Trials and Tribulations: Science in the Law

Susan Haack

University of Miami School of Law, shaack@law.miami.edu

Follow this and additional works at: https://repository.law.miami.edu/fac_articles

Part of the Evidence Commons, and the Law and Society Commons

Recommended Citation

Susan Haack, Trials and Tribulations: Science in the Law, Fall Daedalus 54 (2003).
“I should like to know” [asked Mr. Chichely] “how a coroner is to judge of evidence if he has not had a legal training?”

“In my opinion,” said Lydgate, “legal training only makes a man more incompetent in questions that require knowledge of another kind. People talk about evidence as if it could really be weighed in scales by a blind Justice. No man can judge what is good evidence on any particular subject unless he knows that subject well. A lawyer is no better than an old woman at a post-mortem examination. How is he to know the action of a poison? You might as well say that scanning verse will teach you to scan the potato crops.”

–George Eliot, *Middlemarch* (1872)

Justice requires just laws, of course, and just administration of those laws; but it also requires factual truth. And in determining factual truth, in both criminal and civil cases, courts very often need to call on scientists: on toxicologists and tool-mark examiners, epidemiologists and engineers, serologists and psychiatrists, experts on PCBs and experts on paternity, experts on rape trauma syndrome and experts on respiratory disorders, experts on blood, on bugs, on bullets, on battered women, etc. For, as science has grown, so too has the legal system’s dependence on scientific evidence; it has been estimated that by 1990 around 70 percent of cases in the United States involved expert testimony, most of it scientific. Such testimony can be a powerful tool for justice; but it can also be a powerful source of confusion—not to mention opportunities for opportunism.

Who could have imagined, when DNA was first identified as the genetic material half a century ago, that DNA analysis would by now have come to play so large a role in the criminal justice system, and in the public perception of the law? Even twenty years ago, forensic scientists could tell only whether a blood sample was animal or human, male or female, and, if human, of what type (the least common blood type being found in 3 percent, and the commonest in 43 percent, of the U.S. population). Then, in the mid-1980s, DNA ‘fingerprinting’ made vastly more accurate identification
possible, to probabilities of the order of a billion to one; and by now new techniques have made it possible to amplify and test the tiniest samples.

At first, such evidence was strenuously contested in court; but as its solidity, and its power to enable justice, became unmistakable, the ‘DNA wars’ gradually died down. By the spring of 2002, DNA testing had exonerated more than a hundred prisoners, including a significant number on death row, and helped convict numerous rapists and murderers. In at least one instance, it both exonerated and convicted the same person: after serving nearly eleven years of a twenty-five-to-fifty-year sentence for rape, Kerry Kotler was released in 1992 when newly conducted DNA tests established his innocence; less than three years after his release, he was charged with another rape, and this time convicted on the basis of DNA analysis identifying him as the perpetrator.

Even so, DNA evidence can present problems of its own: police officers and forensic technicians make mistakes—and have been known deliberately to falsify or misrepresent evidence; juries may misconstrue the significance of expert testimony about the probability of a random match with the defendant, or of information about the likelihood that a sample was mishandled—and attorneys have been known to contribute to such misunderstandings; criminals devise devious ways to circumvent DNA identification—and at least one prisoner, apparently hoping to exploit the potential for confusion, has petitioned for a DNA test that, as he must have anticipated, confirmed his guilt.

And who could have imagined, when Hugo Münsterberg urged in his On the Witness Stand: Essays on Psychology and Crime (1908) that the law avail itself of the work of experimental psychologists on the reliability of memory, perception, and eyewitness testimony, that less than half a century later psychological evidence would play a significant role in such landmark constitutional cases as Brown v. Board of Education (1954), or that by now it would have come to play so large a role in the criminal justice system—or that it would be the focus of seemingly endless controversy? For while the work of experimental psychologists on eyewitnesses, memory, etc., has indeed proved useful, clinical psychologists’ and psychiatrists’ diagnoses of this syndrome and that, and especially their theories about the repression and recovery of traumatic memories, have been the subject of heated battles in the courtroom, in the press, and in the academy.

