

An Epistemologist in the Bramble-Bush: At the Supreme Court with Mr. Joiner

Susan Haack
University of Miami

Think before you think! [Stanislaw Lec]¹

“Judges Become Leery of Expert Witnesses,” ran a headline in the *Wall Street Journal* a couple of years ago; they are “Skeptical of Unproven Science”—the “Testimony of Dilettantes” (Schmitt 1997). Intrigued, I began to struggle through thickets of details of exploding tires, allegedly poisonous perfumes, leaking and bursting breast implants, contaminated insulating oil, etc., etc., and through legal developments from *Frye* through the Federal Rules of Evidence to *Daubert*, until eventually I found myself at the U.S. Supreme Court with Mr. Joiner, eavesdropping as the justices—for all the world like a conclave of medieval logicians—disagreed among themselves about whether there is a Categorical Distinction between methodology and conclusions.

Now *that*, I thought, certainly sounds like the kind of question to which

This article is adapted from a paper presented at a conference on epistemology and the law of evidence organized by the School of Law and the Department of Philosophy at the University of North Carolina–Chapel Hill, and in the Schools of Law at Boston University, the University of Pennsylvania, the College of William and Mary, the University of Iowa, the University of Virginia, Creighton University, and the University of Maryland. It was also discussed with faculty in the School of Law at Duke University, and was distributed as part of the briefing packet for a workshop on science-based medical evidence organized by the Institute of Medicine, National Academy of Sciences. I would like to thank Paul Gross, Richard Hausler, Robert Heilbroner, Mark Migotti, and Edgardo Rotman for reading this article in draft form and giving me their reactions; Clare Membiela and Janet Reinke of the University of Miami Law Library for their help in locating relevant materials; and the students in my class on Scientific Evidence in Theory and in Court, who taught me a lot.

1. Gross 1983: 262.

Journal of Health Politics, Policy and Law, Vol. 26, No. 2, April 2001. Copyright © 1999 by Susan Haack.

an epistemologist or philosopher of science ought to be able to make a contribution; and, in due course, I shall have something to say about it. But I soon realized it was only the tip of a very large iceberg.

By now, scientific evidence of just about every kind (from DNA fingerprinting to battered-wife syndrome, from studies of mice injected with potentially carcinogenic chemicals to recovered memories) plays a large and apparently ever-growing role in both criminal and civil cases. The long and tortuous history of efforts to ensure that when the legal system relies on scientific evidence it is not flimsy speculation but decent work, suggests that this interaction of science and the law raises some very tricky problems. And to judge by how often, in that long and tortuous history, explicit or implicit assumptions about the nature of scientific knowledge and the character of scientific inquiry are crucial, those problems are in part epistemological.

The epistemological issues intersect, of course, with problems of other kinds. Peter Huber is preoccupied with greedy tort lawyers hoping to earn huge contingency fees by winning cases with “junk science,”² Kenneth Cheseboro with heartless corporations hoping to avoid compensating the victims of their profitable but dangerous products (Cheseboro 1993). I’m afraid both have a point. Both are well aware, however, that there is something about scientific evidence that encourages and enables the operation of such unsavory motives.

Almost a century ago, Learned Hand argued that the role of the expert witness—who not only may but must offer his opinion, draw conclusions—is anomalous, for if each party presents its own expert witness(es), the jury must decide “between two statements each founded upon an experience foreign in kind to their own”—when “it is just because they are incompetent for such a task that the expert is necessary at all” (Hand 1901: 54). Only a couple of years ago, Justice Breyer—concerned with scientific evidence specifically rather than with expert evidence generally, and focused less on the jury than the judge, on whom a significant gatekeeping burden now falls—suggested an essentially similar diagnosis. Reflecting that *Daubert* requires judges “to make subtle and sophisticated determinations about scientific methodology,” he observes that “judges are not scientists, and do not have the scientific training that can facilitate the making of such decisions.”³

In 1901, Hand had suggested court-appointed experts; in 1997, in his

2. Huber 1991, 1992; Huber and Foster 1997.

3. *General Electric Co. v. Joiner*, 522 U.S. 136, 148, 118 S. Ct. 512, 520 (1997) (Breyer, J., concurring).

concurring opinion in *Joiner*, Justice Breyer urged that judges make more use of their power under Federal Rule of Evidence 706 to appoint scientists to advise them. But, as Hand himself had observed earlier in his article, when there are expert witnesses on both sides we ask the jury to decide “*where doctors disagree*” (ibid.: 154; emphasis mine). And now it begins to appear that there is a problem beyond judges’ or juries’ inability fully to understand scientific evidence. Many scientific claims and theories, at some point in their career, occupy that large grey area of the somewhat-but-far-from-overwhelmingly warranted; so sometimes the scientific determinations that judges or juries are asked to make may be so subtle and sophisticated, so manifold and tangled, that even those competent in the relevant area of science may legitimately disagree—or may agree that there is too little evidence, that they just don’t know.

Legal efforts to winnow decent scientific evidence from the chaff, I shall argue, have often been based on false assumptions about science and how it works. It doesn’t follow, unfortunately, that if we had a better understanding of science, all problems could be easily resolved. A better understanding of scientific evidence and inquiry will reveal why it has proven so difficult to find a legal form of words that will ensure that only decent scientific evidence is admitted, or a simple way to delegate some of the responsibility to scientists themselves; but rather than suggesting any easy solutions it accentuates the need to think hard and carefully about what goals we should be trying to achieve, and what kinds of imperfection in achieving them we are more willing, and what we are less willing, to tolerate.

Here I can offer only some preparatory steps toward such re-thinking: a brief account, first, of scientific evidence and its special complexities; and then—as I cautiously approach that bramble-bush with my philosophical pruning-shears—a brief epistemological commentary on the legal mechanisms that have been devised to handle scientific evidence in court. But I hope, by cutting away some overgrown epistemological dead-wood, to clear the way for potentially healthier new growth.

■ ■ ■

In their descriptive use, the words “science,” “scientific,” etc., refer to a loose federation of disciplines including physics, chemistry, biology, and so forth, and excluding history, theology, literary criticism, and so on. But they also have an honorific use; “scientific,” and “scientifically,” especially, are very often all-purpose terms of epistemic praise, vaguely conveying “strong, reliable, *good*.” They play their honorific role when

the credulous are impressed by actors in white coats assuring them that new, scientific Wizzo will get clothes even cleaner, or that new Smoothex is scientifically proven to get rid of wrinkles faster; and no less so when, skeptical of some claim, people ask: "Yes, but is there any scientific evidence for that?"

Unfortunately this dual usage, descriptive and honorific, has encouraged a damaging preoccupation—especially in Popper and among his admirers—with the "problem of demarcation," of distinguishing *real* science from pretenders.⁴ It has distorted our perception of the place of the sciences within inquiry generally, and disguised what would otherwise be obvious facts: that neither all nor only scientists are good, honest, thorough inquirers; and that scientific claims and theories run the gamut from the thoroughly speculative to the very firmly warranted.

Natural-scientific inquiry is continuous with other kinds of empirical inquiry. The physicist and the investigative journalist, the X-ray crystallographer and the detective, the astronomer and the ethnomusicologist, etc., etc., all investigate some part or aspect of the same world. And scientists, like detectives, or historians, or anyone who seriously investigates some question, make an informed conjecture about the possible explanation of a puzzling phenomenon, check out how well it stands up to the available evidence and any further evidence they can lay hands on, and then use their judgment whether to give it up and try again, modify it, stick with it, or what.

Nor is there any "scientific method" guaranteeing that, at each step, science adds a new truth, eliminates a falsehood, gets closer to the truth, or becomes more empirically adequate. Scientific inquiry is fallible, its progress ragged and uneven. At some times and in some areas, it may stagnate or even regress; and where there is progress, it may be of any of these kinds, or it may be a matter of devising a better instrument, a better computing technique, a better vocabulary, etc.

As human cognitive enterprises go, natural-scientific inquiry has been remarkably successful. But this is not because it relies on a uniquely rational method unavailable to other inquirers; no, scientific inquiry is like other kinds of empirical inquiry—only more so. As Percy Bridgman once put it, "the scientific method, so far as it is a method, is doing one's damndest with one's mind, no holds barred" (Bridgman 1955: 535).

Scientific inquiry is "more so" in part because of the many and various

4. For a brief summary of controversies in recent philosophy of science, see Haack 1995 and Haack 1996. McErlean 2000 is a useful anthology.

helps⁵ scientists have devised to extend limited human intellectual and sensory powers and to sustain our fragile commitment to finding out: models, metaphors, and analogies to aid the imagination; instruments to aid the senses; elaborate experimental set-ups to aid in testing and checking by flushing out needed evidence; mathematical, statistical, and computing techniques to aid our powers of reasoning; and a tradition of institutionalized mutual disclosure and scrutiny that, at its best, enables the pooling of evidence and helps keep most scientists, most of the time, reasonably honest.

