, Haeckel was correct . . . Unfortunately, Haeckel was overzealous. When we compared his drawings with real embryos, we found that he showed many details incorrectly. . . . This does not negate Darwinian evolution. On the contrary, the mixture of similarities and differences among vertebrate embryos reflects evolutionary change in developmental mechanisms inherited from a common ancestor.

M.K. Richardson et al., Letters, *Haeckel, Embryos, and Evolution*, 280 SCIENCE 983, 983 (1998).

134. Examples of websites devoted to criticizing evolution are The Institute for Creation Research (visited Feb. 8, 2001) <<http://www.icr.org>>, The Biblical Creation Society (visited Feb. 8, 2001) <<http://www.pages.org/bcs/index.html>>, and The Center for the Renewal of Science and Culture (visited Feb. 8, 2001) <<http://www.discovery.org/crsc/index.php3>>.

over a three-month period. In the resulting fifty-case sample, approximately 90% of the experts challenged were excluded by the district judge or magistrate. The overwhelming number of challenged experts were testifying on behalf of plaintiffs, which is probably not surprising, given that plaintiffs bear the burden of proof. Next, we looked at a sample of cases discussed in the written decisions of a federal appeals court for the same time period. Once again, the result was that about 90% of the challenged experts had been excluded by district judges.

Perhaps the balance is redressed in the unavailable, unpublished cases.<sup>135</sup> For example, if a judge sustains key witnesses, or even indicates a willingness to sustain them, the case may settle, keeping the judge's decision from being counted in any tabulation based on published opinions. Also, decisions supporting experts may be issued orally. Still, the 90% figure seems high and supports the thesis of this article that current interpretations of *Daubert* will inevitably lead to the exclusion of an extraordinary number of experts in civil cases. Thus, further research is in order. In the meantime, there should be grave concern about the potential impacts of the *Daubert* rule.

Even if this 90% statistic holds up after a deeper analysis, there could be explanations other than a biased *Daubert* process. Perhaps plaintiffs' attorneys have just not yet learned the game or may not have sufficient financial resources to defend against *Daubert* challenges. Maybe defendants only challenge experts whose testimony is particularly implausible. It is difficult, without having the original expert reports in hand, to determine how many of an expert's supposed failings are based on unrealistic expectations by the judge, and how many represent sub-normal practice. Only further research can hope to disentangle the various contributions that each explanatory factor might make to the overall rate at which experts are being excluded. However, the percentage of successful challenges might not be much lower if inept and underfunded

---

135. The Federal Judicial Center found that in 59% of trials in which expert admissibility was disputed, the judge allowed all of the proffered testimony without limitation. See MOLLY T. JOHNSON ET AL., EXPERT TESTIMONY IN FEDERAL CIVIL TRIALS: A PRELIMINARY ANALYSIS (Federal Judicial Ctr. ed., 2000). However, the survey was not, as in our study, restricted to *Daubert* challenges. Because the survey was undertaken in 1998, before the Supreme Court held in *Kumho Tire* that *Daubert* applied to all experts, only a subset of the reported exclusions may have been made on *Daubert* grounds. (Experts from scientific specialties accounted for seven percent of all of the experts; medical/mental health experts accounted for 43% of the total.) See *id.* at 3.

Furthermore, the results only applied to cases that made it to trial. Had a judge emasculated a plaintiff's case by excluding key experts, with the result that the case was dropped or dismissed for lack of expert evidence, the case would not have been counted in the Federal Judicial Center's ("FJC") study. See *id.* at 4. For these reasons, the two studies may not be at all comparable. On the other hand, the FJC study could indicate that the outcomes of some *Daubert* challenges were never memorialized in an opinion, which would have meant they were never counted in our study.

A result closer to our study was obtained in 1999 by Piermattei who found that "in 73 federal product liability/toxic tort cases since 1994 where the plaintiff has sought to admit expert testimony on causation, the district courts admitted only twenty-three experts. Similarly, the federal appellate courts allowed only ten of thirty-one experts to testify as to causation." William Piermattei, *From Frye to Joiner: The Supreme Court Muddies the Waters of Judicial Reasoning*, VT. J. ENV'T. (1999) (visited Feb. 8, 2001) <<http://www.vje.org/roscoe/99/roscoe99c.html>>.

Note that Piermattei does not give details on how the samples he studied were chosen, making it difficult to compare the results precisely with our study.

experts were replaced with the best ones available. A legitimate aura of falsehood and doubt can be created about most any complex expert testimony that has appeared in a court case—a fact that would lead a judge inexperienced with the realities of science to conclude that an expert's testimony was flawed and therefore could not be valid science.