In the mid-1980s, testimony of allegedly repressed and recovered memories came to public attention in the McMartin Preschool case—the longest U.S. criminal trial ever (six years), and one of the most expensive (around $15 million). But in 1990 the seven defendants were acquitted of the ritual sexual abuse that, under the influence of therapists, numerous children at the school had claimed to remember. George Franklin spent nearly seven years in prison for the murder of nine-year-old Susan Nason, convicted on his daughter’s supposed memory of the event, recovered under hypnosis twenty years afterward; he was released in 1996, after his daughter also ‘remembered’ his committing two other murders, with respect to one of which he could be unambiguously ruled out. (Franklin later sued prosecutors and the experts who testified against him for wrongful prosecution and violation of his civil rights.) By the late 1990s, it began to seem that critics such as experimental psychologist Elizabeth Loftus, who had maintained

Science in the courts

Daedalus Fall 2003
55
all along that supposedly repressed and recovered ‘memories’ could be the result of therapists’ suggestive questioning, were vindicated. But recently the ‘memory wars’ have flared up all over again, this time in legal claims filed against Catholic priests accused of sexual abuse of children and young people.

Why has the legal system found scientific testimony hard to handle? Ever since there have been scientific witnesses, lawyers and legal scholars – like Eliot’s Mr. Chichely – have had their doubts about them. The commonest complaint has been that venal scientists brought in by unscrupulous attorneys will testify to just about anything a case demands. In 1858, the Supreme Court observed that “experience has shown that the opposite opinions of persons professing to be experts may be obtained in any amount”; in 1874, John Ordronaux wrote in the American Journal of Insanity that “If Science, for a consideration, can be induced to prove anything which a litigant needs in order to sustain his side of an issue, then Science is fairly open to the charge of venality and perjury, rendered the more base by the disguise of natural truth in which she robes herself.” More than a century later, in Galileo’s Revenge (1991), Peter Huber was sounding a similar theme: junk science – “data dredging, wishful thinking, truculent dogmatism, and, now and again, outright fraud” – was flooding the courts. Some scientists concur. In her study of the silicone breast implant fiasco, Science on Trial (1996), Marcia Angell complains that “[e]xpert witnesses may wear white coats, be called ‘doctor,’ purport to do research, and talk scientific jargon. But too often they are merely adding a veneer to a foregone, self-interested conclusion”; in Whores of the Court (1997), an exposé of flimsy psychiatric and clinical testimony, experimental psychologist Margaret Hagen writes of “charlatans and greedy frauds.”

But other scientists – like Eliot’s Dr. Lydgate – think the real problem is, rather, that jurors, attorneys, and judges are too illiterate scientifically to discriminate sound science from charlatanism. Norman Levitt, for example, commenting in Prometheus Bedeviled (1999) on the “noisome travesty” of the O. J. Simpson trial, complains that “the basic principles of statistical inference were opaque to all concerned except the witnesses themselves. The lawyers..., the judge, the dozens of commentators..., and certainly the woozy public – all seemed utterly ignorant as to what...statistical independence might mean.... All the other scientific issues encountered the same combination of neglect and evasion.”

There surely are venal and incompetent scientific witnesses, and there surely are scientifically ignorant and credulous jurors, attorneys, and judges; but the familiar complaints gloss over many complexities. Scientific testimony may be flawed by outright fraud, or, more often, by the overemphatic presentation of scanty or weak evidence; it may be solid science misapplied by a poorly run laboratory, or serious but highly speculative and controversial science, or sloppily conducted scientific work, or pseudo-scientific mumbo jumbo. The motive may be an expert’s greed, or his desire to feel important, or his anxiety to help the police or a sympathetic plaintiff; or it may be a scientist’s conservatism about new and radical-sounding ideas; or a plaintiff’s attorney’s interest in keeping disputes long settled in science legally alive. Failures of understanding may be due to jurors’ or judges’ or attorneys’ inability to follow complex statistical reasoning, or to their ignorance of the
kind of controls needed in this or that type of experiment or study, or to their excessive deference to science, or their resentment of its perceived elitism. Or the problem may simply be jurors’ sense that someone should compensate the victim of an awful disease or injury, or that someone should be punished for a horrible crime.