E. O. Wilson describes his work on the pheromone warning system of red harvester ants: collect ants; install them in artificial nests; dissect freshly killed workers, crush the tiny gobbets of white tissue released, and present this stuff, on the sharpened ends of applicator sticks, to resting groups of workers: they “race back and forth in whirligig loops.” Enlist a chemist, who uses gas chromatography and mass spectrometry to identify the active substances, and then supplies pure samples of identical compounds synthesized in the laboratory. Present these to the ant colonies: same response as before. Enlist a mathematician, who constructs physical models of the diffusion of the pheromones. Then design experiments to measure the rate of spread of the molecules and the ants’ ability to sense them (Wilson 1999: 69–70).

This illustrates both the continuity of scientific inquiry with other kinds of inquiry, and the remarkable persistence with which good scientists go about solving one problem with the help of solutions to others.⁶ Of course, that carries risks as well as rewards; the earlier results on which a scientist builds could turn out to be mistaken, and possibly in ways that undermine his work. Scientific helps depend on substantive assumptions, and our judgments of their reliability depend on our background information—e.g., our reasons for thinking that gas chromatography reliably indicates chemical composition.

Still, fallible and imperfect as they are, by and large those helps *have* helped, enormously: helped to stretch scientists’ imaginations, to enable their powers of reasoning, to extend their evidential reach, and to stiffen their respect for evidence. Almost every day, it seems, the natural sciences come up with new and better technical helps (from chemical assays through statistical modeling to computer programs). But there are no grounds for complacency. As science has become so expensive that only

5. The term, and the idea, come from Bacon 1620.

6. I borrow this happy phrase from Quine 1995: 16.

222 Journal of Health Politics, Policy and Law

governments and large industrial concerns can afford to support it, as career pressures grow, so too does the temptation to exaggerate results or ignore awkward evidence for the sake of money, prestige, or an easy life.

Like the evidence with respect to any empirical claim, the evidence with respect to a scientific claim includes both experiential evidence (someone's seeing, hearing, etc., this or that) and reasons (background beliefs) ramifying in all directions; and, as "with respect to" was chosen to indicate, normally includes both positive evidence and negative. But, again, it is "more so"—in the complexity of its ramifications, in the dependence of its experiential components on instrumentation, in the pooling of evidential resources within a scientific community, etc.

A press report describes a meteorite found in Antarctica which when heated gives off a mix of gases unique to the Martian atmosphere—it was part of the crust of Mars about four billion years ago. Lasers and a mass spectrometer reveal that it contains polycyclic aromatic hydrocarbons (PAHs); this residue closely resembles what you have when simple organic matter decays, and might be fossilized bacteria droppings. David MacKay of the Johnson Space Center argues: "We have these lines of evidence. None of them by itself is definitive, but taken together, the simplest explanation is early Martian life" (Rogers 1996: 56–57). Other scientists, however, suggest that the PAHs might have been formed at volcanic vents; others agree that they are bacterial traces, but believe they were picked up while the meteorite was in Antarctica; and some think the supposed bacterial traces might be nothing more than artifacts of the instrumentation (Begley and Rogers 1997).

This illustrates both the continuity of scientific evidence with everyday empirical evidence, and the complexities that can make it so strong—or so fragile. All of us, in the most ordinary of everyday inquiry, depend on learned perceptual skills like reading, and many of us rely on glasses, contact lenses, hearing aids; in the sciences, observation is often highly skilled, and usually mediated by sophisticated instruments themselves dependent on theory. All of us, in the most ordinary of everyday inquiry, sometimes depend on what others tell us; a scientist virtually always relies on results achieved by others, from the sedimented work of earlier generations to the latest efforts of his contemporaries—though there is virtually always some disagreement within the relevant scientific community about which results are to be relied on, and which shaky. A firmly anchored and tightly woven mesh of evidence can be a strong indication of the truth of a claim—that is partly why "scientific evidence" has acquired its honorific use; but where anchoring is iffy, where some of the

threads are fragile, where different threads pull in different directions, there will be ambiguity, the potential to mislead.

The structure of evidence, to use an analogy I have long relied on, is more like a crossword puzzle than a mathematical proof.⁷ Einstein, I recently learned, once described a scientist as like a man “engaged in solving a well-designed word puzzle.”⁸ I will add that scientific inquiry is a deeply and unavoidably social enterprise (otherwise, each scientist would have to start the work alone and from scratch); so that scientists, in the plural, are like a bunch of people working, sometimes in cooperation with each other, sometimes in competition, on this or that part of a vast crossword—a vast crossword in which some entries were completed long ago by scientists long dead, some only last week; some are in almost-indelible ink, some in regular ink, some in pencil, some heavily, some faintly; and some are loudly contested, with rival teams offering rival solutions.

The degree to which a scientific claim or theory is warranted, at a time, for a person or group of people, depends on how good that person’s or that group’s evidence is, at that time and with respect to that claim or theory. When there is relevant disagreement within the group—as with several people working on the same crossword and disagreeing over certain entries—the group’s evidence should be construed as including the reasons on which the group is agreed, and the disjunctions of those about which there is dispute. Talk of the degree of warrant of a claim or theory at a time, simpliciter, can be construed as shorthand for the degree of warrant of the claim for the person or group of people whose evidence, at that time, is best.

“Person or group” because, while usually the pooled evidence of a group is better than that of its members, sometimes a single person has learned something which has not yet been shared with other members of the relevant community: the results of his experiment have not yet been published, or have been published in a journal too obscure to reach others in the field, or, etc.

Though the warrant of a claim at a time depends on the quality of the evidence possessed by some person or persons at that time, the quality of evidence, its strength or weakness, is not subjective or community-relative. How reasonable a crossword entry is depends on how well it is supported by the clue and any already completed entries, how reasonable

7. See for example Haack 1990 and Haack 1993. I am also drawing, in this section, on Haack 1995, Haack 1996, and Haack 1999.

8. Einstein 1936, in Einstein 1954: 295; drawn to my attention in 1996 by John Norton.

those entries are, independent of the entry in question, and how much of the crossword has been completed. Analogously, how warranted an empirical claim is depends on how well it is supported by experiential evidence and background beliefs, how reasonable those background beliefs are, independent of the belief in question, and how much of the relevant evidence the evidence includes.

The meteorite example also illustrates the connection between supportiveness of evidence and explanatoriness. Briefly and very roughly, how well evidence supports a claim depends on how well the claim is explanatorily integrated with the evidence. Explanation requires the classification of things into real kinds; so supportiveness, requiring kind-identifying predicates, is vocabulary-sensitive. That is why, though there is supportive-but-not-conclusive evidence, there is no syntactically characterizable inductive logic. Most importantly for our purposes, it is also why scientists so often need to introduce new terms, or to adapt the meaning of old terms, as they try to match their language to the real kinds of thing or stuff. (Friedrich Miescher first found a non-proteinaceous substance in the nucleus of cells and dubbed it *nuclein* in 1856;⁹ now molecular biology has refined its classifications over and over: DNA, with its A, B, and Z forms; messenger RNA, transfer RNA, etc.)

Truth-indicative is what evidence has to be to be good; the better-warranted a claim is, the likelier that it is true.¹⁰ At any time, some scientific claims and theories are well warranted; others are warranted poorly, if at all; and many lie somewhere in between. When no one has good enough evidence either way, a claim and its negation may be both unwarranted (so degrees of warrant don't work just like mathematical probabilities). Most scientific claims and theories start out as informed but speculative conjectures; some seem for a while to be close to certain, and then turn out to have been wrong after all; a few seem for a while to be out of the running, and then turn out to have been right after all. But, as scientific inquiry has proceeded, a vast sediment of well-warranted claims has accumulated.

Ideally, the degree of credence given a claim by the relevant scientific sub-community would be appropriately correlated with the degree of warrant of the claim. The processes by which a scientific community col-

9. For the relevant history (up to the date of its publication, naturally) see Portugal and Cohen 1977.

10. Readers who have reservations about the concept of truth are referred to Haack 1998b: 7–30 and Haack 1999.

lects, sifts, and weighs evidence are fallible and imperfect, so the ideal is not always achieved; but they are good enough that it is a reasonable bet that much of the science in the textbooks is right, while only a fraction of today's speculative frontier science will survive, and most will eventually turn out to have been mistaken.¹¹ *Only* a reasonable bet, however; all the stuff in the textbooks was once speculative frontier science, and textbook science can occasionally be embarrassingly wrong (e.g., the arbitrary tautomeric forms in the chemistry texts on which, before Jerry Donohue set him straight, James Watson relied).¹²

The quality of evidence is objective, depending on how supportive it is, how comprehensive, and how independently secure the reasons it includes; but judgments of the quality of evidence are perspectival, i.e., they depend on the background beliefs of the person making the judgment. If you and I are working on the same crossword, but have filled in the much-intersected 4 down differently, we will disagree about whether the fact that an entry to 12 across ends in an "F," or the fact that it ends in a "T," makes it reasonable. Similarly, if you and I are on the same hiring committee, and you believe that handwriting is an indication of character, while I think that's all nonsense, we will disagree about whether the fact that a candidate loops his *fs* is relevant to whether he should be hired. Whether it *is* relevant, however, depends on whether it is *true* that handwriting is an indication of character.

