

EXPERT EVIDENCE AFTER *DAUBERT*

Michael J. Saks¹ and David L. Faigman²

¹College of Law and Department of Psychology, Arizona State University, Tempe, Arizona 85287; ²Hastings College of the Law, University of California, San Francisco, California 94102; email: Michael.Saks@asu.edu, faigmand@uchastings.edu

Key Words law and science, admissibility, philosophy of science, courts

■ **Abstract** *Daubert* stands for a trilogy of Supreme Court cases as well as revisions of the Federal Rules of Evidence. Together they represent American law's most recent effort to filter expert evidence offered at trial. This review begins by placing the *Daubert* trilogy in the context of earlier judicial efforts to solve the screening problem, which began well before the twentieth century, and then provides a brief explication of evidence law under *Daubert*. Next, we discuss several aspects of the jurisprudence of expert evidence: its connection to debates in the philosophy of science, the practical legal problems courts are trying to solve, and procedural implications. Then we review and discuss varied impacts of *Daubert*: changes in law, marked increases in cases and scholarship relating to expert evidence, and research examining judicial gatekeeping under *Daubert* (civil defendants appear to benefit greatly and criminal defendants hardly at all). We conclude by offering several predictions and prescriptions for the future of expert evidence.

INTRODUCTION

The law of expert testimony provides a lens through which many aspects of modern legal practice can be studied. Every jurisdiction that confronts devising a rule of admission for expert evidence must resolve two basic matters. First, how strict should the rule be? Should it be liberal and allow testimony from virtually all who claim expertise, stopping short perhaps of astrologers and tea-leaf readers? Or should it be conservative and demand rigorous proof of experts' claims of expertise? The second matter that a jurisdiction must resolve is where the real axis of decision making will be. Should courts defer to the professionals in the field from which the experts come, or should they evaluate the quality of the expert opinion for themselves? Implicit in the answers that a particular jurisdiction gives to these two, largely independent, matters are numerous beliefs about legal process and beyond, including its faith in the adversarial process, its confidence in judicial competence, its trust of the jury system, and even its philosophy and sociology of science and empirical knowledge.

Yet, the particular admissibility rules that a jurisdiction adopts might not be reflected in the results reached by the courts that employ them. A jurisdiction,

for example, might set forth a restrictive test that obligates courts to evaluate for themselves the quality of proffered expertise. Such a rule might in theory apply to all cases yet in practice be applied only to civil matters; in criminal prosecutions, perhaps, these same courts, ostensibly applying the same rule, might actually employ a permissive threshold and be highly deferential to the professional field from which the asserted expertise comes.

The study of expert evidence, therefore, must consider both theory and practice. The rules on the books might reflect one set of choices about the legal process, and in practice a wholly different set might operate. In this review of the contemporary state of expert evidence, we consider how the law of expert testimony developed into its current incarnation, describe the processes of and justification for the current state of the law, and examine how it is applied in daily practice.

THE ADMISSIBILITY OF EXPERT EVIDENCE

Courts have long struggled to develop a test to guide their gatekeeping of expert testimony, scientific or otherwise. The task is easily framed: How is a judge to determine which kinds of opinions from which areas of asserted expertise are dependable enough to be permitted at trial? But the task presents what may be an insuperable dilemma: Courts need expert evidence to assist them in making decisions on issues about which they by definition know far less than the expert, yet for that very same reason courts are in a poor position to assess the expertise. The history of rules and procedures for screening expert witnesses represents successive responses to that dilemma.

Before the *Frye* Test

Most discussions of the admissibility of scientific expert testimony begin with *Frye v. U.S.* (1923). This is an odd custom, first because judges had been screening expert evidence for centuries before *Frye*, and second because for decades after *Frye* was decided the case was ignored by both courts and scholars (Faigman et al. 1994). Its influence emerged only when the adoption of the Federal Rules of Evidence drew near, the very time when *Frye* should have become obsolete.

The earliest record of the use in trial of what were then called skilled witnesses was the 1678 trial of the Earl of Pembroke for the murder of Nathaniel Cony (Cobbett 1810), although the use of such witnesses was not regarded as a novelty even in that case. The earliest reported decision affirming the propriety of using expert witnesses proffered by a party occurred in 1782 in *Folkes v. Chadd* (Golan 2004). When such experts did testify, courts were not clear or explicit about the legal principles governing their qualifications or their use. One of the few efforts to discern what test gatekeeping judges (in the nineteenth century) were using

(beyond the witness's qualifications) suggests that the courts employed what could be termed a marketplace test (Faigman et al. 1994). The courts seemed to ask themselves whether expertise had been of value to consumers in the commercial marketplace. If consumers spent their money on an expert or an expertise, then it was presumed to be sound enough for courts as well.

The marketplace test had virtues, but it also had drawbacks (Faigman et al. 2005). The market does not always select for validity. Much that is false, junky, or harmful may nevertheless sell well. The marketplace test honestly applied is unable to distinguish between astronomy and astrology and thus would admit both. In addition, the marketplace test conflates the expert and the expertise. The body of knowledge and the people who purport to possess it tend to be treated as one. A final problem, which ultimately gave rise to the *Frye* test, is that some fields have little or no life in any commercial marketplace. In particular, there are fields that have no function outside of their possible courtroom utility. The courtroom is their marketplace. Where then were judges to look for evaluation help?

Frye and Its Aftermath

The court in *Frye* was confronted with a technology for which there was at the time no commercial market—early polygraph examination. To help it evaluate the admissibility of that testimony, the *Frye* court devised a variation of the marketplace test. It substituted an intellectual marketplace for the commercial one. The court asked whether the principles that underlie the proffered testimony had “gained general acceptance in the particular field in which it belongs.”

The *Frye* variant changed the law's perspective regarding experts in several substantial ways. Principally, by changing the marketplace from the consumers of the expertise to the experts themselves, *Frye* helpfully separated the expertise from the expert. This innovation divided the issue of admissibility more clearly into two parts: (a) the credentials of the expert and (b) the body of knowledge the expert sought to impart. But the *Frye* innovation also, and counterproductively, replaced buyers with sellers as the principal evaluators of the value of what was being offered. Specifically, the test leads the courts to adopt the standards of the field that is the subject of scrutiny. Thus, rigorous scientific fields are judged using strict admissibility standards (because that is how they judge themselves), whereas fields lacking a rigorous tradition are judged using lax admissibility standards (Evet 1993, Saks & Koehler 1991). Even tea-leaf reading is generally accepted if the reference field is composed of practicing tea-leaf readers. Although this transfer of power from consumers to producers may appear peculiar, it was entirely consistent with one of the defining notions of professionalism extant during much of the twentieth century (Haber 1991, Pavalko 1988).

The *Frye* test eventually became a trope for one major notion of the proper criterion for the admissibility of scientific evidence: general acceptance of particular expertise within its field. The test was seemingly easy to apply, required little scientific sophistication on the part of judges, and was to be applied only to

evidence that presented a novel scientific issue, allowing much expert evidence to be scrutinized minimally, if at all.

Offsetting these advantages were limitations: The *Frye* test is vague, is easily manipulated, obscures the relevant inquiry, imposes a protracted waiting period on the use of sound new evidence and techniques, and lacks any definition of when a scientific proposition has become generally accepted. Some products of the most rigorous fields with the healthiest scientific discourse might fail the *Frye* test, while the work of shoddy fields with a great deal of uncritical internal acceptance would easily pass. Moreover, no standards defined what constituted the particular field to which a technique belonged (to one or many fields, and which ones?). Although often criticized for being the most conservative test of admissibility, the *Frye* test could produce the most liberal standards of admission. The more narrowly a court defines the pertinent field, the more agreement it is likely to find. The general acceptance test degenerated into a process of deciding whose noses to count, as well as how many (Black 1988; Faigman et al. 2005, chapter 1; Giannelli 1980; Horrobin 1990; Maletskos & Speilman 1967).