#### IV

##### SQUARING THE UNAVOIDABLE SUBJECTIVE ELEMENTS IN SCIENTIFIC RESEARCH WITH THE VIEW OF SCIENCE AS FORMAL LOGIC: PROPOSALS FOR CHANGE

The Supreme Court has so far failed to synthesize the two views of science—process versus formal-logic that are in tension in *Daubert* and successive opinions. One easy solution, of course, is to shift language from the absolute to the conditional. For example, “if the residual assumptions and inferences made within the research methodology are true (within certain limits), then the results are true (within certain limits).” The conditional nature of scientific statements leads to an uncertainty term, or rate of error, that must be included with the final answer. When the uncertainty terms or rate of error contributed by each assumption or inference are known with high confidence, say from experiment, it is possible to combine the effects of all of them into one uncertainty range that will apply to the conclusion(s).

In many ways, an expert's uncertainty analysis is the most important part of quantitative testimony. Although recognized as one factor under *Daubert* in judging scientific validity,<sup>136</sup> it has not been given the prominence it deserves. The expert's best estimate—for risk, probability, or exposure—is of interest, but it is usually meaningless without an assigned uncertainty range. In effect, if computed fairly, the uncertainty range (rate of error) incorporates most concerns about scientific validity. It gives operational meaning to the concept of scientific validity, at least in those situations where uncertainty ranges are relevant. If scientists were formed as a jury to decide on the scientific validity of expert testimony, their focus would quickly narrow to the assigned uncertainty range. They would examine whether, in their view, the uncertainty range reasonably incorporated the limitations in the underlying methodology. If so, it would be up to the legal system to decide whether the uncertainty range was too large to be useful for legal purposes—a decision that presumably would depend in part on whether a civil or criminal case was involved. A jury of scientists assessing expert testimony would also likely use the uncertainty range assigned by different experts to determine whether there were really any significant differences between their predictions. For example, experts whose

---

136. Uncertainty is introduced by the Supreme Court using the words “rate of error”: “Additionally, in the case of a particular scientific technique, the court ordinarily should consider the known or potential rate of error.” *Daubert v. Merrell Dow Pharms., Inc.*, 509 U.S. 579, 594 (1993). The non-scientist is probably most familiar with the concepts of uncertainty and rate of error as they are stated in the context of public opinion polls, namely as a “margin of error.”

uncertainty ranges overlap do not really differ significantly in the scientific sense, even if their best estimates seem far apart. Choosing between them becomes highly subjective and may be best left to a real jury, because a choice will probably come down to nonscientific factors, such as character, perceptions of intelligence, and possible conflicts of interest.

Thus, one part of synthesizing the process view with the formal-logic view of science is to place greater emphasis on rate of error, because it provides an operational way to measure scientific validity. This is not a complete step for two reasons, however. First, error or uncertainty analysis cannot be done with the 100% rigor that a logician might demand because it requires assumptions and inferences. In a legal proceeding, someone would still have to judge the reasonableness of the uncertainty analysis. If shifting the battles in a *Daubert* challenge meant only that flawless logic would be required for uncertainty analysis, little progress would have been made.

Scientists do not always know the error range very well for all of the underlying assumptions and inferences. To include this class of assumptions into the overall uncertainty analysis, scientists have come up with a pragmatic approach, called Bayesian analysis.<sup>137</sup> In effect, one makes inferences about the uncertainty attached to one's premises:

The Bayesian answer . . . says roughly this: If you can assign a degree of certainty, or personal probability, to the premises of your argument, you may use any of the rules of probability theory to derive a certainty for the conclusion, and this certainty will be a logically valid consequence of your original certainties . . . The Bayesian approach represents a constructive attempt to deal with the dilemma that scientific laws and facts should not be treated as known with certainty, whereas classic deductive logic yields conclusions only when some law, fact, or connection is asserted with 100% certainty . . . The catch is that your concluding certainty . . . may heavily depend on what you used as initial conditions, or prior probabilities. And if those initial certainties are not those of a colleague, that colleague may very well assign a certainty to the conclusion different from the one you derived.<sup>138</sup>

The question then arises: What advantage is gained by adding new inferences to estimate the uncertainties in one's original inferences? First, there is usually less disagreement among scientists about the uncertainty range of individual factors in the analysis than there is about the best estimate or best inference.<sup>139</sup> Second, many of the inferences often turn out to have a negligible effect on the outcome, and can therefore be ignored.<sup>140</sup> Third, the combination of uncertainties from many inferences and assumptions tends to wash out in part, since some raise the final answer and some lower it.<sup>141</sup> In other words, extreme

137. See ROTHMAN & GREENLAND, *supra* note 5, at 21.

138. *Id.*

139. If an analyst does come up against criticism from peers about the uncertainty assigned to a particular assumption or inference, the concern can usually be accommodated by increasing the individual uncertainty range.

140. See M. GRANGER MORGAN & MAX HENRION, *UNCERTAINTY, A GUIDE TO DEALING WITH UNCERTAINTY IN QUANTITATIVE RISK AND POLICY ANALYSIS* 39-40 (1990).