If, as I have maintained, the standards of strong evidence and well-conducted inquiry that apply to the sciences are the very same standards that apply to empirical inquiry generally, doesn't it follow that a lay person should be able to judge the worth of scientific evidence as well as a scientist? Unfortunately, no—far from it; for every area of science has its own specialized vocabulary, dense with theory, and judgments of the worth of evidence depend on substantive assumptions. Very often, the only alternative to relying on the judgment of scientists competent in the relevant field is to acquire a competence in that field yourself.

When a lay person (or even a scientist from another specialty) tries to judge the quality of evidence for a scientific claim, he is liable to find himself in the position of the average American asked to judge the reasonableness of entries in a crossword puzzle where, though some of the clues are in pidgin English, the solutions are all in Turkish and presuppose a knowledge of the history of Istanbul, or are all in Bengali and require a

11. Bauer 1993 chapter 3 is good on this.

12. Watson 1968 chapter 26.

knowledge of Islam, or, etc.¹³ Similarly, to know what kinds of precaution would be adequate to ensure against experimental error requires substantive knowledge of what kinds of thing might interfere. To judge the likelihood that you are not dealing with a real phenomenon but with an artifact of the instrumentation requires substantive knowledge of how the instrument works. And so on.

Still, can't we at least assume that competent scientists in the relevant field will agree whether this is strong or flimsy evidence, whether that experiment is well- or ill-designed, etc.? Unfortunately, no—not always. At the textbook-science end of the continuum, where claims and theories are very well-warranted, competent scientists will agree. But the closer scientific work is to the frontier, the less comprehensive the evidence so far available, the more room there is for legitimate disagreement about what background information is reliable, hence about what evidence is relevant to what, and hence about the warrant of a claim. Even the most competent scientists may be in something like the position of people working on a part of a crossword in which, so far, only a few entries have been completed, leaving open more than one reasonable alternative solution to others. As Crick and Watson began work on the structure of DNA, some scientists in the field still believed that protein was the genetic material. As the work proceeded, Crick and Watson were sure DNA was helical; Franklin remained for a good while unconvinced. Crick and Watson thought the backbone was on the inside of the molecule; Franklin suspected it was on the outside. As soon as he learned of Chargaff's discovery of approximate equalities in the purine and pyrimidine residues in DNA, Watson was convinced of its importance; Crick still had to be persuaded.¹⁴

For most of what follows, the epistemological points that will most concern me are negative, identifying deadwood in need of pruning, misunderstandings about science and how it works which have hampered legal efforts to distinguish decent science from junk: In the descriptive sense of "science," there is bad science as well as good. There is no peculiar method which distinguishes genuine science from impostors. Usually there is no way of judging the worth of scientific evidence without substantive knowledge of the appropriate field. There is no guarantee that specialists in a scientific field won't sometimes legitimately disagree. And

13. This puts me in mind of geneticist S. C. Harland's comment on trying to talk about biology with Trofim Lysenko: "it was like discussing the differential calculus with a man who did not know his 12-times table" (Gardner 1852: 147, referring to Huxley 1949).

14. Again, I rely on Watson 1968.

there is no guarantee, either, that at any given time and for any legitimate scientific question, a warranted answer will be available.

■ ■ ■

Once upon a time, in cases where expert knowledge was required, jurors with the necessary expertise were specially selected—e.g., a jury of butchers when the accused was charged with selling putrid meat; and sometimes specially qualified persons would be summoned to help determine some matter of fact which the court had to decide—e.g., masters of grammar for help in construing doubtful words in a bond. Learned Hand reports that the first case he can find of “real expert testimony”—expert testimony as exception to the rule that the conclusions of a witness are inadmissible—was in 1620.¹⁵ But now, of course, when specialized knowledge is needed, the usual method is calling expert witnesses.

Though it was not cited in a federal or state ruling for a decade, the *Frye* case (1923) gradually began to set the standard of admissibility of scientific evidence, at first mainly in criminal cases but later in civil cases too. Mr. Frye was charged with murder, and had confessed. Later, however, he repudiated the confession; and took, and passed, a polygraph test (or more exactly, a discontinuous test of systolic blood pressure changes under questioning; the technology was in an early and primitive stage).¹⁶ But the trial court judge excluded this evidence, taking the view that deception tests were inadmissible unless there is “an infallible instrument for ascertaining whether a person is speaking the truth or not.”¹⁷ On appeal, the D.C. Court confirmed the exclusion of this lie-detector evidence, ruling that novel scientific evidence “crosses the line between the experimental and the demonstrable,” and so is admissible, only if it is “sufficiently established to have gained general acceptance in the particular field to which it belongs.”¹⁸ This is the “*Frye* rule” or “*Frye* test.”

15. Hand 1901: 40–49; the date (1620) is given on p. 45.

16. “The defendant in *Frye* was subsequently pardoned when someone else confessed to the crime,” writes Paul Giannelli 1980: n. 42. Giannelli cites Wicker 1953; Wicker, he says, cites Fourteenth Annual Report of Judicial Council of the State of New York, 265 (1948). But according to the most complete account I have been able to find of the many twists and turns of Mr. Frye’s story—Starrs 1982—none of this is true.

17. My source in Starrs 1982: 694; he refers to Transcript on Appeal, File 3968, retired files, National Records Center, Suitland, MD.

18. From Judge Van Ordsel’s opinion for the appellate court in *Frye*. At the time, the D.C. Court offered little in the way of rationale for its ruling. Much later, however, when the influence of *Frye* was waning, the same court argued that “[T]he requirement of general acceptance in the scientific community assures that those most qualified to assess the validity of a scientific method will have the determinative voice” (United States v. Addison; my source is Giannelli 1980: 1207).

228 Journal of Health Politics, Policy and Law

As the *Frye* rule was applied and contested in the courts, the effect was sometimes more and sometimes less restrictive. Voice-print evidence, for example, was sometimes admitted under the *Frye* test, sometimes excluded.¹⁹ In *People v. Williams* (1958), the prosecution's own experts conceded that the medical profession was mostly unfamiliar with the use of Nalline to detect narcotic use, but the court upheld the admissibility of its evidence all the same; the Nalline test was "generally accepted by those who would be expected to be familiar with its use," and "in this age of specialization more should not be required."²⁰ In *Copolino v. State* (1968), the prosecution was allowed to introduce the results of a test (for the presence of succinylcholine chloride or its derivatives in human tissues) devised by the local medical examiner specifically for this trial—and so not known to, let alone generally accepted in, any scientific community. The appellate court cited *Frye* but, ruling that the trial judge did not abuse his discretion, nevertheless upheld the admissibility of this evidence (Giannelli 1980: 1222 ff.).

The epistemological assumptions behind the *Frye* test are quite crude; and, while it seems overly restrictive in principle, it is indeterminate in ways that made it nearly inevitable that in practice its application would be, not merely variable in borderline cases, but systematically inconsistent.

Rather than requiring the trial judge to determine in his own behalf whether scientific evidence proffered is solidly established work or unreliable speculation, the *Frye* test had him rely obliquely on the verdict of the appropriate scientific subcommunity. Three assumptions seem to lie behind the test: that there is a definite point at which scientific claims or techniques cease to be "experimental" and become "demonstrable"; that a claim or technique has not achieved this "demonstrable" status unless it is generally accepted in the relevant community; and that only "demonstrable" claims and techniques should be admitted.

The first two assumptions are at best oversimplifications. Rather than a sharp line, there is really a continuum from the unwarranted through

19. See Black, Ayala, and Saffran-Brinks 1994: 735 ff., listing *Reed v. State and United States v. Addison*, excluding voiceprint evidence under the *Frye* test; and *Commonwealth v. Lykus*, admitting voiceprint evidence under the *Frye* test. There is a useful summary of relevant cases in the Symposium on Science and Rules of Evidence, 99 *Federal Rules Decisions* 188 (1983).

20. Giannelli comments: "if the 'specialized field' is too narrow . . . the judgment of the scientific community becomes, in reality, the opinion of a few experts" (1980: 1209–1210).

the poorly-warranted to the well-warranted; and the degree of credence given a claim in the relevant scientific community is only an imperfect indicator of its degree of warrant (which is only an imperfect indicator—albeit the best we can have—of its truth). Sometimes—perhaps in the case of the medical examiner in *Coppolino*—one person has better evidence than the community. General acceptance in the relevant community is only a very rough-and-ready, and a quite conservative, guide to what is well-warranted at the time in question.

The third assumption—that only “demonstrable” scientific evidence should be admitted—seems extremely restrictive. Precluding the possibility that there should be scientific witnesses who disagree but both of whose testimony is admissible, it seems to confine the courts, in effect, to textbook science. A physicist colleague tells me he once testified that the hypothesis was consistent with the laws of mechanics that the deceased wasn’t pushed, but fell; but very often, surely, the relevant science will be quite far from the textbook stage.

However, it takes only a moment’s reflection to realize that how restrictive the *Frye* test would be in practice depends on what exactly was required to be accepted by what proportion of what community. The narrower and more homogeneous the relevant community is taken to be, the likelier it is that there will be agreement; the broader and more heterogeneous the community, the likelier that there will be disagreement. (Unlike the Verification Principle, which is broader if “verifiable” is construed broadly and narrower if “verifiable” is construed narrowly, the *Frye* test is broader if the community is defined narrowly and narrower if the community is defined broadly.) No wonder, then, that, though often criticized as overly restrictive, in practice the test was far from consistent.