And the familiar complaints also gloss over the deep tensions between science and the law that are at the root of these problems. The culture of the law is adversarial, and its goal is case-specific, final answers. The culture of the sciences, by contrast, is investigative, speculative, generalizing, and thoroughly fallibilist: most scientific conjectures are sooner or later discarded, even the best-warranted claims are subject to revision if new evidence demands it, and progress is ragged and uneven. Science doesn’t always have the final answers the law wants, or not when it wants them; and even when science has the answers, the adversarial process can seriously impede or distort communication. It’s no wonder that the legal system often asks more of science than science can give, and often gets less from science than science could give; nor that strong scientific evidence sometimes falls on deaf legal ears, while flimsy scientific ideas sometimes become legally entrenched.

One response to the difficulties has been to try to tame scientific testimony by devising legal rules of admissibility to ensure that judges don’t allow flimsy stuff to be presented to juries. But, as the tortuous history of efforts to frame such formal rules suggests, no legal form of words could guarantee that only good-enough scientific testimony is admitted. Another response has been, instead, to adapt the culture of the law, bringing it more into line with science by compromising adversarialism or the concern for finality. But these pragmatic and piece-meal strategies, though in some ways more promising, raise hard questions about why we value trial by jury, why we want finality, and whether the adversarial process is really an optimal way of ensuring — in the words of the preamble to the Federal Rules of Evidence — “that the truth be ascertained.”

The present practice of relying on experts proffered by the parties not to report on what they saw but rather to give their informed opinion, evolved only gradually, along with the growth of the adversary system, cross-examination, and formal rules governing the admissibility of evidence. For a long time it was required only that a scientific witness, like any other expert witness, establish his qualifications as an expert — until 1923, when the Frye rule imposed new restrictions on the proffered testimony itself.

In Frye, excluding testimony of a then new blood-pressure deception test, the D.C. court ruled that novel scientific evidence was admissible only if it had gained “general acceptance in the field to which it belongs.” At first cited only quite rarely, and almost always with regard to lie-detector evidence, the Frye rule gradually came to be widely followed in criminal trials, and by 1979 had been adopted in a majority of states. (It remains officially the law today in a number of states, Florida included.) Of course, general acceptance is a better proxy for scientific robustness when the field in question is a mature, established scientific specialty than when it is a highly speculative area of research — or, worse, the professional turf of a trade union of mutually supportive charlatans. Moreover, the rule is highly manipula-

---

ble, depending, among other things, on how broadly or narrowly a court construes the field in question. Nevertheless, a main focus of criticism was that the *Frye* test was too restrictive.

The Federal Rules of Evidence (1975) seemed to set a less restrictive standard: the testimony of a qualified expert is admissible provided only that it is relevant, and not legally excluded on grounds of unfair prejudice, waste of time, or potential to confuse or mislead the jury. In line with the Federal Rules’ apparently liberal approach, in *Barefoot*, a 1983 constitutional case, the Supreme Court affirmed that the rights of a Texas defendant were not violated by the jury’s being allowed, in the sentencing phase, to hear psychiatric testimony predicting his future dangerousness – even though an amicus brief filed by the American Psychiatric Association reported that two out of three psychiatric predictions of future dangerousness are mistaken. Justice White, writing for the majority, observed that the Federal Rules anticipate that courts will admit relevant evidence and leave it to juries, with the help of cross-examination and presentation of contrary witnesses, to determine its weight. In dissent, however, noting that a scientific witness has a special aura of credibility, Justice Blackmun averred that “[i]t is extremely unlikely that the adversary process will cut through the facade of superior knowledge.”

By the late 1980s, as legal scholars debated whether the Federal Rules had or hadn’t superseded *Frye*, and whether a more or a less restrictive approach to scientific testimony was preferable, there was rising public and political concern that the tort system was getting out of hand; a crisis due in large measure, Huber argued in his influential book, to

---


---

The *Daubert* case – the first case in the Court’s 204-year history where the central issue was the standard of admissibility of scientific testimony.

*Daubert* was a tort action against Merrell Dow Pharmaceuticals brought by parents who claimed that their children’s severe birth defects had been caused by their mothers’ taking the company’s morning sickness drug, Bendectin, during pregnancy. In excluding the plaintiffs’ expert testimony, the lower court had cited *Frye* (which up till then, contrary to Huber’s diagnosis, had almost always been cited in criminal, not civil, cases). Remanding the case, the Supreme Court held that the Federal Rules had superseded *Frye*, but added that the Rules themselves required judges to screen proffered expert testimony not only for relevance, but also for reliability.