141. This property of combined uncertainties is a result of what statisticians call the "Central Limit Theorem." See ROGER PORKESS, *THE HARPERCOLLINS DICTIONARY OF STATISTICS* 26 (1991).

combinations of uncertainties, those tending toward a worst-case scenario, are unlikely to occur.

Nevertheless, one always ends up with some assumptions and inferences that cannot be handled through either formal error analysis or Bayesian error analysis. For example, consider the use of DNA testing in the courtroom to match samples containing DNA to an individual. Initially, critics attacked the way probabilities for a DNA match were being stated, asserting that an undue confidence in the results was being assumed due to significant uncertainties around the match probabilities that most courts used.<sup>142</sup> A committee of the National Research Council proposed a conservative way to estimate the match probabilities, using upper bounds on “allele” frequencies and racial makeup that made it very unlikely that uncertainty remained a problem.<sup>143</sup> This proposal ended the controversy. However, DNA analysis, even with conservative estimates of match probabilities, still relies on the assumption that our current understanding of DNA and genes is correct.

It is again necessary to assign some personal probability to the residual assumptions and inferences that cannot be handled in any other way. This class of assumptions and inferences is referred to as “residuals.” These assumptions can neither be handled by propagating their known uncertainties to the final answer, nor by consensus among scientific disputants. To deal with the residual assumptions and inferences in law, it is useful to consider the following modification of the conditional statement with which this section began: If the residual assumptions and inferences made within the research methodology are deemed by the fact finder to be plausible (or reasonable, or very likely to be true, or more likely to be true), then the results are also to be deemed plausible (or reasonable, very likely to be true, or more likely to be true).<sup>144</sup>

This is a plausible interpretation of *Daubert* that reconciles the inconsistencies in the original opinion. If the Supreme Court had clearly articulated such a

---

142. “This report describes both the science behind DNA profiling and the data on the frequency of profiles in human populations, and it recommends procedures for providing various statistics that may be useful in the courtroom.” The committee made a series of detailed, scientific recommendations for estimating probabilities, the first of which was

in general, the calculation of a profile frequency should be made with the product rule. If the race of the person who left the evidence-sample DNA is known, the database for the person’s race should be used; if the race is not known, calculations for all the racial groups to which possible suspects belong should be made. For systems such as VNTRs, in which a heterozygous locus can be mistaken for a homozygous one, if an upper bound on the frequency of the genotype at an apparently homozygous locus (single band) is desired, then twice the allele (bin) frequency,  $2p$ , should be used instead of  $p^2$ . For systems in which exact genotypes can be determined,  $p^2 + p(1 - p)\emptyset$  should be used for the frequency at such a locus instead of  $p^2$ . A conservative value of  $\emptyset$  for the US population is 0.01; for some small, isolated populations, a value of 0.03 may be more appropriate. For both kinds of systems,  $2p_{HP}$  should be used for heterozygotes.

NATIONAL RESEARCH COUNCIL, THE EVALUATION OF FORENSIC DNA EVIDENCE (1996) [hereinafter NRC, *Evaluation*].

143. *See id.*

144. *See discussion supra* Part IV.

view, it would no longer be appropriate for parties in a *Daubert* proceeding to argue whether the assumptions embedded in an expert's report were true. Instead, opposing counsel would be arguing whether the residual assumptions met the relevant criteria, such as reasonable or more likely than not. The criterion to be substituted in the conditional statement above would presumably depend on the nature of the legal action and the importance of a particular methodology in the case. The appropriate criterion is for the courts and legislatures to choose, not the scientists.

The conditional approach thus provides a conceptual way to bridge the gap between the logicians and those who see science as process.<sup>145</sup> It is not new, having been routinely used in the field of "risk assessment" for many years.<sup>146</sup> Risk assessment evolved in the conditional direction to handle the recognition that subjectivity and professional judgment are unavoidable in relating health effects to toxic emissions. The task of judging the reasonableness of the residual assumptions and inferences is passed on to the policy makers or to whom-ever requested the risk assessment.<sup>147</sup> If this Bayesian approach were used in law, the fact-finder or *Daubert* judge would make an assessment of the certainty of the unprovable assumptions and inferences. Possibly, guidance from scientists could be of assistance with this task.

#### A. Proposals for Synthesizing the Two Schools

There are at least three ways that an assessment of the residual scientific assumptions and inferences could be handled.

*1. Make Use of Generic Recommendations from Scientific Committees.* At least for generic issues that crop up over and over again, courts could seek advice from balanced committees of scientists established by professional societies on such issues as animal testing, meta-analysis, and exposure reconstructions.