■ ■ ■

The Federal Rules of Evidence (1975) encapsulate a (less ostensibly restrictive) relevancy approach. Rule 104 (a) affirms the gatekeeping role of the court in ruling on admissibility of evidence. But Rule 401 states that relevant evidence—evidence which has any tendency to make the existence of any fact of consequence to the determination of the action either more or less probable than it would otherwise be—is admissible unless otherwise provided by law. Rule 702 states that expert evidence, including but not restricted to scientific evidence, is admissible subject to exclusion under Rule 403. Rule 403, specifying the grounds for exclusion, mentions the danger of unfair prejudice, confusion of the issues, or misleading the jury, but does *not* mention any requirement of general

acceptance in the appropriate scientific community. Rule 706 allows the court to appoint expert witnesses of its own selection.

The *Frye* rule didn't wither away immediately. Scholars debated whether the Federal Rules were compatible with the *Frye* test: some arguing that they weren't, because they didn't mention consensus in the relevant community; and some arguing that they were, because they didn't mention consensus in the relevant community (!).²¹ The 1987 edition of a textbook on the Federal Rules suggests ironically that the *Frye* test be reconstrued under Rule 403 as "an attempt to prevent jurors from being unduly swayed by unreliable scientific evidence" (Graham 1987: 92).

Most to the point of the present narrative, in *Daubert* (1993) the trial court relied almost exclusively on *Frye* in ruling the plaintiff's expert evidence inadmissible. The plaintiffs were two minor children and their parents, and the claim was that the children's birth defects were caused by their mothers' having taken the morning-sickness drug Bendectin during pregnancy. But the plaintiffs' expert evidence (based on animal studies, pharmacological studies of the chemical structure of Bendectin, and an unpublished "re-analysis" of previously published human statistical studies) was disqualified under the *Frye* test. The Ninth Circuit confirmed the trial court's decision to exclude.

But in 1993, reversing the exclusion of Daubert's expert testimony, the majority of the U.S. Supreme Court repudiated the *Frye* test as an "austere standard, absent from, and incompatible with, the [Federal Rules]. . . . [U]nder the Rules the trial judge must ensure that any and all scientific testimony or evidence admitted is not only relevant, but reliable."²² Jurors, whose job it is to determine sufficiency, are to concern themselves with expert witnesses' conclusions; but judges, whose job it is to determine admissibility, must focus "solely on principles and methodology" to make "a preliminary assessment of whether the reasoning or methodology underlying the testimony is scientifically valid and . . . properly can be applied to the facts in issue."²³

In determining whether what is offered is really scientific knowledge—*knowledge*, not mere opinion, and genuinely *scientific* knowledge, "with a grounding in the methods and procedures of science"—a key

21. I rely on Giannelli 1980: 1229–1230. He mentions Saltzburg and Redden 1977: 426 as holding that the Federal Rules are compatible with the *Frye* test because they don't mention general acceptance; and Wright and Graham 1978: 92, as holding that the Federal Rules are incompatible with the *Frye* test because they don't mention general acceptance.

22. *Daubert*, 509 U.S. at 598, 113 S. Ct. at 2794.

23. *Daubert*, 509 U.S. at 580, 113 S. Ct. at 2790.

question will be “whether it can be (and has been) tested.”²⁴ Justice Blackmun’s opinion for the majority quotes Green: “‘Scientific methodology today is based on generating hypotheses and testing them to see if they can be falsified; indeed, this methodology is what distinguishes science from other fields of human inquiry’,”²⁵ and refers to Popper and Hempel. Retaining something of the *Frye* test in the liberalized form of indications, rather than necessary conditions, of admissibility, the *Daubert* ruling also mentions peer-review, a “known or potential error rate,” and “widespread acceptance.”

However, dissenting in part from the majority, after pointing out that there is no reference in Rule 702 to reliability, and urging that the question of expert testimony generally not be confused with the question of scientific testimony specifically, Justice Rehnquist remarks:

I defer to no one in my confidence in federal judges; but I am at a loss to know what is meant when it is said that the scientific status of a theory depends on its ‘falsifiability,’ and I suspect some of them will be, too. . . . I do not think [Rule 702] imposes on them either the obligation or the authority to become amateur scientists.²⁶

Those reservations are well-founded; for the epistemological assumptions on which the *Daubert* ruling rests are badly confused.

Unlike the *Frye* test, the Federal Rules as interpreted in *Daubert* require the trial judge to make determinations about scientific methodology in his own behalf. But what the *Daubert* Court has to offer by way of advice about how to make such determinations is—well, a little embarrassing.

The justices are apparently unaware that Popper gives “falsifiable” a very narrow sense, “incompatible with some basic statement” (a basic statement being defined as a singular statement reporting the occurrence of an observable event at a specified place and time); and that according to Popper no scientific claim or theory can ever be shown to be true or even probable, but is at best “corroborated.” In Popper’s mouth, this is not equivalent to “confirmed,” and does not imply truth or probable truth, but means no more than “tested but not yet falsified.”²⁷ If Popper were

24. *Daubert*, 509 U.S. at 593, 113 S. Ct. at 2796.

25. Green 1992: 645. A footnote (12) refers to Popper, but I can find no reference to Hempel.

26. *Daubert*, 509 U.S. at 600–601, 113 S. Ct. at 2800.

27. In ordinary speech, of course, *corroborated* usually means “confirmed by another witness,” but Popper has given the word a quite different, technical meaning. Black, Ayala, and Saffran-Brinks 1994: 750 ff. seem to have confused *corroboration*, in Popper’s sense, with

right, no scientific claim would be well-warranted. In fact, it is hard to think of a philosophy of science less congenial than Popper's to the relevance-and-reliability approach (or to the admissibility of psychiatric evidence, but that is a whole other can of worms). And if the reference to Popper is a *faux pas*, running Popper together with Hempel—a pioneer of the logic of confirmation, an enterprise the legitimacy of which Popper always staunchly denied—is a *faux pas de deux*.

In and of itself, of course, the *Daubert* Court's mixing up its Hoppers and its Pempels is just a minor scholarly irritation. A more serious problem is that neither Popper's nor Hempel's philosophy of science will do the job they want it to do. Popper's account of science is in truth a disguised form of skepticism; if it were right, what Popper likes to call "objective scientific knowledge" would be nothing more than conjectures which have not yet been falsified. And, though Hempel's account at least allows that scientific claims can be confirmed as well as disconfirmed, it contains nothing that would help a judge decide either whether evidence proffered is really scientific, or how reliable it is.

And the most fundamental problem is that the *Daubert* Court (doubtless encouraged by the dual descriptive and honorific uses of "scientific") is preoccupied with specifying what the method of inquiry is that distinguishes the scientific and reliable from the nonscientific and unreliable. There is no such method. There is only making informed conjectures and checking how well they stand up to evidence, which is common to every kind of empirical inquiry; and the many and various techniques used by scientists in this or that scientific field, which are neither universal across the sciences nor constitutive of real science.

The *Daubert* Court runs together (1) the tangled and distracting questions of demarcation and scientific method with (2) the question of the degree of warrant of specific scientific claims or theories and (3) the question of the reliability of specific scientific techniques or tests—which is different again, for the claim that this technique is unreliable may be well warranted, the claim that this other technique is reliable poorly warranted. Unlike determining whether a claim is falsifiable, however, determining whether a scientific theory (e.g., of the etiology of this kind of

confirmation. Green—who, incidentally, introduces Popper's philosophy of science in Kuhnian terms, as "the existing paradigm under which scientists work"—acknowledges that Popper holds that "[t]heoretically . . . hypotheses are never affirmatively proved" but continues, "of course, if a hypothesis repeatedly withstands falsification, one may tend to accept it, even if conditionally, as true" (1992: 645–646).

cancer) is well warranted, or whether a scientific test (e.g., for the presence of succinylcholine chloride) is reliable, requires substantive scientific knowledge. Justice Rehnquist is right: the reference to falsifiability is no help, and judges are indeed being asked to be amateur scientists.

Furthermore, despite the majority's reassuring noises to the effect that juries can handle scientific evidence well enough, and can always be directed by the judge if they look like going off the rails, one is left wondering: if judges need to act as gatekeepers to exclude scientific evidence which doesn't meet minimal standards of warrant because juries may be taken in by flimsy scientific evidence, how realistic is it to expect juries to discriminate the better from the worse among the half-way decent?

■ ■ ■

One of the many subsequent cases²⁸ in which the Federal Rules as interpreted in *Daubert* are applied to the question of the admissibility of scientific evidence is the one that first drew my attention—the case of Mr. Joiner.

Robert Joiner had worked for the Water and Light Department of the City of Thomasville, Georgia, since 1973. Among his tasks was the disassembly and repair of electrical transformers in which a mineral-based dielectric fluid was used as a coolant—dielectric fluid into which he had to stick his hands and arms, and which sometimes splashed onto him, occasionally getting into his eyes and mouth. In 1983 the city discovered that the fluid in some of the transformers was contaminated with PCBs, which are considered so hazardous that their production and sale has been banned by Congress since 1978.