Justice Blackmun wrote for the majority that courts must look not to an expert’s conclusions, but to his methodology, to determine whether proffered testimony is really “scientific . . . knowledge,” and hence reliable. Citing law professor Michael Green citing philosopher of science Karl Popper, and adding a quotation from Carl Hempel for good measure, the ruling suggested four factors for courts to consider: falsifiability, i.e., whether the proffered evidence can be, and has been, tested; the known or potential error rate; peer review and publication; and (in a nod to *Frye*) acceptance in the relevant scientific com-

---

munity. Dissenting in part, however, Justice Rehnquist pointed out that the word ‘reliable’ nowhere occurs in the text of Rule 702; anticipated that there would be difficulties over whether and how Daubert should be applied to non-science expert testimony; worried aloud that federal judges were being asked to be amateur scientists; and questioned the wisdom of his colleagues’ foray into the philosophy of science.

That foray was indeed (if you’ll pardon the expression) ill judged. As Justice Blackmun’s ellipses acknowledge, Rule 702 doesn’t speak of “scientific knowledge,” but of “scientific or other technical knowledge.” However, doubtless influenced by the honorific use of “science” and “scientific” as all-purpose terms of epistemic praise, the majority apparently took for granted that there is some mode of inference or procedure of inquiry, some methodology, that is distinctive of genuinely scientific, and hence reliable, investigation. And so they reached for Popper’s criterion of demarcation, according to which the hallmark of genuine science is that it is falsifiable, i.e., could be shown to be false if it is false; and for his account of the scientific method as conjecture and refutation, i.e., as making bold hypotheses, testing them as severely as possible, and, if they are falsified, giving them up and starting again rather than protecting them by ad hoc maneuvers. Unfortunately, however, Popper’s philosophy of science is singularly ill suited as a guide to reliability; for, if he were right, scientific theories could never be shown to be true or even probable, but at best “corroborated,” by which Popper means only “tested but not yet falsified.” And so the Court ran Popper together with Hempel, whose logic of confirmation does allow that scientific claims can be confirmed as well as disconfirmed.

But Popper’s and Hempel’s philosophies of science are not compatible. Worse, neither can supply the hoped-for crisp criterion to discriminate the scientific, and hence reliable, from the un-scientific, and hence unreliable. No philosophy of science could do this; no such criterion is possible, for not all scientists, and not only scientists, are good, reliable inquirers. Nor is there a uniquely rational mode of inference or procedure of inquiry used by all scientists and only by scientists – no ‘scientific method’ in the sense the Court assumed. Rather, as Einstein once put it, scientific inquiry is “a refinement of our everyday thinking,” superimposing on the inferences, desiderata, and constraints common to all serious empirical inquiry a vast variety of amplifications and refinements of human cognitive powers: instruments of observation, models and metaphors, mathematical and statistical techniques, experimental controls, etc., devised by generation upon generation of scientists, constantly evolving, and often local to this or that area of science.

So perhaps it is no wonder that in the two subsequent decisions in which it has spoken on the admissibility of expert testimony, the Supreme Court quietly backed away from the confused philosophy of science built into Daubert. In the Court’s ruling in Joiner (a toxic tort case involving PCB exposure), references to Hopper, Hempel, falsifiability, scientific method, etc., are conspicuous by their absence; and the distinction between methodology and conclusions, crucial to Daubert, is repudiated as not really viable after all. And in response to inconsistent rulings across the circuits over the applicability of Daubert to non-scientific experts, in Kumho (a product liability case)

involving a tire blowout) the Court ruled that *Daubert* applies to all expert testimony, not only the scientific. According to the *Kumho* Court, the key word in Rule 702 is “knowledge,” not “scientific”; what matters is whether proffered testimony is reliable, not whether it is science.

However, the Supreme Court certainly didn’t back away from its commitment to federal judges’ gatekeeping responsibilities. Far from it. In *Joiner*, the Court affirmed that a judge’s decision to allow or exclude scientific testimony, even though it may determine the outcome of a case, is subject only to review for abuse of discretion, not to any more stringent standard. And in *Kumho*, stressing that the factors listed in *Daubert* are “flexible,” the Court ruled that a judge may use any, all, or none of them. So, abandoning the false hope of finding a form of words to discriminate “reliable, scientific” testimony from the rest, the *Kumho* Court left federal judges with wide-ranging responsibility and considerable discretion in determining whether expert testimony is reliable enough for juries to hear, but with little guidance about how to do this.