It is especially inefficient to have each separate court repeatedly wrestle with the same scientific issues. For generic assumptions, such as those needed to apply animal data to humans or to use meta-analysis reliably, the federal judiciary could get professional assistance. It could call on the National Research Council ("NRC") to establish committees to give guidance as to what assump-

---

145. The conditional approach is, at least, useful in situations where there are only a limited number of scientific experts needed to prove a legal point. In the situation where a number of pieces of scientific evidence must be chained together, the likelihood of a final answer is the product of the likelihoods of each piece of the chain. Thus, if the opinion of five different scientists had to be reviewed for evidentiary reliability, and each opinion were part of the causal chain of logic, the overall likelihood might only be "more likely than not," even though each opinion were judged to be likely beyond a reasonable doubt. Should a group of scientists be making the assessment of scientific assumptions in a case, this problem could be handled in a straightforward manner, which argues in favor of the panel idea mentioned later.

146. Risk assessment attempts to quantify the risk to human and ecological health from various activities, such as release of toxic substances from a facility.

147. Typically, risk assessors will explore the sensitivity of their results to variations in initial assumptions, but choosing a range for each input. The choice of range, then, becomes the subjective part of the input.

tions are reasonable for each generic methodology.<sup>148</sup> Only with such guidance could courts make sensible decisions about excluding scientific testimony put forward by qualified experts in fields where there is not a true consensus.

In the absence of guidance from a balanced committee of scientists, the courts should rarely exclude proposed testimony, only doing so in egregious cases where, for instance, the expert continually fails to provide a basis for his or her opinions or concedes an error that cannot be corrected.<sup>149</sup> The court should recognize, instead, that the different scientific assumptions that experts make in these cases fall into the “range where experts may disagree.” The fact-finder—usually the jury—should be instructed to do its best to establish (1) which of the experts truly believe their stated assumptions; (2) which expert has the most credible assumptions; and (3) which expert seems most credible overall.

2. *Court-appointed Advisor.* Although it is preferable to view scientific experts as imperfect eyewitnesses to the scientific enterprise and to have judges spend their time instructing juries about the inherent limitations of scientists or, when available, comparing an expert’s assumptions to guidelines established by a balanced professional committee, some judges may not be willing to take a hands-off approach in the situations discussed here. Some commentators have suggested a court-appointed scientific advisor.<sup>150</sup> A court-appointed scientific advisor would provide guidance on the reasonableness of the assumptions and inferences made in the case at hand. There are problems, however, with the appointment of a single advisor. For instance, a single advisor may belong to a particular school of thought, which could lead to large variation in the guidance provided by different advisors given the contested nature of scientific knowledge.<sup>151</sup>

---

148. The NRC has recently established a permanent committee on Science, Technology, and Law, which would likely be the oversight committee for subcommittees created to focus on specific tasks of interest to the judiciary. Co-Chairs of the committee are Donald Kennedy, Bing Professor of Environmental Science, Stanford University; and Richard Merrill, Daniel Caplin Professor of Law, University of Virginia School of Law. National Research Council studies are not inexpensive, so congressional support might be necessary to fund them.

The NRC’s report on DNA testing is an example of an effort to provide guidance. See NRC, *Evaluation*, *supra* note 142. NRC committees focused on scientific issues relevant to the judiciary will produce more balanced reports if they include a number of scientists who have defended a *Daubert* challenge, as well as a number of scientists who have filed affidavits in support of a *Daubert* challenge. Although it sometimes takes some nudging by concerned parties to achieve balance, the NRC will usually respond appropriately when a gap in representation is pointed out.

149. An expert should always be given an opportunity to provide a basis for a challenged assumption, as well as to correct errors or temper a conclusion. As for “errors” in data or logic that an expert refuses to admit or correct, the gatekeeper judge should ensure that opposing counsel have (1) ample opportunity to explore the expert’s reasoning (or lack of it), and (2) ample opportunity to “expose” the weaknesses to the jury. Only if the judge rules that the supposed errors are too complex for a jury to comprehend should an expert in such a situation be excluded.

150. See Breyer, *supra* note 14, at 6.

151. See Ellen E. Deason, *Court-Appointed Expert Witnesses: Scientific Positivism Meets Bias and Deference*, 77 OR. L. REV. 59 (1998).

3. *Teams of Scientific Reviewers.* To get around the problems surrounding a single court advisor, it might be desirable to move to a review process, much like the process used by journals to peer review submitted articles.<sup>152</sup> Although the following proposal would increase the parties' court costs and could no doubt be improved and streamlined, it does provide what scientists call an "existence proof," namely proof that there is at least one solution to the problem at hand. At the very least, the proposal may encourage further research and stimulate new thinking.

The foundational premise is that a jury of scientists would be the gold standard for deciding whether expert testimony should be considered scientifically valid. If anyone is to make such a subjective decision, it should be scientists. All arguments that favor a judge over a jury fade away, provided that this hypothetical jury of scientists were given as much time as a judge to consider the issues and could direct questions to the experts. Scientists would not be intimidated by experts; nor would they be intimidated by the material. Furthermore, the scientists would have a context for rating the testimony against standard scientific practice. Having a group of scientists rather than an individual decide whether expert testimony is scientifically valid reduces that chance that individual bias or temperament would skew the decision.