In 1991 Mr. Joiner was diagnosed with small-cell lung cancer; he was thirty-seven. He had been a smoker for about eight years, and there was a history of lung cancer in his family. He claimed, however, that had it not been for his exposure to PCBs and their derivatives, furans and dioxins, his cancer would not have developed for many years, if at all. On this basis he sued Monsanto, which had manufactured PCBs from 1935 to 1977, and General Electric and Westinghouse, which manufactured transformers and dielectric fluid. His case relied essentially on expert witnesses who testified that PCBs alone can cause cancer, as can furans and dioxins, and that since he had been exposed to PCBs, furans, and dioxins, this exposure had likely contributed to his cancer.

Removing the case to federal court, GE et al. contended that there was

28. By January of 1998, according to Frankel 1998: 3, there had been more than eleven hundred such cases.

234 Journal of Health Politics, Policy and Law

no evidence that Mr. Joiner suffered significant exposure to PCBs, furans, or dioxins, and that in any case there was no admissible scientific evidence that PCBs promoted Joiner's cancer. The district court granted summary judgment, holding that the testimony of Joiner's experts was no more than "subjective belief or unsupported speculation."²⁹

The court of appeals reversed. Federal Rule 702, governing expert testimony, displays a "preference for admissibility," and in the present instance, the question of admissibility was "outcome-determinative": if the scientific evidence offered were excluded, Mr. Joiner would simply have no case. So a "particularly stringent standard of review" should apply to the trial judge's exclusion of expert testimony.³⁰

But in 1997, reversing the admissibility of Mr. Joiner's expert evidence, the U.S. Supreme Court held that the appeals court erred in applying an especially stringent standard of review. The appropriate standard was abuse of discretion; and it was *not* an abuse of discretion for the district court to have excluded Mr. Joiner's experts' testimony.³¹

And now it begins to appear how the question of the legitimacy of the distinction between methodology and conclusions came to be a hotly contested issue. The *Daubert* Court, taking the distinction for granted, had interpreted the gatekeeping role of trial judges as requiring them to focus solely on methodology, not conclusions. But, Mr. Joiner's lawyers argue, the District Court had no objection to the methodology of the studies cited, only to the conclusions that their experts drew; and this was a reversible error.

GE's brief argues that the court of appeals treated *Daubert's* requirement of scientific methodology "at such a superficial level as to leave it meaningless—calling for no more than the invocation of scientific materials."³² Mr. Joiner's experts rely on the "faggot fallacy": the fallacy of supposing that "multiple pieces of evidence, each independently being suspect or weak, provide strong evidence when bundled together."³³ Mr.

29. General Electric Co. v. Joiner, 522 U.S. 136, 136, 118 S. Ct. 512, 514 (1997), citing Joiner v. General Electric Co., 864 F. Supp. 1310, 1326 (N.D. Ga. 1994), which in turn cites *Daubert*, 509 U.S. 579, 113 S. Ct. 2786 (1993), where the phrase occurs three times: at 597 and 2786; at 590 and 2795; and at 599 and 2800.

30. See General Electric Co. v. Joiner, 522 U.S. at 140, 118 S. Ct. at 516.

31. But the question with regard to furans and dioxins, according to the Supreme Court ruling, remained open.

32. Brief for Petitioners, General Electric Co. v. Joiner, 47.

33. Brief for Petitioners, General Electric Co. v. Joiner, 49, citing Skrabanek and McCormick 1997: 35, quoted in Huber and Foster 1997: 142. I notice that on the same page, Skrabanek and McCormick refer to what they call the "weight of evidence fallacy"; this, they claim, is not scientific because science, according to Popper, focuses on negative evidence (which cannot be outweighed by confirming instances). While I am noting that GE's lawyers cite Peter Huber, I will also note that Kenneth Cheseboro was one of Mr. Joiner's lawyers.

Joiner's lawyers reply that his experts "were applying a methodology which is well established in the scientific method. It is known as the weight of evidence methodology. . . . There are well-established protocols for this . . . published as the EPA's guidelines. There are similar guidelines for the World Health Organization."³⁴ GE's lawyers never challenged Mr. Joiner's experts' methodology before; indeed, they use the "weight of evidence" methodology themselves.

Rather than challenging Mr. Joiner's claim that the District Court failed to restrict its attention to methodology as *Daubert* requires, the majority of the *Joiner* Court sustains its ruling that there was no abuse of discretion by holding that "*conclusions and methodology are not entirely distinct from each other.*"³⁵

Justice Stevens, however (concurring on the question of the correct standard of review but dissenting from the majority's ruling on whether the district court erred) protests that this is neither true nor helpful. "The difference between methodology and conclusions is just as categorical as the distinction between means and ends." The district court ruling on reliability in *Joiner*, in particular, is "arguably not faithful" to the statement in *Daubert* that the focus must be on methodology rather than conclusions. The majority "has not adequately explained why its holding is consistent with Federal Rule of Evidence 702 as interpreted in *Daubert v. Merrell Dow Pharmaceuticals.*"³⁶

In the *Joiner* ruling, *Daubert's* epistemological chickens come home to roost: with the references to falsifiability gone and the distinction between methodology and conclusions dropped, it is starkly obvious that judges will sometimes be obliged to determine substantive scientific questions.

Given the difficulties with the *Daubert* Court's efforts to specify what makes evidence genuinely scientific, perhaps the knots in which everyone ties themselves in *Joiner* (not to mention the absence from the ruling of any reference whatever to falsifiability, testability, Hepper, Pompel, etc.)³⁷ are not so surprising. What *is* surprising, to me at any rate, is that the *Joiner* Court should offer, as an interpretation of *Daubert*, a ruling

34. Oral Argument of Michael H. Gottesman, *General Electric Co. v. Joiner*, 43–44. Mr. Gottesman was also one of the attorneys for Mr. *Daubert*.

35. *General Electric Co. v. Joiner*, 522 U.S. at 155, 118 S. Ct. at 523 (emphasis mine).

36. *General Electric Co. v. Joiner*, 522 U.S. at 151, 118 S. Ct. at 521 (Stevens, J., dissenting).

37. And of any reassuring noises about jurors' ability to assess the weight of scientific evidence.

that denies the legitimacy of a distinction *Daubert* presupposed. I have no difficulty with the idea that a later ruling may make an earlier ruling determinate in respects in which it was formerly indeterminate (which, incidentally, explains why the *Daubert* Court could rule that the *Frye* test is incompatible with the Federal Rules, which at first raised my logical eyebrows quite far). But the idea that a later ruling which flatly denies a clear presupposition of an earlier ruling could qualify as an interpretation, rather than a revision, of it, still strikes me as very strange indeed.

However. What about the distinction between methodology and conclusions presupposed in *Daubert*, but repudiated in *Joiner*? In these cases the concept of **methodology** (never exactly well-defined in the philosophy of science) seems to have turned into an accordion concept,³⁸ expanded and contracted as the argument requires. Is the judge, in determining the validity of experts' "methodology," to decide whether the mouse studies on which Mr. Joiner's experts in part relied were well-conducted, with proper controls and good records, using specially bred genetically-uniform mice, etc., etc.; or what weight to give mouse studies with respect to questions about humans; or what weight to give those mouse studies in the context of other studies of the effects on humans of PCB and other contaminants; or what? There are so many ambiguities that everyone is right—and everyone is wrong.

Mr. Joiner's lawyers are right to suggest that drawing the reasonable conclusion from a conglomeration of disparate bits of information (mouse studies, epidemiological evidence, etc.) requires, well, weighing the evidence. But of course it matters whether you weigh the evidence *properly*; and GE's lawyers are right, too, when they complain that Mr. Joiner's attorneys use "methodology" so loosely as to make *Daubert*'s requirements practically vacuous.

But GE's accusation that Mr. Joiner's experts commit the "faggot fallacy" relies on an equivocation. There is an ambiguity in the reference to "pieces of evidence, each independently . . . suspect or weak": this may mean *either* "pieces of evidence each themselves poorly warranted" (which seems to be the interpretation intended by Skrabanek and McCormick, to whom the phrase "faggot fallacy" is due), *or* "pieces of evidence each by itself inadequate to warrant the claim in question" (which seems to be the interpretation most relevant to the case). True, if the reasons for a claim are themselves poorly warranted, this lowers the degree of warrant of the claim itself. But GE's brief offers no argument

38. The term, and the idea, come from Sellars 1965: 172.

that the reasons based on the studies to which Mr. Joiner's experts refer are themselves poorly warranted. True again, none of those reasons by itself strongly warrants the claim that PCBs promoted Mr. Joiner's cancer. But GE's brief offers no argument that they don't do so jointly.