Despite *Frye's* defects, it remains the standard by which expert evidence is evaluated by courts in many jurisdictions. Increasingly, however, the alternative perspective articulated by the U.S. Supreme Court in *Daubert v. Merrell Dow Pharmaceuticals* (1993), and now the standard codified in the Federal Rules of Evidence, is gaining ascendance.

The *Daubert* Trilogy

The Supreme Court's gatekeeping revolution came in three cases, each of them essentially unanimous. *Daubert v. Merrell Dow Pharmaceuticals* (1993) held that the admissibility of scientific evidence depends mainly on its evidentiary reliability (its scientific merit)—suggesting that courts consider whether the scientific basis has been tested empirically, the methodological soundness of that testing, and the results of that testing. These were flexible criteria, so that if courts thought of more appropriate criteria they could use the alternatives. Lower courts were later cautioned, however, against taking flexibility as a license to scrutinize sloppily or not at all: “Though. . .the *Daubert* factors are not holy writ, in a particular case the failure to apply one or another of them may be unreasonable and hence an abuse of discretion” [*Kumho Tire Co. v. Carmichael* (1999), Scalia concurring]. *General Electric v. Joiner* (1997) held that appellate courts must review trial court admission decisions under *Daubert* deferentially and that the logic by which the expert traveled from principles and evidence to a conclusion also is subject to appraisal by the court. Finally, *Kumho Tire Co. v. Carmichael* (1999) held that *Daubert's* essential evidentiary reliability requirement applies to all fields of expert evidence, not only to science. *Daubert* retained the general acceptance criterion, though in downgraded status, and *Kumho Tire* demoted it further.

Although the standards set forth in the *Daubert* trilogy were ostensibly mere interpretations of the applicable Federal Rules of Evidence, those rules were amended

in 2000 to reflect trial courts' obligations to insure the soundness of expert evidence as prescribed in these cases. Specifically, three rules governing opinion testimony—Rules 701, 702, and 703—were amended. Rule 701, which permits lay witness opinions under certain circumstances, was strengthened to ensure that testimony that should be evaluated under Rule 702 did not slip in through the back door of Rule 701. Rule 702 essentially codified *Daubert* by adding three new numbered clauses. The rule now states:

If scientific, technical, or other specialized knowledge will assist the trier of fact to understand the evidence or to determine a fact in issue, a witness qualified as an expert by knowledge, skill, experience, training, or education, may testify thereto in the form of an opinion or otherwise, if (1) the testimony is based upon sufficient facts or data, (2) the testimony is the product of reliable principles and methods, and (3) the witness has applied the principles and methods reliably to the facts of the case.

Rule 703 had often been used to import otherwise inadmissible hearsay statements into evidence (Carlson 1992). The rule has now been amended to prevent this. The new Rule 703 provides, in relevant part: "Facts or data that are otherwise inadmissible shall not be disclosed to the jury by the proponent of the opinion or inference unless the court determines that their probative value in assisting the jury to evaluate the expert's opinion substantially outweighs their prejudicial effect."

The federal courts have averaged about 500 decisions per year on *Daubert*-related issues, and law journals have published thousands of articles at least touching on the subject. The states also are actively involved in this area, with more than half the state courts now following the basic *Daubert* approach (though not all of them the full trilogy) (Bernstein & Jackson 2004), and many other state courts are deeply influenced by *Daubert* (Faigman et al. 2005). Some *Frye* states have on occasion interpreted their test in ways that bear a strong resemblance to *Daubert* principles (Florida and New York are the most notable).

The Meaning of *Daubert*

In essence, the *Daubert* trilogy adopts a changed perspective and relocates the axis of decision. With the old commercial marketplace test, judges piggy-backed onto what consumers seemed to think about a proffered expertise and expert. Under *Frye's* general acceptance test, judges took a rough nose count and deferred to what the producers of knowledge thought about the knowledge they had to offer. *Daubert* finally places the obligation to evaluate the evidence where one might have expected it to be all along: on the judges themselves. For empirical or scientific proffers, *Daubert* requires judges to evaluate the research findings and methods supporting expert evidence and the principles used to extrapolate from that research to the task at hand (Risinger 2000a). And for nonscience expertise (that is, expertise on questions that are seldom the topic of systematic empirical investigation), courts might have to develop new criteria for evaluating the soundness of proffered expert

evidence. [But see Sanders (2001), discussed below, who argues that the basic principles of *Daubert* can be applied, if only by analogies still to be worked out, to all expert knowledge.] This obligation on the part of judges is daunting. It may be more apparent now than it was for centuries before why judges sought ways to avoid such responsibility (and why, notwithstanding the commands of *Daubert*, many of them still do).

Daubert, in many respects, appeared to be a revolutionary decision. Certainly judges', scholars', and lawyers' reactions to it support this view (see, e.g., *Weisgram v. Marley* 2000). The core principle of *Daubert* is its changed focus from *Frye's* deference to the experts to a more active judicial evaluation of a particular field's claims of expertise. Under *Frye*, judges did not need to understand research methodology because it was sufficient to inquire into the conclusions of professionals in the pertinent fields. *Daubert* mandates that judges query which methods support the scientific opinions that experts seek to offer as testimony, and this requires that they understand those methods and data. In *Daubert v. Merrell Dow Pharmaceuticals*, for instance, questions about testing and error rate led to responses about comparison groups, standard deviations, relative risk, statistical significance, and many other concepts foreign to the average lawyer. The revolutionary core of *Daubert* is in this call for judges to become knowledgeable about basic research methods. *Daubert*, in effect, brought the scientific revolution into the courtroom.

Revolutions inevitably produce partisans having widely varying views, including some who defend the old regime, others who seek to justify the new order, and still others on either side of the barricades, who determine its ultimate fate. Many of the battles over the *Daubert* revolution have been carried out in the law review literature, where the debate moved quickly from whether a revolution had occurred at all to the nature of that revolution, and, even more so, to the philosophical justifications for it. For instance, some commentators argue that in *Daubert*, "the U.S. Supreme Court took it upon itself to solve, once and for all, the knotty problem of the demarcation of science from pseudoscience" (Goodstein 2000), or that the Court adopted and imposed a specifically experimental or Newtonian or Popperian view of science. But, in the legal context in which the *Daubert* trilogy arose and to which it pertains, the Court can be seen as trying to solve more flexibly a more modest (though similarly enduring and knotty) problem of trial evidence, namely, how to filter proffered expert opinion testimony so that reliable evidence is admitted and unreliable evidence is not. *Daubert* confronted a particular type of expertise, empirical claims, that lends itself to evaluation by scientific methods. *Daubert's* answer, in essence, is that if the proponent of such evidence cannot supply good grounds for concluding that the expert opinion is sufficiently trustworthy—cannot supply appropriate validation—then the testimony should be excluded. It added that the obligation to test the soundness of expert proffers is applicable to timeworn as well as to novel testimony. Given that *Daubert* itself was a case about epidemiological (correlational) data, the charge that it wrongheadedly demands experimental data is hard to support. Still, one can debate whether the

best filter has been chosen. In addition, one can debate the philosophical justifications for the revolution itself. How long this philosophical debate will endure only time can reveal, but it certainly occupied a prominent place in the *Daubert* era's first decade.

JURISPRUDENCE OF EXPERT EVIDENCE

In the course of writing *Daubert*, the Court, perhaps somewhat improvidently, cited Sir Karl Popper and his notion of falsifiability as the defining criterion of a scientific statement. As a consequence, the Court seemed to step into the quagmire of the philosophy of science. Justice Blackmun's citation to Popper has been roundly criticized, beginning with the concurring opinion of Chief Justice Rehnquist. The chief justice complained that neither he nor, he supposed, most federal judges understood what falsifiability was. Moreover, the majority opinion, he complained, appeared to call upon judges to be amateur scientists, a role for which they were not trained and in which they were not likely to excel. Scholars also have challenged Blackmun's conception of the philosophy of science, alternating between pointing out that by 1993 Popper was no longer *au courant* among philosophers and that, in any case, Popper's philosophy was largely inapposite to the Court's ultimate holding. Not surprisingly, no citations to Popper are to be found in the subsequent two legs of the *Daubert* trilogy. Despite the protestations and general hand-wringing among courts and scholars, the Popperian perspective offers useful insights into the regime enacted by *Daubert* and provides a basic justification for the Court's decision.