Of course such an ideal situation would be extremely burdensome on the community of scientists, and rebellion could be anticipated. Fortunately, the essence of a scientific jury for expert testimony could be retained, while reducing the burden on impressed scientists,<sup>153</sup> by allowing the panel to review expert testimony in their offices in the same way that they review papers for a journal. To make the reviewing job easier, testifying scientific experts could be required, as part of their report, to lay out the parts of their argument that required subjective judgments. For example, the experts might be required to (1) list the assumptions and inferences that were needed to make their expert testimony hold together in a seamless, logical chain, and (2) estimate the uncertainty (rate of error) introduced into any quantitative conclusions. In principle, these steps would make it easy for a scientifically educated reviewer to reach a conclusion, albeit a subjective one, based on the reasonableness of an expert's testimony. The gatekeeping judge could divide the list of assumptions into two groups. The first set (Set I) would contain those assumptions and inferences made by

---

152. For other ideas concerning how scientists might provide reviews for a court, see Lawrence S. Pinsky, *The Use of Scientific Peer Review and Colloquia to Assist Judges in the Admissibility Gatekeeping Mandated by Daubert*, 34 HOUS. L. REV. 527 (1997).

The first suggestion is to give judges the option to obtain external scientific peer review of the proffered evidence. As an alternative, or in addition to such peer review, the second suggestion is that judges use colloquium-type preliminary proceedings to get the background information and education needed to make an informed decision.

*Id.* at 529.

153. Conceivably, scientists who have pre-registered for this function could be used. For example, in January of 1999, the Private Adjudication Center of Duke University actively began to establish a registry of independent scientific and technical experts who are willing to advise courts.

the expert that the judge decided were within the capabilities of the jury to analyze for reasonableness. An example of a possible determination to be left for the jury would be whether it was reasonable for the expert to discount certain data collected by the defendant on the grounds that it appeared to have been fudged. Judging the reasonableness of such an assumption does not require scientific training.

The second set of assumptions and inferences (Set II) would not be given to the jury, but would be sent in writing to a randomly chosen set of, perhaps, five scientists identified by a scientific society. Attached would be a definition of the criteria that the reviewers were to use in judging the assumptions and inferences for the particular case, for example, “reasonable,” “very likely,” or “beyond a reasonable doubt.” These scientists would be asked to write a review, much like they do when asked to review a scientific paper for a journal.<sup>154</sup> They would comment on the reasonableness and likelihood of each assumption and inference, as well as on the overall assignment of quantitative uncertainty. They would be asked to make their judgments in the context of the case as presented, in writing, to them by the judge. They would accept as valid all of the Set I assumptions, because Set I would be judged by the actual jury, if the testifying expert got that far.

If a reviewer found that a particular assumption in the context of the case failed to meet the assigned criteria—that is, was not reasonable—the reviewer would indicate in writing the changes in language or analysis that would be required for him or her to accept the assumption or inference as meeting the assigned criteria. Just as with a paper submitted to a journal, a challenged expert would be shown the reviews, have the option to revise his or her report, and re-submit it. If, on the second pass, a majority of the reviewers could not state that on balance that the Set II assumptions were reasonable, the expert would not be allowed to testify.

On the other hand, if the reviewers found on balance that the Set II assumptions met the assigned criteria and the uncertainty analysis was reasonable, the expert would pass the first *Daubert* screening and could present his or her testimony to a jury. But the first screening would not be the final test for an expert, because the first set of assumptions would not yet have been scrutinized. At trial, the judge would instruct the jury that it had to find the Set I assumptions reasonable—or whatever criterion was relevant to the case—before it could accept the expert’s testimony as reliable.

The advantage of a proposal like this is that the scientifically subjective parts of an expert’s testimony in a complex case are judged by people whose scientific qualifications to sit as a jury cannot be doubted. Such a process would make the district judge more of a referee than a gatekeeper in some ways, although

---

154. Kapsa & Meyer argue that some of this review function is already available in the civil law systems in Continental Europe, which leaves “ample time for obtaining written reviews by local, and sometimes world-wide professional peers and for further responses that are available to the trier of facts for review prior to trial.” Kapsa & Meyer, *supra* note 40, at 317 n.36.

the role of dividing up the questions and ensuring that the experts properly listed their assumptions, inferences, and uncertainty factors would not be trivial. Finally, the *Daubert* judge would still have to rule on relevance and other legal factors, which are squarely within the court's expertise. The judge would also have to decide if the assigned uncertainty and caveats were too great to be of legal value.

This proposal serves as an example of how "junk science" can be kept out of the courtroom, while preserving the advantages of the jury system, such as averaging out the biases of individual decision makers. However, the proposal fails to correct two crucial problems with the current *Daubert* review. As Frank McClellan has pointed out, a *Daubert* review may preclude a jury from lowering the barrier for causation when a civil defendant has acted improperly,<sup>155</sup> and *Daubert* does not allow for some of the nonlegal purposes of a jury trial.<sup>156</sup> Nothing would change in this regard with the proposed substitute, at least with regard to the science.