Though the *Daubert* ruling spoke of the Federal Rules’ “preference for admissibility,” it imposed significantly more stringent requirements than Justice White had envisaged in *Barefoot*; arguably, indeed, more stringent requirements than *Frye*. (In 2000, revised Federal Rules made explicit what, according to *Daubert*, had been implicit in Rule 702 all along: admissible expert testimony must be based on “sufficient” facts or data and be the product of “reliable” principles or methods, which the witness has “reliably” applied to the facts of the case.) And, despite the usual rhetoric about the Court’s confidence in the adversarial system and in jurors’ ability to sift strong scientific testimony from weak, the *Daubert* ruling involved a significant shift of responsibility from juries to judges, a shift Justice White had resisted. As Judge Alex Kozinski, to whom *Daubert* was remanded,6 caustically observed, he and his colleagues “face a far more complex and daunting task in a post-*Daubert* world…. [T]hough we are largely untrained in science and certainly no match for any of the witnesses whose testimony we are reviewing, it is our responsibility to determine whether the experts’ proposed testimony amounts to ‘scientific knowledge,’ constitutes ‘good science,’ and was derived by the ‘scientific method.’” In a post-*Kumho* world, the task is even more daunting.

In the wry words of Federal Judge Avern Cohn: “You do the best you can.” A sensible layperson might suspect that an expert witness is confused, self-deceived, or dishonest, or that he has failed to take account of readily available relevant information; and should be capable of grasping the importance of double-blinding, independence of variables, etc. But the fact is that serious appraisal of the worth of complex scientific evidence (as Dr. Lydgate pointed out long ago) almost always requires much more than an intelligent layperson’s understanding of science: the specialized knowledge needed to realize that an experimenter failed to control for this subtle potentially interfering factor; that these statistical inferences failed to take account of that subtle dependence of variables; that new work has cast doubt on this widely accepted theory; that this journal is credible, that journal notorious for such-and-such editorial bias.

Since *Daubert* there have been various efforts to educate judges in science – such as the two-day seminar on DNA for

---

Massachusetts Superior Court judges at the Whitehead Institute for Biomedical Research, after which, the director of the institute told The New York Times, they would “understand what is black and white . . . what to allow in the courtroom.” But while a bit of scientific education for judges is certainly all to the good, a few hours in a science seminar will no more turn judges into scientists competent to make subtle and sophisticated scientific judgments than a few hours in a legal seminar would transform scientists into judges competent to make subtle and sophisticated legal judgments; and may risk giving judges the false impression that they are qualified to appraise specialized and complex scientific evidence.

As judges’ gatekeeping responsibilities have grown, so too has their willingness to call directly on the scientific community for help. Since 1975, under FRE 706, a court has had the power to “appoint witnesses of its own selection.” Used in a number of asbestos cases between 1987 and 1990, the practice came to public attention in the late 1990s, when Judge Sam Pointer, to whom several thousand federal silicone breast implant cases had been consolidated, appointed a National Science Panel to report on whether these implants were implicated in the systemic connective-tissue diseases attributed to them. In 1998, the four-member panel reported that the evidence did not warrant claims that the implants caused these diseases. (Six months later, a thirteen-member committee of the Institute of Medicine reached the same conclusion.) The plan had been for the videotaped testimony of panel members to be presented at trial; after the contents of the report became known, however, and before the testimony had been transcribed, most of the cases were settled.

When the report was made public, a headline in The Washington Post hailed it as a “Benchmark Victory for Sound Science,” and an editorial in The Wall Street Journal announced that “reason and evidence have finally won out.” And it is not only those whose sympathies lie with defendant companies in danger of being bankrupted by baseless tort claims who welcome the idea; so do the many scientists impatient with what they see as lawyers’ pointless wrangling over well-known scientific facts. Indeed, where mass torts involve vast numbers of litigants on the same issue, where the science concerned is especially complex, and where hired scientific guns are entrenched on both sides, court-appointed experts may well be the best way to reach the right upshot (and more uniform results than the kind of legal lottery in which some plaintiffs win huge awards and others nothing) – especially if judges learn from Judge Pointer’s experience about the pitfalls of choosing scientists to advise them, and about instructing those scientists on record-keeping, conflict of interest, etc.