The *Daubert* Court's Philosophy of Practicality

As is true in many legal contexts, especially when it comes to Supreme Court opinion construction, Blackmun's citation to Popper was meant to be more illustrative than necessary to understanding or applying the opinion. For Blackmun, Popperianism was a synecdoche for the Court's desire to distinguish proper scientific inquiry from the perceived detritus of "junk science" (Edmond & Mercer 2002). Good science, according to the Court, follows certain methodological conventions. Bad science does not. Accordingly, the Court suggested guideposts for lower courts in their new obligations to admit good science and exclude the junk. The vaunted *Daubert* four factors—(a) testing, (b) peer review and publication (which Blackmun suggested mainly as a means of assisting judges in evaluating research methodology), (c) error rate, and (d) general acceptance—are essentially aspects of the ordinary conduct of scientific investigation.

The *Daubert* Court, therefore, was engaged in the rather pedestrian activity of articulating a test by which lower courts could make decisions regarding the admissibility of expert evidence. The decision must be understood in those terms. Unlike philosophers of science, trial courts must make concrete decisions in particular

cases. But in articulating this evidentiary standard, Justice Blackmun effectively entered the science wars (Haack 2003a). By citing Popper, rather than, say, Thomas Kuhn (1996), the Court was signing on to scientific realism. Indeed, it is not terribly surprising that the Court, in seeking to establish standards for decisions in concrete cases, would reject relativism in favor of a more positivistic and objective form of realism.

Although commentators' criticisms of the Court's seeming philosophy are not illegitimate, they are largely beside the point (Leiter 1997). Popper's notion of falsification was concerned with a fairly narrow philosophical issue involving the problem of demarcating science from pseudoscience. Only peripherally was Popper's analysis relevant to demarcating admissible from inadmissible expertise (O'Connor 1995). And this was precisely how the Court used Popper. Falsifiability was shorthand for how scientists largely understand their jobs (Lewontin 1994), and the Court was trying to incorporate scientific sensibilities into the legal culture (Faigman 1999, Haack 2003b). The four factors the Court chose are immediately recognizable as central to the scientific enterprise. Few scientists would disagree that their task is to test hypotheses with a view to describing quantitatively the world around them, whether it is the effects of Vioxx[®] or the causes and effects of global warming. Moreover, virtually all scientists are concerned with the quality of research design and methods, and they understand peer review and publication as a standard component of the process of checking methodology before disseminating research. Finally, general acceptance of one's findings is the hoped-for end result of the process. Thus, the Popper reference in *Daubert* can be seen to be more a synecdoche for good science than a literal answer to some unasked question in the philosophy of science.

Interestingly, when the Court eventually reached the legal question that was more synonymous with the one for which Popper is so closely associated—demarcating science and pseudoscience—the Court neither cited him nor attempted to draw a bright line around which statements count as scientific. In *Kumho Tire*, the Eleventh Circuit, as had several other circuits, concluded that *Daubert's* gatekeeping standard applied only to scientific testimony and did not pertain to “technical or other specialized knowledge” that constitutes the other kinds of expert testimony contemplated by Rule 702. The Supreme Court in *Kumho Tire*, however, held that *Daubert* applied to all expert testimony, not just the determinedly scientific variety. Whereas Popper had good philosophical reasons for trying to define those statements that qualify as scientific, the Court had good legal reasons for eschewing that task (Jonakait 1997).

The basic challenge for trial courts in the area of expert testimony is to define the boundary between admissible and inadmissible evidence. As the *Kumho Tire* Court understood, the definition of adequate science is only a subpart of this greater task. Expertise comes to court in myriad forms, ranging from the most traditionally rigorous fields, such as physics, to the most traditionally lax, such as clinical medicine. Some experts dress in the guise of science, such as forensic document examiners, whereas others claim expertise by virtue of experience alone,

such as police officers. The one thing all these ostensible experts have in common is their claim to opinions that are relevant and sufficiently accurate to be helpful to the trier of fact.

Daubert's holding, in fact, was limited to this helpfulness assessment. *Daubert* described trial courts as gatekeepers whose responsibility is to preliminarily assess whether proffered expert testimony is relevant and whether its basis is reliable. Many courts and commentators, however, have confused this basic holding with the factors that Justice Blackmun suggested might be useful to carrying out this gatekeeping duty (see, e.g., Crump 2003). Consequently, these scholars have accused the Court of naively applying a positivistic model of science onto all expert testimony. Reading *Daubert* as requiring the application of the very same criteria to any and all expert evidence is plainly absurd. Auto mechanics, for example, are not likely to publish their ideas in peer-reviewed journals or know what error rates apply to the technology they rely upon. But the so-called four-factor test of *Daubert* was never more than a set of suggested criteria by which to evaluate ostensibly scientific evidence. As the Court has repeatedly said, including in the *Daubert* opinion itself, no one set of criteria would be useful to assess the validity of every kind of science (from physics to biology), much less every kind of expertise (from engineering to real estate appraisals).

The holding of *Daubert*, as made clear in *Kumho Tire*, applies to all expert testimony. It provides, simply enough, that trial courts are obligated to determine whether the basis for proffered expert testimony is, more likely than not, reliable and valid. The four *Daubert* factors will often help courts make that determination, and sometimes they will not. In *Kumho Tire*, the Court declined any attempt to set forth a single set of criteria that might be useful in assessing the myriad kinds of expertise the courts hear. The point is that trial judges are obligated to carry out the gatekeeping function; how they do so is a separate question. Therefore, in the *Daubert* trilogy, the Court was engaged in the task of defining a rule of procedure that would apply to all forms of expertise. Philosophers and sociologists of science could offer insights into the difficulty of the task, but their views have limited relevance to whether the Court chose the correct rule for its purposes.

The *Daubert* trilogy, however, has left much that still needs to be done. The four *Daubert* factors offer some guidance regarding a large proportion of experts, particularly those from professional fields in which quantitative empirical methods can be, and ordinarily are, employed. The Court in *Kumho Tire*, however, made no attempt to offer similar sorts of criteria for evaluating experts for whom some or all of the *Daubert* criteria might not be decisive or sufficient. Auto mechanics, historians, accountants, clinical medical doctors, and scores of others have traditionally testified but would not be able to meet one or more *Daubert* criteria. Clearly, the Court and the Rules of Evidence contemplate that many experts from these fields would still be permitted to testify, but, at the same time, trial courts must in some way determine if the bases for the opinions they intend to offer are sufficiently valid to admit.

Evaluating Expertise

Commentators have suggested various ways for courts to evaluate this wide variety of expert opinion. The first challenge is to establish expectations regarding the testing that should have been done. In making the necessary validity assessment, should courts consider whether an opinion could be rigorously tested, or should they simply accept the standard practice of the particular field, which might include relying on experience as a basis of expertise? For example, perfume testers might believe they gain expertise through long experience with different kinds of perfumes, and it might not occur to them to carry out validity tests to measure this belief. Clearly, such testing could be easily done, but just as clearly professional perfume testers generally do not do so. Should a court demand that these tests be done, or that a particular perfume tester undergo proficiency testing before his or her opinion will be admitted into evidence (Faigman 2002)?