The above alternatives to the amateurish nature of the current expert review are meant solely as a demonstration. This demonstration is intended to show that, in principle, it is not necessary to tolerate either the conceptual problems with *Daubert* or the replacement of juries as the evaluators of subjective decisions about scientific testimony. Nor is it necessary to take complex cases from a jury completely, as proposed by Paul Miller and Bert Rein, and assign them to a special court modeled after the patent courts.<sup>157</sup> After all, a jury of peers should be making the subjective decisions in law as often as possible.

## B. The Problem of Scientific Uncertainty

One last issue must be addressed in connection with expert review. Given the incomplete nature of science, there will be legal cases where experts on both sides are eminently reasonable and credible.<sup>158</sup> Possibly, neither side is completely right nor completely wrong. Because the plaintiff has the burden of proof, does that mean that a civil case should be thrown out by a reviewing judge in such a situation? Certainly not, from a policy perspective.

A de facto policy of blindly excluding expert testimony in the face of uncertainty would have very serious ramifications. Such a policy would encourage potential defendants to increase uncertainty as much as possible and would en-

---

155. "The Bendectin progeny violates the important tort policy that tailors the burden of proof on causation to deter wrongful conduct." RESTATEMENT (SECOND) OF TORTS § 323(a) (1965). Section 323 relaxes the burden of proof on causation in circumstances where a tortfeasor's failure to fulfill a duty to act 'increases' the risk of harm to the plaintiff." Frank M. McClellan, *Bendectin Revisited: Is There a Right to a Jury Trial in an Age of Judicial Gatekeeping?*, 37 WASHBURN L.J. 261, 274 n.55 (1998).

156. "[A] civil trial is not only a search for historical truth, but also a kind of popularly approved umpired game or ritual for the peaceful settlement of disputes." Farrell, *supra* note 6, at 2206.

157. See Paul S. Miller & Bert W. Rein, *Whither Daubert? Reliable Resolution of Scientifically-Based Causality Issues In Toxic Tort Cases*, 50 RUTGERS L. REV. 563, 568.

158. "After all, different experts, in total honesty, often can interpret the same data differently." Stephen Breyer, *The Interdependence of Science and Law*, 280 SCIENCE 537, 538 (1998).

courage cover-ups and the underfunding of studies to muddy the waters.<sup>159</sup> Furthermore, exclusion of testimony in the face of uncertainty is not a necessary or inevitable conclusion of *Daubert*, at least as interpreted in this article.<sup>160</sup> Uncertainty surrounding evidence occurs regularly in legal cases. True, juries may be awed by experts, as has been claimed by supporters of *Daubert*. If so, the solution in science-dominated cases is to explain more to the jury about the limitations of experts, not to take the cases away from them.<sup>161</sup> Also, a number of reforms could be introduced to give opposing counsel sufficient ammunition to expose a witnesses' weaknesses to a jury, without having to rely on a court to screen out supposedly invalid science. In particular, the nature of expert reports could be revised to require experts to (1) place themselves within schools of thought in their field, and (2) list their assumptions and inferences.

### C. Summaries of Additional Proposals

In addition to the reforms proposed so far in this article, an additional set of modest suggestions are included below in an overall summary, along with proposals made by other commentators.

1. *Earlier Proposals.* A number of proposals have been discussed by others in the legal profession, some of which have already been mentioned: (1) Courts could rely more on directed verdicts to deal with problems of scientific evidence, which would allow for de novo review;<sup>162</sup> (2) The standard for review could be changed to reasonably reliable;<sup>163</sup> (3) The response of third-party

---

159. Manufacturers already have an incentive to avoid learning about negative information. See Wendy E. Wagner, *Choosing Ignorance in the Manufacture of Toxic Products*, 82 CORNELL L. REV. 773, 811-25 (1998); see also Margaret A. Berger, *Eliminating General Causation: Notes Towards a New Theory of Justice and Toxic Torts*, 97 COLUM. L. REV. 2117, 2119 (1997). Once negative information does come out, the current expert review procedures encourage manufacturers to get as many critical reviews into the literature as possible, as well as new studies that contradict the negative ones. This can be done by funding researchers who are unusually skeptical and have particular biases favorable to the company. In addition, manufacturers have an incentive to underfund studies so that the sample sizes are too small to find any results with statistical significance, which then get branded as no-effect studies. See Wagner, *supra*, at 820-25.

160. For a more pessimistic view, see Heidi Li Feldman, *Science and Uncertainty in Mass Exposure Litigation*, 74 TEX. L. REV. 1, 9 (1995) (predicting that exclusion of expert testimony in toxic tort cases would become inevitable once "we fully understand the vision of science that informs *Daubert*"). She argues that reform of tort law may be necessary to handle the problem of scientific uncertainty.