Still, though the conclusion the Pointer panel reached was almost certainly correct, it is troubling to think that just four scientists – all of whom combined this work with their regular jobs, and one of whom revealed poor judgment, to say the least, in signing a letter, while serving on the panel, to ask for financial support for another project from one of the defendant companies – were in effect responsible for the disposition of thousands of cases. More radically than Frye’s oblique deference to the relevant scientific community – more radically even than Daubert’s (and Joiner’s and Kumho’s) extension of judges’ gatekeeping powers – reliance on court-appointed scientists departs from the adversarial culture of the common-law
approach. Proponents have recognized this from the beginning: “[t]he expert should be regarded as an amicus curiae” (John Ordronaux); a court should have the power to appoint “a board of experts or a single expert, not called by either side” (Judge Learned Hand, 1901). So have contemporary critics of the practice, such as Sheila Jasanoff, who complain that it is elitist, undemocratic, a move in the direction of an inquisitorial system.

Then there are the ripple effects of those disturbing DNA exonerations, which have prompted not only renewed scrutiny of forensic laboratories, renewed concern about how lineups are conducted and photographs presented to eyewitnesses, moves to videotape interrogations, and so on—all, surely, welcome developments—but also legislation to overcome obstacles to admitting ‘new’ evidence, i.e., the results of new DNA tests on old material. Notwithstanding the law’s traditional emphasis on (in Justice Blackmun’s words) “quick, final, and binding” solutions, some states have mandated post-conviction DNA testing, and others have extended or eliminated the statute of limitations where DNA evidence may be available.

The basic purpose of a trial is the determination of truth,” the Supreme Court averred in a 1966 ruling. “Our system of criminal justice is best described as a search for the truth,” Attorney General Janet Reno affirmed in her introduction to the 1996 National Institute of Justice report on DNA evidence, Convicted by Juries, Exonerated by Science. So we like to think; but it would be more accurate to say that the law seeks resolutions that correspond as closely as possible to the ideal of convicting X if and only if X did it, or obliging Y to compensate Z if and only if Y caused harm to Z, given other desiderata of principle or policy: that it is worse to convict the innocent than to free the guilty; that constitutional rights must be observed; that legal resolutions should be prompt and final; that people should not be discouraged from making repairs that, if made earlier, might have prevented the events for which they are being sued; etc. We also like to think that our adversarial system (under which a jury is asked to decide, on the basis of evidence presented by competing advocates, held to legally proper conduct by a judge, whether guilt or liability has been established to the required degree of proof) is as good a way as we can find to reach the desired balance. But problems with scientific testimony oblige us to think harder both about exactly what balance is most desirable and about the best means to achieve it.

There is no question about the desirability of prompt and final legal decisions; think of totalitarian regimes where people routinely languish in jail without trial, or of Dickens’s Jarndyce v. Jarndyce. Nevertheless, if new scientific work makes it possible to establish that an innocent person has been convicted, it seems obtuse to refuse to compromise finality in the service of truth. And, while it is salutary to remember that the brouhaha over recovered memories also prompted some modifications of statutes of limitations, with DNA analysis there really are the strongest grounds for such an adaptation of the culture of the law.

There is no question, either, that trial by jury is a vastly superior way of getting at the truth than the trials by oath, ordeal, or combat that gradually came to an end after 1215, when the Fourth Lateran Council prohibited priests from participating in such theologically grounded tests. Our adversarial system is a dis-
tant and highly evolved descendant of the first English jury trials; but it is not perfectly adapted for an environment in which key factual questions can be answered only with the help of scientific work beyond the comprehension of anyone not trained in the relevant discipline. We value trial by jury in part because we think it desirable that citizens participate in public life not only by voting, but also by jury service; still, though such participation is a desirable expression of the democratic ethos, civics education for jurors hardly seems adequate justification for tolerating avoidable, consequential factual errors.

But we also value trial by jury for a more fundamental reason: the protection it affords citizens against partial or irrational determinations of fact. Court-appointed experts are no panacea, and there are both legal and practical problems to be worked out; but if, where complex scientific evidence is concerned, we can sometimes do a significantly better job of determining the truth with their help, adapting the culture of the law in this way might afford better protection, and thus better serve the fundamental goal.