In addition to the level of rigor to be expected, courts must determine how (or pursuant to what criteria) they should evaluate the numerous expertises that rely on a wide variety of methods, from casual experience to controlled experiment (Seton Hall Symposium 2003). Scholars have begun to address this issue, with several setting forth taxonomies that might be employed for different categories of expertise (Gross & Mnookin 2003, Risinger 2005). Sanders (2001) argues for a different approach. He suggests that “[t]he judicial task will be made more manageable if in both scientific and nonscientific testimony, the courts assess the expert’s reasoning from an objective, rational processing perspective with a single test” (p. 409). In favoring this rational form of information processing, Sanders rejects as undependable the basic alternative foundation of most expert opinion: experientially based information processing. As a practical matter, he states that this approach means that courts “should focus on the first *Daubert* criterion” (Sanders 2001, p. 409). The testing criterion, he explains, has at its core a concern with methodology. It asks the expert to describe objectively how the hypothesis at issue can be tested and how the expert put the hypothesis to the test (Sanders 2001, p. 409).

Social Construction and Gatekeeping

Implicit in much of the debate about the nature of scientific knowledge is another controversial matter that concerns whether judges as gatekeepers usurp the role of jurors in the trial process. Scholars who reject the scientific realism of *Daubert*, preferring a more Kuhnian-based relativism, have also bemoaned what they believe to be *Daubert*’s conferral of excessive power on judges. For instance, sociologists of science who believe that scientific knowledge is largely socially constructed have generally doubted the wisdom of the gatekeeper function at the core of *Daubert*. However, the tenets associated with a social constructionist view of science are not necessarily inconsistent with the gatekeeping role defined in *Daubert*. Consider, for example, the views of Jasanoff (1992), a particularly influential sociologist of science, who has questioned the wisdom of adopting a

restrictive rule of admissibility, such as that of *Daubert*. According to Jasanoff (1992, p. 347),

[t]he most significant insight that has emerged from sociological studies of science in the past 15 years or so is the view that science is socially constructed. . . . [T]he “facts” that scientists present to the rest of the world are not simply reflections of nature; rather, these “facts” are produced by human agency, through the institutions and processes of science, and hence they invariably contain a social component.

This insight leads ineluctably to the conclusion that judges should exercise their power to exclude expert evidence with restraint. When judges exclude experts,

[t]hey help shape an image of reality that is colored in part by their own preferences and prejudices about how the world should work. Such power need not always be held in check, but it should be exercised sparingly. Otherwise, one risks substituting the expert authority of the black robe and the bench for that of the white lab coat—an outcome that poorly serves the causes of justice or of science (Jasanoff 1992, p. 359).

The view that science might be socially constructed is not foreign to courts’ understanding, and, indeed, the *Daubert* Court cited Jasanoff when it cautioned that peer review and publication are not the *sine qua non* of good science. At the same time that the Court’s reliance on Popper and Hempel reflected its scientific realism, its acknowledgment of a role for social construction reflected courts’ traditional, cautious approach to experts and advocates. Nor is social construction foreign to the understanding of scientists. The very purpose of scientific method is to try to minimize the contribution of bias (borne of social or personal construction) and maximize the contribution of evidence of the phenomenon under study. This reflects what is probably the dominant view among thoughtful scientists, scholars, and courts: a realist-constructionist view of science (social construction constrained by the empirical world) (Cole 1992, Haack 2003a, Sanders et al. 2002).

But assessing whether research methodology is successful, and therefore whether scientific knowledge, or empirical knowledge more generally, is socially constructed, does not answer the question of what standard of admissibility should apply to expert testimony. Indeed, from a legal-structural perspective, *Daubert*’s gatekeeping standard is entirely consistent with at least the nonradical version of the social construction of science. Jasanoff, for example, cites the fear that the “preferences and prejudices” of the “black robe and bench” will replace “the white lab coat.” But the more that science is socially constructed, the less the black robe should defer to the white lab coat. Judges have the institutional and, in most respects, the constitutional obligation to ensure due process and fair and balanced trial procedures. To the extent that expert testimony is infused with “preferences and prejudices,” they should be those of the judge and not the expert. The responsibility to exercise such preferences and sometimes impose such prejudices devolves upon judges in our constitutional system. Jasanoff, in contrast, would

invest this authority first in experts who are not accountable in the way that judges are and, second, in juries to judge whether the evidence given them is worth any consideration. Experts are fact witnesses, and their value to the trial process lies exclusively in their capacity to assist jurors to adjudicate disputed facts. Judges are charged with the discretion to interpret the law and apply the values inherent in these interpretations to particular disputes. If anyone's preferences and prejudices are going to infuse the trial process, it should be those of the judges, whose biases (such as they are) are imposed with political legitimacy. And, if we hope to limit the effect of bad, biased, or seriously misleading testimony, judicial gatekeeping is our best hope.

Law uses expertise (whether scientifically, technically, or experientially based) as a tool for its own purposes. A philosophy of science answers only a small portion of the questions involved in developing a philosophy of expert evidence. Whereas science strives for truth, truth is only one component of the law's mission.

Expertise and Procedural Considerations

Ordinarily, rulings regarding the admissibility of evidence are firmly within the trial court's discretion. The principal reason for this is that trial courts are in a better position than appellate courts to screen evidence, an essential part of conducting a trial. In *Joiner* and *Kumho Tire*, the Supreme Court followed the conventional wisdom and held that appellate courts owe substantial deference to the trial court, both in the criteria used to assess the validity of proffered expertise (*Kumho Tire*) and the ultimate admissibility decision (*Joiner*). Under the *Daubert* trilogy, rulings on expert evidence, like other evidentiary rulings, can be overturned on appeal only for an abuse of discretion. This approach makes jurisprudential sense if expert evidence is like other kinds of evidence, but it is not.

Some questions are case specific, whereas others are relevant across a broad spectrum of cases. For example, whether or not a car in a particular case went through a red light has no implications for what color a traffic light was in other cases. But if "Which color grants a driver the right of way?" has a different answer from case to case, that would be arbitrary and lawless. Questions of fact, which typically affect only the case before the court and do not have meaning for other future cases, can be altered on appeal only when clear error is found. In contrast, matters of law, which apply across cases, are reviewed *de novo*. This differential treatment grants deference on some kinds of questions and consistency on other kinds, and facilitates judicial efficiency along with the rule of law: Once green is declared by a legislature or appellate court to indicate go, that question is decided for all trial courts in the jurisdiction.

Similarly, in most evidentiary contexts, admissibility decisions are case specific. Scientific evidence, however, does not conform to this traditional wisdom. Many scientific findings transcend individual cases. Questions such as whether Bendectin is a teratogen, smoking causes lung cancer, or polygraphs detect lying do not, in principle, vary from case to case. (Conversely, such questions as the fit of a body

of knowledge to issues in a particular case or whether a given expert has correctly applied a technique in a given instance are case specific.)

Monahan and Walker, writing before the *Daubert* decisions, were the first to explore the implications of this insight (Monahan & Walker 1986; Walker & Monahan 1987, 1988). Monahan and Walker argue that although facts and law differ in that one is positive and the other normative, facts sometimes share an important similarity with law: Some factual issues are case specific and some transcend individual cases. The Monahan-Walker analysis has procedural implications. Facts that are trans-case in nature should be treated much as law is treated: Subject to de novo review on appeal, courts are not obligated to rely on the record developed by the parties but can engage in their own inquiries, and the holdings of higher courts should be binding on lower courts.

To allow issues such as whether smoking increases the risk of lung cancer to be decided one way in one case in one trial court and differently in another case in another trial court in the same jurisdiction would strike most observers as plainly irrational. Moreover, to require appellate courts to declare that contradictory answers to such questions by lower courts are, in the absence of clear error, both right, as the holding in *Joiner* requires, merely highlights the problem. Were courts to make such rulings, scientists, editorial writers, and the general public would wonder out loud about how courts can possibly believe that two contradictory propositions about unchanging general phenomena can both be regarded by the law as true. When the problem presents itself starkly, courts certainly see the implications. Thus, for example, when confronted by two different district courts reaching opposite conclusions about whether billboard advertising of alcohol increases auto accidents (and would therefore be subject to regulation), the U.S. Court of Appeals for the Fifth Circuit ruled that it had no choice but to examine the underlying empirical question for itself and reach one decision for both cases below (*Dunagin v. City of Oxford* 1983).