161. See Ronald D. Myers et al., *Complex Scientific Evidence and the Jury*, 83 JUDICATURE 150, 151-52 (1999). An additional part of the solution may be to copy the reforms adopted in Arizona and other states to give juries more of a role in questioning witnesses and therefore a better ability to be their own gatekeepers.

162. In addition, we note that while courts might acknowledge the existence of devices for controlling jury decisions when the evidence is insufficient to permit a rational decision, such as summary judgment, directed verdict and judgment notwithstanding the verdict, the same courts seriously underestimate the utility of these devices in dealing with expert-testimony based on natural or social science evidence.

Scallen & Wiethoff, *supra* note 95, at 1144-55.

163. See Kapsa & Meyer, *supra* note 40.

professional peers could be solicited,<sup>164</sup> optimally with more than one reviewer involved. The proposal by Margaret Berger to substitute “failure to provide substantial information relating to risk” for proof of causation would probably not qualify as a modest change,<sup>165</sup> but it is listed here for completeness.

2. *Rate of Error.* More emphasis could be put on rate of error, as discussed earlier.<sup>166</sup>

3. *Review of Generic Methodologies.* As much as possible, generic methodologies, such as animal testing and meta-analysis, should be reviewed by balanced committees of the NRC to establish guidelines for use in a courtroom. NRC committees might also provide guidance on how to handle the difficult problem of potential biases in the scientific literature. The New England Journal of Medicine (“NEJM”) is one of the few journals that has a policy of eliminating review writers who have received funding from manufacturers, but even the NEJM has had difficulty making this policy work.<sup>167</sup> It is difficult to find people free of conflicts of interest. For instance, a deputy editor of the British Medical Journal has complained that “[i]t’s almost impossible to find a very informed commentator on a medical topic who hasn’t received money from the pharmaceutical industry.”<sup>168</sup> Nobel Prize winner David Baltimore has said that “there are few pure academics left” in molecular biology.<sup>169</sup> The intermingling between engineering faculty and industry is nothing new. Ellen Deason has reviewed studies that show that these potential conflicts actually lead to differences in results: “[R]esearch funded by the chemical industry is more likely than government-funded research to conclude that occupational exposure to chemicals is not hazardous. Only 32% of the industry-sponsored studies found an increased risk of dying, compared to 59% of the studies conducted by other epidemiologists.”<sup>170</sup>

---

164. See *id.*; see also Lawrence S. Pinsky, Comment, *The Use of Scientific Peer Review and Colloquia to Assist Judges in the Admissibility Gatekeeping Mandated by Daubert*, 34 HOUS. L. REV. 529 (1997).

165. See Berger, *supra* note 159, at 2117.

166. See discussion *supra* Part IV.

167. See Constance Holden, *NEJM Admits Breaking its Own Tough Rules*, 287 SCIENCE 1573 (2000).

168. *Id.*

169. Marcia Barinaga, *Asilomar Revisited: Lessons for Today*, 287 SCIENCE 1584 (2000). Also, as reported in Deason, “a recent study of articles on cellular and molecular biology and genetics by authors from institutions in Massachusetts found that lead authors of one-third of the papers had a financial interest in the research they were describing.” Deason, *supra* note 151, at 112 (discussing Sheldon Krinsky et al., *Financial Interests of Authors in Scientific Journals: A Pilot Study of 14 Publications*, 2 SCIENCE & ENGINEERING ETHICS 395, 402 (1996)).

170. Deason, *supra* note 151, at 110 n.205 (citing G.M.H. Swaen & J.M.M. Meijers, *Influence of Design Characteristics on the Outcome of Retrospective Cohort Studies*, 45 BRITISH J. INDUS. MED. 624, 626 (1988)).

Similarly, clinical trials comparing two nonsteroidal anti-inflammatory drugs almost always report that the drug manufactured by the sponsor is more effective and less toxic, although these claims are often not supported by the data.<sup>171</sup>

A committee of the NRC could help make these potential biases widely known as well as give guidance concerning how to handle them in the courtroom. Without such generic assistance, it will be very difficult for anyone but the most highly funded parties to deal with these issues in an individual case.

4. *Education.* Courts can become better educated about the limitations of science, including its subjective elements. In addition to learning about the conditional nature of scientific papers and testimony, the courts need to learn relevant information about the concept of statistical power.<sup>172</sup>

5. *Burdens for Acceptance.* *Daubert* proceedings for civil cases should require different criteria for acceptance of expert reports than is required for criminal cases, allowing for the different level of certainty that is required for the fact finder to reach a judgment. The level of confidence expected in the assumptions underlying expert testimony should be higher in criminal cases than in civil cases.

6. *Formal Review.* Courts can adopt some of the formal methods used by journals that review submitted articles. For instance, experts should be given a chance to rewrite their reports completely after hearing critiques from both opposing experts and the reviewing judge. It is unscientific to make technical criticisms, as some judges do in their *Daubert* opinions, to which experts cannot respond. Similarly, experts should have the opportunity to respond to issues raised in depositions. Although cross-examination in depositions is appropriate for legal purposes, it is not conducive to good science. Scientists need time for reflection if they are to have the ability to respond carefully.