Sometimes bits of evidence which are individually weak are jointly strong; sometimes not—it depends what they are, and whether or not they reinforce each other (whether or not the crossword entries interlock). Chargaff's discovery that there are approximate regularities in the relative proportions of adenine and thymine, guanine and cytosine in DNA is hardly, by itself, strong evidence that DNA is a double-helical, backbone-out macromolecule with like-with-unlike base pairs; Franklin's X-ray photographs of the B form of DNA are hardly, by themselves, strong evidence that DNA is a double-helical, backbone-out macromolecule with like-with-unlike base pairs. That the tetranucleotide hypothesis is false is hardly, by itself, strong evidence that DNA is a double-helical, backbone-out macromolecule with like-with-unlike base pairs; and so on. But put all these pieces of evidence together, and the double-helical, backbone-out, like-with-unlike base pairs, structure of DNA is very well-warranted indeed (in fact, the only entry that fits).

Neither party seriously addresses this question of interlocking. But in the very complex EPA guidelines to which Mr. Joiner's attorneys so causally refer, I find this: "Weight of evidence conclusions come from the *combined strength and coherence* of inferences appropriately drawn from all of the available evidence."³⁹

Justice Stevens is right to say that there is a difference between methodology and conclusions, as there is between ends and means; there is a difference, certainly, between a technique and its result, or between premises and conclusion. But on a more charitable interpretation, the majority's point is not that there is literally no distinction, but that it is impossible to judge methodology without relying on some substantive scientific conclusions. And this is both true and important.

To determine whether this evidence (e.g., of the results of mouse studies) is relevant to that claim (e.g., about the causes of Mr. Joiner's cancer) requires substantive knowledge (e.g., about the respects in which mouse physiology is like human physiology, about how similar or how different the etiologies are of small-cell lung cancer and alveogenic adenomas, etc.). And to determine the reliability of a scientific experiment, technique, or test, it is necessary to know what kinds of thing might

39. 61 *Federal Register* 19,760–1: 17972 (1996) (my italics).

interfere with the proper working of this apparatus, what the chemical theory is that underpins this analytical technique, what factors might lead to error in this kind of experiment and what precautions are called for, or to possess a sophisticated understanding of statistical techniques or of complex and controversial methods of meta-analysis pooling data from different studies. And so on.

Which takes us back to that old worry of Justice Rehnquist's of which Justice Breyer's observation that judges are not scientists reminds us: judges are neither trained or qualified to do this kind of thing.

■ ■ ■

Already at the time of *Joiner*, the *Daubert* ruling, requiring judges to make a preliminary evaluation of scientific evidence proffered, had prompted wider use of Rule 706, allowing judges to appoint their own experts.

In 1992, the FDA had banned silicone breast implants, formerly "grandfathered in." They were not known to be unsafe; but manufacturers had not, as required under FDA regulations, supplied evidence of their safety. Understandably, the ban caused a good deal of anxiety, and provoked a wave of fear, greed, and litigation. In 1996, Judge Sam Pointer of the U.S. District Court in Birmingham, Alabama, who had been in charge of all several thousand federal implant cases for more than six years, convened a panel of four scientists—an immunologist, an epidemiologist, a toxicologist, and a rheumatologist—to review evidence of the alleged connections between silicone implants and various systemic and connective tissue diseases.

Judge Pointer's carefully phrased remit asks: "to what extent, if any and with what limitations and caveats do existing studies, research, and reported observations provide a reliable and reasonable scientific basis for one to conclude that silicone-gel breast implants cause or exacerbate any . . . 'classic' connective tissue diseases [. . . or] 'atypical' presentations of connective tissue diseases. . . . To what extent, if any, should any of your opinions . . . be considered as subject to sufficient dispute as would permit other persons, generally qualified in your field of expertise, to express opinions that, though contrary to yours, would likely be viewed by others in the field as representing legitimate disagreement within your profession?"⁴⁰

40. Submission of Rule 706 National Science Panel Report: 2, In re: Silicone Gel Breast Implant Products Liability Litigation (N.D. Ala. 1998) (No. CV 92-P-10000-S), available online at fjc/BREIMLIT/SCIENCE/report.htm.

Two years and (only) \$800,000 later,⁴¹ after selecting from more than two thousand published and unpublished studies those they thought most “rigorous and relevant,” in December 1998 the panel submitted a long report. Their conclusion was that the evidence studied and reanalyzed (apparently the forty or so studies submitted by each side plus about one hundred others, including unpublished studies, Ph.D. dissertations, and letters) does not warrant the claim that silicone breast implants cause these diseases. They add, however, that in some respects “the number and size of studies is inadequate to produce definite results”; that animal testing “may not fully predict the human effects”; that some evidence suggests that silicone implants are not entirely benign (they can cause inflammation, and droplets can turn up in distant tissues); and that while most people in the field would agree with their conclusions, a few might not.⁴²

Despite Judge Pointer’s efforts to ensure that his experts were unimpeachably neutral, the plaintiffs’ lawyers objected that the rheumatologist on Pointer’s panel had undisclosed connections with one of the defendants, Bristol-Meyers Squibb (BMS), while a member of the panel: in August 1997, apparently, he signed a letter soliciting up to \$10,000 in support of a rheumatology meeting he co-chaired, stating that “the impact of sponsorship will be high, as the individuals invited for this workshop, being opinion leaders in their field, are influential with the regulatory agencies”; in October 1998 he signed a \$1,500-a-day fee arrangement with BMS, and in November 1998 he received \$750 for participating in a company seminar.⁴³

In April 1999, averring that there was no actual bias, though acknowledging that there might be a regrettable appearance of bias, Judge Pointer ruled against the plaintiffs’ motion that the panel’s report be excluded. The members of the panel will give videotaped sworn statements that may be used as evidence in courts nationwide.

The bramble-bush, of course, is alive and well, growing new fruit, and new thorns, almost every day.⁴⁴ In *Kumho* (1999), considering judges’ responsibility for making a preliminary reliability assessment of the tes-

41. “Only” not only because the sum is trivial relative to the compensation awarded in some implant cases, but also because the amount is quite modest relative to the task undertaken.

42. Submission of Rule 706 National Science Panel Report: 8.

43. Pointer Rules Federal Science Panel Report Not Tainted by Payments to Panelist, 7.5 *Legal Aspects of Breast Implants*: 4 (April 1999). I conjecture that the discrepancy between reports about the sum of money involved—plaintiffs say \$750, court says \$500—may be a matter of Canadian versus U.S. dollars. Plaintiffs also object that a colleague who had assisted Dr. Tugwell in his work for the panel had received support from a company wholly owned by BMS.

44. It is just this that, as a philosopher, I find most disturbingly unfamiliar when I tackle legal

timony of engineers and other non-scientific experts, the Supreme Court stressed that *Daubert's* test of reliability is "flexible," and that its list of specific factors (falsifiability, peer review, etc.) "neither necessarily nor exclusively applies to all experts or in every case"; thus partially addressing the issues about the place of scientific evidence within expert evidence generally raised by Justice Rehnquist's dissent from the *Daubert* ruling.⁴⁵

There have also been some efforts to educate judges scientifically. In April 1999 about two dozen Massachusetts Superior Court judges attended a two-day seminar on DNA at the Whitehead Institute for Biomedical Research. A report in the *New York Times* quotes the director of the institute: in the O. J. Simpson trial lawyers "befuddle[d] everyone" over the DNA evidence; but after this program, "I don't think a judge will be intimidated by the science." Judges will "understand what is black and white . . . what to allow in the courtroom" (Goldberg 1999: 10).

And in May 1999 the American Association for the Advancement of Science inaugurated a five-year project to make available to judges "independent scientists who would educate the court, testify at trial, assess the litigants' cases, and otherwise aid in the process of determining the truth" (Bandow 1999).

Disentangling "reliable" from "scientific," as *Kumho* begins to do, is certainly all to the good. But a bit of scientific education for judges is at best a drop in the bucket; and court-appointed panels of experts, though potentially helpful, are no panacea.

Not that educating judges about DNA or whatever mightn't do some good. But a few hours in a science seminar will no more transform judges into scientists competent to make subtle and sophisticated scientific determinations than a few hours in a legal seminar would transform scientists into judges competent to make subtle and sophisticated legal determinations. ("This kind of thing takes a lot of training," as Mad Margaret sings in *Ruddigore*.) And, to be candid, that *New York Times* report has me a little worried about the danger of giving judges a false impres-

matters. Perhaps that is why, though Karl Llewellyn can write (1930: 141), "To me there is more joy than pain, by a good deal, in the thorns of such a thicket as that through which I have just dragged you," I am starting to feel as if I have been dragged through a hedge backwards!

45. And *United States v. Starzecpyzel* (880 F. Supp. 1027 [S.D.N.Y. 1995]) raises some interesting epistemological issues about learning and skill in perception, the relation of knowing-that and knowing-how, etc. But I shall have to set these aside.

sion that they *are* qualified to make those “subtle and sophisticated determinations.”

“[N]either the difficulty of the task nor any comparative [*sic*] lack of expertise can excuse the judge from exercising the ‘gatekeeper’ duties that the Federal Rules impose,” Justice Breyer avers.⁴⁶ More directly than the *Frye* test, calling on court-appointed panels of scientists turns part of the task over to those who are more equipped to do it. Isn’t this a whole lot better than asking judges to be amateur scientists? Sometimes, probably, significantly better—the more so, the closer the work at issue is to black-letter science; not, however, as straightforwardly or unproblematically better as some hope.