Of course, the scientific questions and their answers are not always, or ever, straightforward. The procedural suggestions of Monahan and Walker work best when the answers are reasonably well studied and accessible to the courts, which may require allowing the science and judicial experience with the matter to ripen before an appellate court rules as a matter of law. Even then, an appellate court might get the science wrong, so that courts at all levels need to be ready to revisit issues they thought had been resolved (whenever a party makes a sufficient showing)—just as they do when an error of law has been made. Moreover, we might expect scientific conclusions to be more subject to change than are normative conclusions because the former are more vulnerable to the growth of knowledge and convincing demonstrations of past errors than normative conclusions generally are.

The preceding discussion about the law-likeness of some scientific questions applies to decisions whether or not to admit expert evidence. A certain kind of medical knowledge, applied to a certain class of problems, cannot be valid in some jurisdictions and invalid in others. Basic principles of microscopic hair comparison

evidence or the identification of handprinting or fragmentary fingerprints cannot be invalid in some places and valid in others.

THE IMPACT OF *DAUBERT*

What difference does *Daubert* actually make? The answer is multifaceted. Although changes in the gatekeeping behavior of judges might be the kinds of effects that first come to mind, one ought not to overlook changes in the behavior of lawyers, litigants, and expert witnesses—either in anticipation of or in reaction to changes in judicial treatment of expert evidence—or the discussions that scholars and those in the legal system generally have about expert evidence.

Anticipation of Increased Scrutiny

The most unmistakable impact of *Daubert* has been the changes in federal evidence law, changes that have been radiating into state law. Another unmistakable impact has been the production of many hundreds of scholarly articles on the admissibility of expert testimony, most of them clearly undertaken in the wake of *Daubert*, concerning the meaning and role of the *Daubert* trilogy and their application by lower courts. The federal courts alone have recently averaged about 500 decisions per year on *Daubert*-related issues. Many of those decisions involved *Daubert* hearings at which scrutiny was given to proffers of expert evidence, many of which, before *Daubert*, entered court with little if any scrutiny. If nothing more, one can certainly say that the law has changed and that people have been talking about those changes.

Some courts and commentators assumed that *Daubert*, like *Frye*, applied only to novel scientific evidence. Those who thought so were wrong. The *Daubert* opinion explicitly mentions (though only in a footnote) that the Court did “not read the requirements of Rule 702 to apply specially or exclusively to unconventional evidence.” It continues that “well-established propositions are less likely to be challenged than those that are novel, and they are more handily defended.” This plainly suggests that long veneration was no protection from scrutiny. And, as the Court later taught in *Kumho Tire*, *Daubert* was not even limited to science.

Many courts and commentators were unsure or disagreed about whether *Daubert*, in contrast to *Frye*, raised or lowered the threshold of admission. Gatowski and colleagues’ (2001) survey of judges found that 32% believed the intent was to raise the threshold of admissibility for scientific evidence, 23% believed the intent was to lower the threshold, and 36% believed the intent was neither to raise nor to lower but instead to articulate a framework for admissibility. Numerous courts have expressed surprise to discover that their application of a supposedly more liberal test led them to the brink of excluding evidence that had never before appeared so excludable. The better answer probably is that “it depends.”

Permit us to oversimplify here to illustrate an essential point. The *Frye* and *Daubert* tests look at different attributes of scientific propositions. *Frye* asks how

TABLE 1 Outcomes of scrutiny of expert evidence under the *Frye* and *Daubert* tests

<i>Frye</i> : General acceptance in scientific community	<i>Daubert</i> : Valid scientific foundation	
	Strong	Weak
High	Both admit	<i>Frye</i> admits <i>Daubert</i> excludes
Low	<i>Frye</i> excludes <i>Daubert</i> admits	Both exclude

generally accepted the proposition is in a reference community or communities. *Daubert*, in contrast, inquires directly into the proposition’s scientific foundations. As Table 1 illustrates, these two questions should usually lead to the same answer about admissibility. That which has a strong scientific foundation usually will be generally accepted; that which has a weak scientific foundation usually will not be widely accepted. In either situation, both tests should lead to the same decision to admit or to exclude.

But there are circumstances in which the underlying attributes of the expert evidence diverge. When a scientific proposition is sound but not generally accepted, *Daubert* should admit while *Frye* should exclude. This is the category of cases that most commentators and courts had in mind when they suggested that *Daubert* is more liberal than *Frye*. But when a scientific proposition has not been shown to be sound yet nevertheless has gained general acceptance in its field, then *Daubert* excludes even though *Frye* admits. This latter category is not a null set; it contains, perhaps most notably, many of the forensic sciences.

The anticipated impact of *Daubert*’s filtering of expert evidence prompted concern, litigation, and even empirical research in fields that had done little previous testing of their theories. For a time, some fields tried to evade scrutiny by re-defining themselves as nonscience or by emphasizing their art over their science. Pursuing this strategy, a consortium of law enforcement organizations—fearing that the assertedly expert testimony of forensic scientists and police officers would be excluded if they were to be required to prove that what they were saying had a sound basis—submitted an amicus brief to the Supreme Court in *Kumho Tire*. The brief urged the Court to exempt from *Daubert* scrutiny prosecution expert evidence, “the great bulk of [which] does not involve scientific theories, methodologies, techniques, or data in any respect. . .” but instead offers opinions “about such things as accident reconstruction, fingerprint, footprint and handprint [identification], handwriting analysis, firearms markings and toolmarks and the unique characteristics of guns, bullets, and shell casings, and bloodstain pattern identification” (Am. Eff. Law Enforc. et al. 1997). Nevertheless, the decision in *Kumho Tire* extended *Daubert* scrutiny to all fields of asserted expertise.

Daubert and *Kumho Tire* prompted expectations that if fields such as the forensic sciences that lacked a basis in sound research were to survive scrutiny they would

need to undertake the necessary research. Such thinking led the National Institute of Justice to launch several funding initiatives, inviting researchers to begin to fill the considerable gaps in the knowledge claims of these fields that they were sure the courts would now discover in the glare of *Daubert* and *Kumho Tire* scrutiny (e.g., Natl. Inst. Justice 2000).

Similarly, civil litigants, especially tort plaintiffs, anticipated that *Daubert* would raise barriers to their proffered expert evidence—although less dramatically than the prosecution sciences because most civil plaintiff experts come from fields such as medicine and engineering, where research and education are better established and more systematic. Plaintiffs' attorneys were advised to choose their experts more carefully and prepare their experts (and themselves) more thoroughly to meet *Daubert* scrutiny; some were advised to avoid filing their cases in the federal courts to avoid *Daubert* scrutiny entirely (Assoc. Trial Lawyers Am. 2004).

One can find articles by practitioners and scholars in various fields reflecting anxieties about the prospect of having their offerings challenged and tested under *Daubert*. Some scholars and researchers point out that the Supreme Court's decision in *Daubert* has had the salutary effect of precipitating improvements in their own fields (e.g., Shuman & Sales 1999). Yet others argue that some fields are unlikely to make any improvements if not impelled to do so by the threat of exclusion or limitation (e.g., Faigman et al. 2005, chapter 1).

Admissibility on the Books versus Admissibility in Action

That the law in action can sometimes be quite different from the law on the books is something legal realists noted decades ago. The implementation of the *Daubert* trilogy can be viewed as an immense case study of this phenomenon. But the picture is not a simple one.