A peer-reviewed journal could be created to focus on science in the courtroom and publish some expert testimony. Currently, much of the testimony prepared for the courtroom, even in complex cases, is too mundane for its contents to be of interest to standard scientific journals. In time, a specialized journal could help to define good scientific and engineering practice for experts testifying in legal proceedings, just as in the field of risk assessment. Such a journal would also provide an avenue for pointing out scientific errors made by judges.

---

171. See *id.* (citing Paula A. Rochon et al., *A Study of Manufacturer-Supported Trials of Nonsteroidal Anti-inflammatory Drugs in the Treatment of Arthritis*, 154 ARCH. INTERNAL MED. 157 (1994)).

172. "Power" is a statistical concept that indicates the extent to which a study is (or was) capable of detecting a statistically significant result. Because studies that do not find a statistically significant effect, even if they find a trend, are branded as null studies, it is clearly important to know the statistical power of a study reporting a "null" result. Unfortunately, this is rarely done, which adds a special burden on plaintiffs in a toxic tort case who must deal with supposedly null-studies.

7. *Self-identification.* Experts should be asked in their reports to place themselves within schools of thought in their field, so that everyone is made conscious of the inevitable disagreements that exist among even the best scientists.

8. *Use of Science Writers.* Courts could make use of scientific writers to help judges and juries understand points of view within various branches and sub-branches of science.

9. *Exchange of References.* When reports are to be filed simultaneously, experts should be required to exchange references prior to submission of reports. Requiring such an exchange would improve the quality of reports as well as narrow some of the differences between experts.

10. *Empirical Research.* Research into the impact of the *Daubert* procedures is urgent. Experiments with mock trials and mock in limine hearings should be used to test the effectiveness of existing procedures and proposed reforms. The use of panels of scientists could be helpful in comparing judgments reached by lay persons and scientists, although the criteria for judgment given to scientists would be crucial in determining the outcome.

11. *Rejecting Expert Testimony.* Finally, it is important to recognize and correct the unfair treatment of scientists that can occur under the current system. As discussed, withholding judicial criticisms from experts remains a problem. It is particularly outrageous to expect that all responses to criticism should have been in the original report, as if an expert in a finite report could anticipate everything that an opposing counsel could attack. Just as serious is branding an expert's testimony invalid or unreliable as a matter of science. This problem could be alleviated if judges would couch their rulings on experts in terms of the law, for instance, by using language such as, "the testimony is unreliable from an evidentiary point of view for use in this case."<sup>173</sup> If judges frame their decisions in terms of law, the grounds for their opinion will be more reliable and those opinions will be more fair.

## V

### CONCLUSION

Subjective elements and professional judgment cannot be removed from expert scientific testimony, but can only be highlighted and assessed for reasonableness. Both judges and juries need to be educated on the limitations of sci-

---

173. Some commentators have argued that an expert should be embarrassed if a lay judge excludes their testimony. Jan Beyea has vigorously disputed that view, arguing that it ignores the fact that a judge may not understand the science, may have case-related biases, and may actually have excluded testimony for reasons of relevance. See Jan Beyea, *Scientists in the Courtroom*, 284 SCIENCE 1125 (1999) (letter to the editor). Such partisan views will encourage scientists not to participate in the courtroom. See *id.*

entific experts. The current paradigm adopted by many courts that science falls into a formal, logical structure is naïve. At best, scientists are imperfect eye-witnesses to the scientific enterprise. They need to be judged by juries accordingly. Under a *Daubert*-style review, scientific experts should be judged on reasonableness, instead of subjected to a de novo review by amateurs that can dig down to any level of science. If necessary, assessment of reasonableness of the scientific assumptions in testimony should be carried out by scientific reviewers, not lay judges.

The Supreme Court's opinions in the *Daubert* Trilogy<sup>174</sup> contain two inconsistent views of science, "process" vs. "formal logic." It is time that the two views are synthesized into a consistent whole. In the absence of a synthesis of the process and logic views of science, the inconsistencies within the Supreme Court's rulings on science are likely to lead to the enshrinement of scientific error in the law, as has already occurred with the doubling-of-risk standard adopted in some jurisdictions. A Bayesian approach using expert opinions expressed as conditional statements may provide the synthesis of the process and logic views needed for science-dominated civil cases. Certainly, greater focus on rate of error as an operational measure of scientific validity would be helpful. Reform of expert review should attempt to mimic processes followed by the scientific review process. Research into the impact of *Daubert* should be increased. Experiments with mock trials and mock in limine hearings should be used to test the effectiveness of existing procedures and proposed reforms. To the greatest extent possible, judicial decisions on expert testimony should be grounded in legal reasoning instead of on scientific reasoning alone.

---

174. See discussion *supra* notes 1-3.