As Judge Pointer’s panel’s report was made public, an optimistic headline in the *Washington Times*⁴⁷ proclaimed “Benchmark Victory For Sound Science”; and under the headline “An Unnatural Disaster,” an editorial in the *Wall Street Journal* announced that “reason and evidence have finally won out.”⁴⁸ ABCNEWS.com’s “Health and Living” was considerably more cautious: under the headline “No Implant-Disease Link?” a sideline adds, “The panel found no definite links, but it also left the door open for more research.”⁴⁹ Neither quite captures my reaction.

I should be quite surprised if it turned out that silicone implants do, in fact, cause the various diseases they have been alleged to (so far as I can tell it isn’t just, as the panel’s report says, that there is no evidence that they do; but that there is pretty good evidence that they don’t).⁵⁰ And I don’t think it very likely that that \$750 seriously affected Dr. Tugwell’s opinion (though I must say that—even if this kind of thing is routine in funding applications, as for all I know it may be—that letter boasting of the applicants’ influence with regulatory bodies leaves a bad taste in my mouth).

I don’t feel equally confident, however, that a really good way has yet been found to delegate part of the responsibility for appraising scientific evidence to scientists themselves. Besides the worry about ensuring neu-

46. *General Electric Co. v. Joiner*, 522 U.S. at 148, 118 S. Ct. at 520 (Breyer, J., concurring).

47. *Washington Times*, 11 December 1998, byline by Eric Peters.

48. *Wall Street Journal*, 10 December 1998, A22.

49. Jay Reeves, No Implant-Disease Link? ABC News on-line at www.abcnews.go.com/sections/living/DailyNews/breastimplants981201.html (visited 8 July 1999).

50. I note that in June 1999, a thirteen-member committee of the Institute of Medicine also concluded that “silicone breast implants do not cause chronic disease, but other complications are of concern.” News from the National Academies on-line at www.nationalacademies.org/news.nsf/isbn/0309065321?opendocume. June 21, 1999 (visited 10 August 1999).

trality, and the appearance of neutrality,⁵¹ there is the worry about how much responsibility falls on how few shoulders—just four people, in the case of Judge Pointer's panel, all of whom combined this work with their regular full-time jobs, each of them in effect solely responsible for a whole scientific area; and the worry about what jurors will make of court-appointed experts' testimony. The history of the *Frye* test should warn us, also, of potential pitfalls in determining the relevant area of specialization.

■ ■ ■

Here is Justice Blackmun, struggling valiantly if not quite successfully to articulate the mismatch between science and law that lies at the root of the trouble:

[T]here are important differences between the quest for truth in the courtroom and the quest for truth in the laboratory. Scientific conclusions are subject to perpetual revision. Law, on the other hand, must resolve disputes finally and quickly. The scientific project is advanced by broad and wide-ranging consideration of a multitude of hypotheses, for those that are incorrect will eventually be shown to be so, and that in itself is an advance. Conjectures that are probably wrong are of little use, however, in the project of reaching a quick, final and binding legal judgment—often of great consequence—about a particular set of events in the past.⁵²

Yes, we want the law to settle disputes in a timely manner, while scientific inquiry takes—well, it takes the time it takes. Of course, we want cases settled not just promptly but rightly: Mr. Frye to be acquitted if and only if he didn't do it, Mr. Coppolino to be convicted if and only if he did do it, Mr. Joiner to be compensated if and only if his cancer was promoted by his exposure to PCBs, and so on. When scientific evidence is pertinent, we want scientific evidence which is probably right.⁵³ As Jus-

51. Dr. Diamond, one of the members of Judge Pointer's panel, acknowledging that she knows X, Y, and Z, who turn out to be connected in some way with the defendants, remarks that she feels "extraordinarily naive." Transcript of Rule 706 Panel Hearing: 91, available on-line at womnhlth.home.mindspring.com/706/transcript_of_rule_706_panel_hear.htm. February 4, 1999 (visited 9 July 1999). I suspect that the involvement of large industrial concerns in the funding of scientific research is at this point so ubiquitous that it may be quite difficult to find scientists who are both competent to the task required and without such associations as Dr. Diamond acknowledged.

52. *Daubert*, 509 U.S. at 596–97, 113 S. Ct. at 2798.

53. While according to Popper, remember, no scientific claim is ever probable.

tice Breyer reminds us, one of the goals that the Federal Rules of Evidence set themselves is “that the truth be ascertained.”

I don’t mean to suggest that juries can never (perhaps with the help of a cross-examining attorney) spot inconsistencies in scientific testimony, realize that a scientist’s credentials are dubious, notice that the studies relied on were not controlled, or form a reasonable suspicion that a scientific witness is stretching the facts for the sake of a large fee, or, etc.;⁵⁴ nor, of course, that mistakes are only made where scientific witnesses are involved. But as I have been maintaining all along, scientific evidence is “more so”—complex, esoteric, often expressed in an unfamiliar and deeply theoretical vocabulary, and hence unusually difficult for a jury or a judge adequately to assess. (On average, that is; nothing I have said implies that it is more difficult for a judge or jury adequately to assess relatively simple scientific evidence than, say, extremely complicated evidence about accounting procedures.)

No legal form of words can come close to ensuring that only the probable-enough is admitted. *Of course* we want relevant and reliable scientific evidence; but that form of words doesn’t tell a judge anything about what, specifically, to exclude and what to admit (as Peirce might have put it, it reaches only the second grade of clarity, not the third, pragmatic or operational grade). *Of course*, also, scientists in the relevant field are nearly always better judges of the quality of scientific work than the rest of us; but finding a good way to delegate some of the responsibility isn’t trivial, and nothing can ensure that even the most competent and honest scientists will always agree about what is probably right, or that they won’t sometimes agree that, at the moment, they just don’t know.

No wonder scientific evidence provides so many opportunities for opportunism! Often, we are trying to arrive at justice on the basis of imperfect and imperfectly understood information; and not so rarely, we are trying to create justice out of ignorance.

I’m afraid I have been something of an epistemological wet blanket—so much so that by now you may think me an incurable pessimist. So I had better remind you of that nice old Leibnizian joke: “What’s the difference between an optimist and a pessimist? They both think this is the best of

54. The truthfulness of a witness is a matter of (1) whether what he is saying is what he believes to be true and (2) whether what he believes to be true is true. Where scientific testimony is concerned, in general I would think juries are more likely to be able to judge the former than the latter.

all possible worlds”—and assure you that in my opinion this is quite far from the best of all possible worlds.

There are no easy answers; but there are, certainly, better questions and worse. Rather than worrying fruitlessly about the problem of demarcation or the distinction of methodology versus conclusions and all that, we would do better to turn our attention to questions of other kinds—and to keep firmly in mind that, though perfection is impossible, better is better than worse; that the cumulative effect of small improvements can be quite large; and that it is inadvisable to restrict our attention too exclusively to issues and strategies internal to the legal system.

Some of the fruitful-looking questions are practical in orientation: What could be done to help jurors deal better with scientific evidence: e.g., consistent with filtering out legally unacceptable questions, to allow them to ask for clarification when they can't follow an expert witness? What could scientists' professional associations do to help serious scientific witnesses communicate better with judges and juries, or to discourage those who abuse their expertise? Could the legal profession and legal educators do more to discourage unscrupulous witness-shopping and related abuses? What could we learn from the experience with Judge Pointer's panel about bridging some of the gaps between the folkways of science and of the legal system? What advice might best be given to court-appointed scientists about what connections should be disclosed, or what kinds of record-keeping will be expected of them? (Should we consider asking court-appointed scientists to provide details of the qualifications and affiliations of any assistants on whom they relied; of which studies they decided to look at in detail, and why; of which studies seemed most strongly to indicate the contrary conclusion to theirs, and why, in their opinion, those studies were flawed?)

Could we make the legal system more responsive when new evidence comes in to the scientific community?⁵⁵ Could the scientific community be more responsive when legal disputes turn on scientific issues irresolvable by the presently available evidence? Can we think of ways to provide incentives for scientists to study such issues even when they are of much less scientific than practical interest?

Other fruitful-looking questions are more policy-oriented: How significant a gatekeeping role is it appropriate for judges to take? (What

55. Michael Graham observes: "Once [the status of this or that scientific evidence] is set in appellate concrete, a long time might be required to change it when scientific skepticism begins to overtake the original scientific optimism about the validity of the principle or procedure." 99 *Federal Rules Decisions* 188: 222–23 (1983). Or vice-versa.

exactly do we value about trial by jury, and why?) Given that mistakes are inevitable, should we be more willing to tolerate some kinds than others—not forgetting that scientific evidence plays a role both in civil and in criminal cases, and on both sides?⁵⁶ Do we think it appropriate for policy considerations about, for example, how to manage the risks inherent in our reliance on synthetic materials, chemicals, drugs, etc., also to determine what evidence is admissible in criminal cases? (What exactly do we value about uniformity in the legal system, and why?) Are the problems of scientific evidence significantly exacerbated by the contingency-fee system? If so, is it worth the price—presumably, more limited access to the legal system for those without large resources—of changing it? What, ideally, would be the role of tort litigation vis à vis other means of ensuring that, when there is a question about the safety of this or that product, it is carefully looked into, and appropriate action taken?⁵⁷—a question prompted in part by the singularly unfortunate interaction of the FDA and the tort litigation system in the silicone-implant affair.