Several studies have examined the patterns of admissibility decisions in cases decided prior to and after the adoption of *Daubert*, casting at least some light on the behavior of both federal and state courts in several categories of cases. The patterns of change and nonchange provoke one to think about the nature of courts and the society in which they are embedded.

Comparisons of the rate of pretrial challenges to the admissibility of expert evidence before and following *Daubert* found, overall, a marked increase [Risinger (2000b), both civil and criminal case samples in both federal and state courts; Dixon & Gill (2001), federal civil cases; Krafka et al. (2002), federal civil cases]. But, in the civil arena, Risinger (2000b) found that nearly 90% of the challenges were raised by defendants against plaintiffs' expert evidence. Among the criminal cases, where the overwhelming bulk of expert evidence is offered by the government, defendants are far less active in bringing challenges, often failing to raise objections that would have been reasonable and available, and which presumably would have been raised in a civil case involving evidence with similarly weak foundations. In federal courts, fewer than 10% of the challenges to expert evidence were in criminal cases. Of those, the prosecution brought more challenges to defense evidence than

vice versa by a ratio of 7:2, even though the government presents the far larger target for attack.

When a challenge is mounted, how do the courts respond? In civil cases, post-*Daubert* courts are more likely to exclude challenged expert evidence than they had been before. Dixon & Gill (2001) found that challenged expert evidence was excluded about 50% of the time pre-*Daubert*; that figure rose to as much as 70% in years post-*Daubert*. In surveys conducted by Krafka et al. (2002) before and after *Daubert*, federal judges reported excluding or limiting challenged expert evidence 25% of the time pre-*Daubert*, compared with 41% of the time post-*Daubert*. But the data on excluded evidence reveal a notable lack of symmetry between the success of plaintiffs compared with defendants. Risinger (2000b) found that defendants succeeded about two thirds of the time in the many federal cases in which they challenged plaintiff experts. In the smaller set of cases in which plaintiffs challenged defense-proffered expertise, the challenges succeeded less than half the time. This pattern was repeated on appeal. In state civil cases, Risinger found that challenges by plaintiffs and by defendants succeeded at about the same 40% rate, but of course defendants were more active in bringing challenges (82% of the challenges on appeal were by defendants). A reading of the cases confirms that courts have become more aggressive in their scrutiny and exclusion of evidence in civil cases. The reader can get a good taste of this from reading the "Notes of Decisions" accompanying chapters on economics, engineering, survey research, epidemiology, toxicology, and medicine in the *Annotated Reference Manual on Scientific Evidence, Second* (Saks et al. 2004).

On the criminal side, the picture is quite different. Risinger found that, post-*Daubert*, in federal district courts defense challenges to government evidence succeeded less than 10% of the time. Government challenges to defense evidence succeeded two thirds of the time. On appeal, defense-proffered expertise was found to have been properly excluded 83% of the time. Prosecution-proffered expertise that had been admitted at trial was excluded only once on appeal. Defendants did somewhat better in state courts than in federal courts, winning a quarter of their challenges. Prosecution challenges to defense expertise succeeded about three quarters of the time.

Groscup and colleagues' (2002) data (criminal cases drawn from federal appellate courts) suggest that patterns of admission and exclusion are unchanged from before *Daubert*, and that this constancy has held true for each category of expert testimony examined. Much of the difference between the conclusions of this study and that of the others can be attributed to the universe of cases on which this study focused. That is, the findings reflect the behavior of the courts in the body of cases examined. As Groscup et al. realized, by focusing on appellate cases they are missing most of the action (or inaction) at the trial court level. The only trial rulings their selection method captures are those that resulted in appeal, meaning cases in which defendants lost *Daubert* challenges and then also lost their trials. However, Groscup et al.'s method allows us to see more precisely how the decisions of appellate courts compare with the decisions of trial courts on the same cases.

“Of course, none of this,” as Risinger notes, “goes directly to the validity of any given decision,” but the data “are fairly striking in their own right” (Risinger 2000b, p. 108). One possible explanation for the differences between civil and criminal cases is that there are meaningful differences in the quality of the science being offered in the different groups of cases and that there are systematic differences between the factual issues that arise in civil and criminal cases. Or, perhaps, the differential outcomes are attributable to differences in the quality of advocacy (borne of differences in resources) in the two realms. However, some commentators suggest that social and political differences easily explain the differential treatment: As a general proposition, judges disfavor civil plaintiffs and criminal defendants and are more likely to rule against them than against their opposites even when presenting equivalent evidence or arguments. A more definitive explanation of the pattern awaits future research.

Meanwhile, one can ponder some tantalizing hints. Risinger’s analyses of decisions by type of expert evidence proffered in criminal cases, as well as other reviews of government science in criminal cases (e.g., Faigman et al. 2005, forensic science chapters), suggest that if *Daubert* gatekeeping were rationally based on the quality of the underlying expert evidence, the exclusion rate pursuant to defense challenges would be higher than it is. The irony is more pronounced in light of data examining the trial evidence in DNA exoneration cases, which find that faulty forensic science is second only to eyewitness errors as the leading cause of erroneous convictions (Saks & Koehler 2005).

What Are Judges Doing When Undertaking *Daubert* Review?

Can we explain the contrast between the success of challenges by civil defendants and lack of success by criminal defendants? Recalling the discussion that accompanied Table 1, for well-developed disciplines the same results should occur whether the filter applied is *Frye* or *Daubert*. No change should occur for most of the evidence proffered in civil cases if *Frye*-like criteria were being employed pre-*Daubert*. Apparently, such criteria were not being employed. Notwithstanding that rules of evidence apply equally to civil and criminal cases (with few exceptions), courts often acted as though *Frye* applied only to criminal cases, whereas many judges understood *Daubert* to apply especially to civil cases. So *Daubert* had much to offer to civil cases, even if logically it should not have. But the cell in Table 1 that would have called for the most dramatic changes in admissibility pre- to post-*Daubert* is the low-validity/high-general-acceptance cell, which is largely populated with weak forensic individualization science. For fields in this category, the move from *Frye* to *Daubert* would (logically) have produced a noticeable (if not dramatic) shift from admission to exclusion. But, as both counts and readings of the cases suggest, it has not. Can this be explained?

A number of studies suggest that judges do not employ *Daubert* as the directive it seems by its terms to be—a directive to conduct meaningful and sincere analyses of the substance of proffered expert evidence, using rational criteria and following

them to their logical destination. Instead, judges have taken *Daubert* to be a vague call to arms against junk science in civil cases while keeping hands off of the government's proffers in criminal cases.

Several studies have looked beneath the surface of mere counts of decisions to admit or exclude to find the basis of the courts' less-than-straightforward treatment of admissibility issues in the post-*Daubert* era. Krafka et al.'s (2002) surveys found that in deciding expert admissibility in civil cases, federal judges were far more likely to rest their appraisal of the evidence on considerations that were not new with *Daubert* (e.g., relevance, qualifications, ability to assist trier of fact) than they were to employ the new *Daubert* factors (testing, quality of underlying research, error rates). Thus, in the wake of *Daubert*, judges were excluding more expert evidence, but doing so using legal doctrines that were available before *Daubert*. The new guidance developed in *Daubert* played little role in these courts' analyses of the expert evidence they were now excluding.

Similar conclusions emerged from a very different research approach: Groscup et al. (2002) analyzed the amount of time (word counts, actually) a large sample of federal criminal appellate opinions devoted to discussing the factors they were employing to evaluate the proffered expert evidence in the case. After *Daubert*, appellate opinions concerning the admissibility of expert evidence spent less time discussing general acceptance and more time discussing Rule 702 and the procedures under which pretrial challenges are conducted, but they did not increase the amount of attention paid to the specific factors suggested by *Daubert* for evaluating scientific evidence. "There was no comparable increase in the discussion of the four *Daubert* criteria in evaluating expert testimony. Discussion of *Daubert* was lengthy, but the discussion devoted to the three new criteria was relatively abbreviated" (Groscup et al. 2002, p. 365).