And, of course: Are these things done differently elsewhere, specifically in the legal systems of other scientifically and technologically advanced countries? If so, what are the benefits, and what the drawbacks?

But it might be prudent, before I begin to tackle such questions, to take Mr. Lec's very shrewd advice, and Think Before I Think . . .

Dedicated to the memory of Richard A. Hausler

56. In the same 1983 Symposium on Science and Rules of Evidence (99 *Federal Rules Decisions* 188: 206), Paul Giannelli writes: "For me *Frye* functions much like a burden of proof. . . . If [in criminal cases] we are going to make mistakes in assessing the validity of a novel technique, they should be mistakes of excluding reliable evidence rather than mistakes of admitting unreliable evidence." Ironically enough, however, in *Frye*, where the novel scientific evidence was proffered by the defense, Giannelli's argument would go exactly the other way.

57. In his concurring opinion in *General Electric Co. v. Joiner* (522 U.S. at 148, 118 S. Ct. at 520), Justice Breyer remarks shrewdly on our ubiquitous dependence on synthetic substances and the importance of ensuring that the "powerful engine" of tort litigation discourages the production only of the harmful stuff (though it spoils the effect somewhat that the case in question concerns PCBs, so dangerous that they have been banned for decades!).

Appendix: Cases Cited

- Commonwealth v. Lykus, 327 N.E. 2d 671 (Mass. 1975)
 Coppelino v. State, 223 So. 2d 68 (Fla. Dist. Ct. App. 1968), *appeal dismissed*, 234 So. 2d 120 (Fla. 1969), *cert. denied*, 399 U.S. 927 (1970)
 Daubert v. Merrell Dow Pharmaceuticals Inc., 509 U.S. 579, 113 S. Ct. 2786 (1993)
 Frye v. United States, 293 F. 1013 (D.C. Cir. 1923)
 Joiner v. General Electric Co., 864 F. Supp. 1310 (N.D. Ga. 1994), *reversed*, General Electric Co. v. Joiner, 78 F.3d 524 (11th Cir. 1996), *reversed and remanded*, 522 U.S. 136, 118 S. Ct. 512 (1997)
 Kumho Tire Co., Ltd., v. Carmichael, 526 U.S. 137, 119 S. Ct. 1167 (1999)
 People v. Williams, 164 Cal. App. 2d. Supp. 858, 331 P.2d 251 (Cal. App. Dep't Super. Ct. 1958)
 Reed v. State, 391 A.2d 364 (Md. 1978)
 United States v. Addison, 498 F.2d 741, 744 (D.C. Cir. 1974)
 United States v. Starzecpyzel, 880 F. Supp. 1027 (S.D.N.Y. 1995)

References

- Bacon, F. 1620. *The New Organon*.
 Bandow, D. 1999. Keeping Junk Science Out of the Courtroom. *Wall Street Journal*, 26 July, A23.
 Bauer, H. 1993. *Scientific Literacy and the Myth of Scientific Method*. Urbana: University of Illinois Press.
 Begley, S., and A. Rogers. 1997. War of the Worlds. *Newsweek*, 10 February, 56–58.
 Black, B., F. J. Ayala, and C. Saffran-Brinks. 1994. Science and the Law in the Wake of *Daubert*: A New Search for Scientific Knowledge. *Texas Law Review* 72:715–802.
 Bridgman, P. 1955. *Reflections of a Physicist*. New York: Philosophical Library.
 Cheseboro, K. 1993. Galileo's Retort: Peter Huber's Junk Scholarship. *American University Law Review* 42:1637–1726.
 Einstein, A. 1936. Physics and Reality. *Journal of the Franklin Institute* 221.3; reprinted in Einstein 1954.
 Einstein, A. 1954. *Ideas and Opinions*, translated by Sonja Bargmann. New York: Crown.
 Frankel, M. S. 1998. The Role of Science in Making Good Decisions. American Association for the Advancement of Science. Testimony before the House Committee on Science, 10 June.

- Gardner, M. 1952. *Fads and Fallacies in the Name of Science*. New York: New American Library.
- Giannelli, P. 1980. The Admissibility of Scientific Evidence: *Frye v. United States*, a Half-Century Later. *Columbia Law Review* 80:1197–1250.
- Goldberg, C. 1999. Judges' Unanimous Verdict on DNA Lessons: Wow! *New York Times*, 24 April, A10.
- Graham, M. 1987. *Federal Rules of Evidence*. St. Paul, MN: West.
- Green, M. D. 1992. Expert Witnesses and Sufficiency of Evidence in Toxic Substance Litigation: The Legacy of Agent Orange and Bendectin Litigation. *Northwestern Law Review* 86:643–699.
- Gross, J., ed. 1983. *The Oxford Book of Aphorisms*. Oxford: Oxford University Press.
- Haack, S. 1990. Rebuilding the Ship While Sailing on the Water. In *Perspectives on Quine*, ed. R. Barrett and R. Gibson. Oxford: Blackwell.
- . 1993. *Evidence and Inquiry: Towards Reconstruction in Epistemology*. Oxford: Blackwell.
- . 1995. Puzzling Out Science. *Academic Questions* spring: 25–31; reprinted in *Manifesto of a Passionate Moderate: Unfashionable Essays*. Susan Haack. Chicago: University of Chicago Press.
- . 1996. Science as Social?—Yes and No. In *Feminism, Science, and Philosophy of Science*, ed. J. Nelson and L. Hankinson Nelson. Dordrecht: Kluwer; reprinted in Haack 1998a, 194–22.
- . 1998a. *Manifesto of a Passionate Moderate: Unfashionable Essays*. Chicago: University of Chicago Press.
- . 1998b. Confessions of an Old-Fashioned Prig. In *Manifesto of a Passionate Moderate: Unfashionable Essays*. Susan Haack. Chicago: University of Chicago Press.
- . 1999. Staying for an Answer. *Times Literary Supplement*, 9 July, 12–14.
- Hand, L. 1901. Historical and Practical Considerations Regarding Expert Testimony. *Harvard Law Review* 15:40–58.
- Huber, P. 1991. *Galileo's Revenge: Junk Science in the Courtroom*. New York: Basic Books.
- . 1992. Junk Science in the Courtroom. *Valparaiso University Law Review* 26:732–755.
- Huber, P., and K. B. Foster, eds. 1997. *Judging Science: Scientific Knowledge and the Federal Courts*. Cambridge: MIT Press.
- Huxley, J. 1949. *Heredity, East and West: Lysenko and World Science*. New York: H. Schuman.
- Institute of Medicine. 1999. News from the National Academies. Available on-line at www.nationalacademies.org/news.NSF/isbn/0309065321?opendocume. 21 June.
- Lec, S. 1962. *Unkempt Thoughts*. New York: St. Martin's.
- Llewellyn, K. 1930. *Bramble Bush: Our Law and Its Study*. 2d. ed. 1951. New York: Oceana.
- McErlean, J., ed. 2000. *Philosophies of Science: From Foundations to Contemporary Issues*. Belmont, CA: Wadsworth.

248 Journal of Health Politics, Policy and Law

- Peters, E. 1998. Benchmark Victory for Sound Science. *Washington Times*, 11 December.
- Portugal, F. H., and J. Cohen. 1977. *A Century of DNA: A History of the Discovery of the Structure and Function of the Genetic Substance*. Cambridge: MIT Press.
- Quine, W. V. O. 1995. *From Stimulus to Science*. Cambridge: Harvard University Press.
- Reeves, J. 1999. No Implant-Disease Link? Available on-line at www.abcnews.com/sections/living/DailyNews/breastimpants981201.html.
- Rogers, A. 1996. Come In, Mars. *Newsweek*, 20 October, 56–57.
- Saltzburg, K., and K. Redden. 1977. *Federal Rules of Evidence Manual: A Complete Guide to the Federal Rules of Evidence*. Charlottesville, VA: Michie Co.
- Schmitt, R. B. 1997. Witness Stand. *Wall Street Journal*, 17 June, A1 and A8.
- Sellars, W. 1965. Scientific Realism or Irenic Instrumentalism? In *Boston Studies in the Philosophy of Science*, eds. R. Cohen and M. Wartofsky. Dordrecht: Kluwer.
- Skrabanek, P., and J. McCormick. 1997. *Follies and Fallacies in Medicine*. Buffalo, NY: Prometheus.
- Starrs, J. E. 1982. A Still-Life Water-Color: *Frye v. United States*. *Journal of Forensic Sciences* 27(3):684–694.
- Wall Street Journal*. 1998. “An Unnatural Disaster” (editorial). 11 December, A22.
- Watson, J. D. 1968. *The Double Helix: A Personal Account of the Discovery of the Structure of DNA*; critical edition, ed. G. Stent, 1980. New York: W. W. Norton.
- Wicker, W. 1953. The Polygraphic Truth Test and the Law of Evidence. *Tennessee Law Review* 22(6):711–742.
- Wilson, E. O. 1999. *Consilience: The Unity of Knowledge*. New York: Alfred Knopf.