The limited attention actually paid in judicial opinions to the vaunted (or reviled) *Daubert* factors is less surprising once one realizes that judges do not understand what they mean. Gatowski et al. (2001) surveyed and interviewed a large sample of state court judges. Nearly all the respondents strongly supported the gatekeeper role and nearly two thirds asserted that they were in fact active in making expert evidence admissibility decisions, and asserted high regard for *Daubert's* provision of "the basis. . . for justifying or explaining the decision-making process" and for the usefulness of the specific *Daubert* guidelines. But when asked to define or explain each of the *Daubert* factors, and probed further when they had difficulty, only 5% of the respondents demonstrated a working understanding of falsifiability, and only 4% demonstrated an understanding of error rate. And, when presented with specific examples of expert testimony to evaluate, the criterion relied upon most heavily was general acceptance. The authors summed up their findings by suggesting that the judges' "responses reflected more of the rhetoric of *Daubert* than the substance." Other research suggests that judges have difficulty distinguishing good research designs from poor ones (Kovera & McAuliff 2000).

More traditional legal analysis—that is, close reading of cases—generally supports the quantitative findings summarized above. Whether excluding or admitting

expert evidence, judicial opinions displaying sophisticated application of *Daubert* or other thoughtful focus on the validity of the proffered expertise are few and far between. The major exceptions to this generalization seem to be toxic tort litigation, where judicial sophistication is more evident, and cases involving economic analysis of damages (compare relevant chapters in Faigman et al. 2005). These exceptions suggest that the evolution of judicial sophistication is associated with increasing sophistication of counsel and with growing experience with debate on a subject.

As an example of judicial casualness about validity, the forensic science area is particularly telling. Logically, many of these fields are highly vulnerable under *Daubert* (Thornton & Peterson 2005). But those fields have endured little other than anxiety. Three examples suffice [but for one notable exception, see *U.S. v. Crisp* (2003), dissenting opinion].

First, a review of cases found that although not a single court could cite any systematic empirical evidence supporting critical propositions underlying fingerprint identification, the courts all nevertheless found the proffered testimony regarding fingerprint evidence not only admissible but often worthy of high praise [“the very archetype of reliable expert testimony under [*Daubert*],” (*U.S. v. Havvard* 2000, p. 855)]. That review summed up the body of cases it examined as little more than “a catalog of evasions” of the duty to scrutinize under *Daubert* (Faigman et al. 2005, chapter 27, p. 432). An explanation might be that the field of fingerprint identification has been so effective in its public mythology that courts cannot suspend their belief long enough to examine the real basis of the claims (Cole 2001).

Second, not long after the Kentucky Supreme Court adopted *Daubert* as that state’s standard for testing expert evidence, the same court had occasion to evaluate the admissibility of microscopic hair comparison evidence (*Johnson v. Commonwealth* 1999). Although the court could not cite any studies at all, it nevertheless held the evidence to be fully admissible based on its assumption of general acceptance by past Kentucky cases—even though no prior Kentucky cases had found the evidence to be generally accepted or had even addressed the issue. The *Johnson* court reasoned that silence bespoke general acceptance (Saks 2004).

Third, in the very case by which the Alaska Supreme Court adopted *Daubert* as its admission doctrine (*State v. Coon* 1999), and even though the court remanded to the trial court for the taking of further evidence in anticipation of adopting and applying *Daubert*, no court involved in the case built its holding on a foundation of relevant studies or other data. Voiceprint identification expert evidence was admitted as unthinkingly as ever, even in Alaska’s inaugural *Daubert* decision, and even though for this forensic identification science a good bit of relevant data does exist (Faigman et al. 2005).

In sum, *Daubert* has had somewhat paradoxical effects. Judges overwhelmingly say they subscribe to the gatekeeper role and endorse *Daubert’s* framework for analyzing scientific (and other) expert evidence. It has precipitated a great increase in judicial examination of expert evidence. Yet judges often appear to have little understanding of the basis of the expertise at issue, and all indications are that they

invest little of their scrutiny and decision making in seriously applying *Daubert* or in bringing any other kind of thoughtful examination to bear. Nevertheless, the decisions reveal a pattern of impact: *Daubert* has led to increased exclusion of expert evidence, mostly in civil cases, and most of that excluding plaintiffs' evidence. The questionable sciences of criminal cases, often among the weakest of the scientific evidence that comes to court, are by one device or another usually admitted (or perhaps it is more accurate to say they are granted exemption from serious scrutiny).

Daubert has precipitated a pattern of gatekeeping that is impossible to explain in terms of *Daubert's* doctrinal elements or the relative quality of the underlying science presented for scrutiny. Thus, *Daubert's* impact may have more to do with the sociology of judging than with the law of *Daubert* (Kaye et al. 2004). The future of expert evidence will need to take into account these odd patterns of decision making.

THE FUTURE OF EXPERT EVIDENCE

One can contemplate the future of expert evidence either predictively or prescriptively. The former, of course, is the more challenging and risky.

Predictions

Dramatic national change in the law of expert evidence is unlikely to occur again soon. The Supreme Court rarely changes its mind shortly after making a grand pronouncement. But once the Court becomes aware of problems in the implementation of its earlier rulings, it might adjust the law in ways it thinks will solve those problems. This is especially so if the lower courts split in regard to how they handle certain kinds of evidence. The most likely candidate for high court intervention is clinical medicine, a subject currently dealt with differently in different jurisdictions. Some jurisdictions, for example, allow expert medical opinion with little supporting research (see, e.g., *Heller v. Shaw* 1999), whereas others exclude such opinion pending the completion of sufficient research to support the proffered opinion (see, e.g., *Black v. Food Lion, Inc.* 1999).

The more profound problem in the implementation of *Daubert*, as discussed in the previous section, is the lower courts' apparent inconsistency in applying *Daubert* in civil and criminal cases. If the causes are in any sense sociological or political, and the Court shares in the cultural assumptions and biases that led to that pattern of differential treatment, then the Court may have little desire to alter them.

Although *Daubert* had potential to press various fields toward improvements, including the first serious research they have done on their claims, little of the potential has been realized, and it seems to be declining. Indeed, one scholar has pointed out that once the courts approve shoddy science under the banner of

Daubert, the chances of improvement in those fields, like the chance of judicial re-examination, are less than ever (Berger 2003).

The deferential standard of review announced in *General Electric v. Joiner* cannot survive in the long run. Courts will find ways to fudge, to slow the contradictory or repetitious examinations of the same evidence again and again. Eventually, the Supreme Court can be expected to authorize less deferential review, at least in some classes of cases.

The frequent calls for the increased use of court-appointed witnesses (e.g., Breyer 2000) will go largely unheeded. Although the net use of court-appointed experts and, possibly more so, technical advisers is likely to rise over time, this reform is unlikely to be as transforming as its advocates hope. The power to appoint experts has long been available to the courts, was codified in the Federal Rules of Evidence, and, despite periodic calls for making more use of the power, remains largely unused (Cecil & Willging 1993). This pattern is a testament to the courts' commitment to the adversarial process or to their disinclination to become more managerial. Gross (1991) has offered an impressive analysis of the failure, and continuing failure, of calls for more use of court-appointed experts. And he suggests an alternative for accomplishing much of the good of court appointment while preserving the adversarial imperative: Invite the parties to nominate a certain number of experts to be court-appointed, appoint everyone the parties nominated (who would then understand that their first loyalty is to the court), and then require that all meetings with these experts be open to all parties (and forbid any contact outside of those open meetings). To our knowledge, Gross's suggestions have not yet been followed by any court.

As the fraction of the population of lawyers and judges consisting of people with scientific training slowly grows—due as much as anything else to the advent of more technologies (requiring lawyers to protect or challenge intellectual property claims) or downturns in the market for scientists and engineers—there will be more lawyers and judges who are capable of understanding what *Daubert* is aiming to do and able to see where it has been failing most.

Kumho Tire has potentially deep and demanding implications, as illustrated by the taxonomies of expert evidence that several scholars have started the law thinking about. We suspect that, given the difficulties courts have had in using the essentially ready-made criteria of empirical science handed to them by *Daubert*, the development of nonscience criteria for nonscience fields will be a much steeper and thornier path for the courts to travel.

Prescriptions

In one particular type of case, courts could and might and, we think, should make increased use of experts appointed to serve by (and for the educational benefit of) the court and not by the parties: In consolidated class actions, such as mass toxic torts, the dispute over certain empirical claims is to be resolved in one grand proceeding, rather than in a lengthy stream of individual cases. On such occasions,

courts might feel an unusual obligation to reach the best possible answer. On such occasions they can appoint advisory juries (or similar panels composed of experts) to consider the evidence and offer the court suggested findings. Expert panels have been appointed, for example, in the silicone breast implant litigation (*In re SGBI Litigation* 1997) and the Parlodel litigation (*Soldo v. Sandoz* 2003). The use of such panels should be expanded.

Although there have been efforts to teach judges to become better students of natural and social science as well as statistics, crash courses and checklists will probably not accomplish much. To ensure that courts have judges with scientific acumen, the best method is to recruit scientifically educated lawyers to the legal profession and then to the bench. We note above that more such persons are becoming lawyers. Perhaps the process could be accelerated. Perhaps the judiciary—which employs a great many students of the humanities—already has all the personnel it needs to assess the offerings of nonscience expertises.

Judicial gatekeeping has unavoidable effects on the creation of new knowledge. We believe the courts should act in ways that promote the growth of knowledge that is important to resolving major or frequent disputes that come before the courts.

For example, some fields will do no more research than is required of them. If the courts set a low threshold of admission, some fields will develop little or no fundamental new knowledge. They can remain in business with what they already have and, indeed, risk setting themselves back in the eyes of courts by producing real data that can never show them to be as flawless as they have long claimed themselves to be. For these fields, most often seen on the criminal side of the docket, the courts should set higher thresholds, or set time limits (a period of years) for the production of research on fundamental questions about the field. If nothing else, courts should require parties to remain within the bounds of the knowledge they have, forbidding wishful exaggerations, and requiring statements of the limits of what is known, whether those statements are informed by data showing error rates or by the absence of data on error rates. A court could ask parties for briefs on these matters and issue its own instruction to the jury on the limits of expertise.

On the civil side, a similar problem of ignorance-is-bliss exists, but it requires a different solution. A manufacturer, such as a pharmaceutical company, has no inherent incentive to test a product for safety or effectiveness. That is why regulations sometimes exist to compel such testing. Once the product is approved and is in the market, then the less the company learns about it the better. When plaintiffs begin to suspect harmfulness, under *Daubert* their claims will often die with the pretrial *Daubert* hearing because the limited evidence will mean that their experts cannot even testify at trial. (In the past, such suits would sometimes get to juries and, under that scenario, manufacturers had an incentive to conduct additional research in the hope of acquiring evidence with which to defend against the claim.) Thus, *Daubert* ironically acts as a disincentive to improve the body of scientific knowledge about products. To ameliorate this problem, one procedural device might be considered: When one party has a substantially greater ability to collect data about a matter, the burden of producing needed evidence could be

placed on that party. In its absence, the other side's experts could be allowed to testify to the limited knowledge that does exist, across some lowered threshold of admission.

Finally, we urge that the serious study of expert evidence by social scientists continue and expand. To understand the interaction of experts and the courts, more research is needed.

**The Annual Review of Law and Social Science is online at
<http://lawsocsci.annualreviews.org>**

LITERATURE CITED

- Am. Eff. Law Enforc., et al. 1997. Brief Amici Curiae, submitted in *Kumho Tire Company, Ltd. v. Carmichael*, 526 U.S. 137 (1999)
- Assoc. Trial Lawyers Am. 2004. *Annual Convention Reference Materials*, Vol. 2, *Products Liability*. Washington, DC: Assoc. Trial Lawyers Am.
- Berger MA. 2003. Expert testimony in criminal proceedings: questions *Daubert* does not answer. *Seton Hall Law Rev.* 34:1125–40
- Bernstein DE, Jackson JD. 2004. The *Daubert* trilogy in the states. *Jurimetr. J.* 44:351–66
- Black B. 1988. A unified theory of scientific evidence. *Fordham Law Rev.* 56:595–695
- Black v. Food Lion, Inc.*, 171 F.3d 308 (Fifth Cir. 1999)
- Breyer S. 2000. Science in the courtroom. *Issues Sci. Technol.* 2000:52–56
- Carlson RL. 1992. Experts as hearsay conduits: confrontation abuses in opinion testimony. *Minn. Law Rev.* 76:859–75
- Cecil JS, Willging TE. 1993. *Court-Appointed Experts: Defining the Role of Experts Appointed Under Federal Rule of Evidence 706*. Washington, DC: Fed. Judic. Cent.
- Cobbett W. 1810. The Trial of Philip, Earl of Pembroke and Montgomery, at Westminster, for the murder of Nathaniel Cony (1678). In *Cobbett's Complete Collection of State Trials*, 6:1310–50. London: R. Bagshaw
- Cole S. 1992. *Making Science: Between Nature and Society*. Cambridge, MA: Harvard Univ. Press
- Cole SA. 2001. *Suspect Identities: A History of Fingerprinting and Criminal Identification*. Cambridge, MA: Harvard Univ. Press
- Crump D. 2003. The trouble with *Daubert-Kumho*: reconsidering the Supreme Court's philosophy of science. *Missouri Law Rev.* 68:1–42
- Daubert v. Merrell Dow Pharmaceuticals, Inc.*, 509 U.S. 579 (1993)
- Dixon L, Gill B. 2001. *Changes in the Standards for Admitting Expert Evidence in Federal Civil Cases since the Daubert Decision*. Santa Monica, CA: RAND
- Dunagin v. City of Oxford*, 718 F.2d 738 (5th Cir. 1983)
- Edmond G, Mercer D. 2002. Conjectures and exhumations: citations of history, philosophy and sociology of science in U.S. federal courts. *Law Lit.* 14:309–52
- Evelt IW. 1993. Criminalistics: the future of expertise. *J. Forensic Sci. Soc.* 33:173–78
- Faigman DL. 1989. To have and have not: assessing the value of social science to the law as science and policy. *Emory Law J.* 38:1005–95
- Faigman DL. 1999. *Legal Alchemy: The Use and Misuse of Science in the Law*. New York: Freeman
- Faigman DL. 2002. Is science different for lawyers? *Science* 197:339–40
- Faigman DL. 2004. *Laboratory of Justice: The Supreme Court's 200-Year Struggle to Integrate Science and the Law*. New York: Henry Holt (Times Books)
- Faigman DL, Kaye DH, Saks MJ, Sanders J. 2005. *Modern Scientific Evidence: The*

- Law and Science of Expert Testimony*. Minneapolis, MN: West/Thompson. 2nd ed. Republished
- Faigman DL, Porter E, Saks MJ. 1994. Check your crystal ball at the courthouse door, please: exploring the past, understanding the present, and worrying about the future of scientific evidence. *Cardozo Law Rev.* 15:1799–835
- Folkes v. Chadd*, 3 Dougl. 157, 99 Eng. Rep. 589 (K.B. 1782)
- Frye v. U.S.*, 293 F. 1013 (D.C. Cir. 1923)
- Gatowski SI, Dobbin SA, Richardson JT, Ginsburg GP, Merlino ML, Dahir V. 2001. Asking the gatekeepers: a national survey of judges on judging expert evidence in a post-*Daubert* world. *Law Hum. Behav.* 25:433–58
- General Electric v. Joiner*, 522 U.S. 136 (1997)
- Giannelli PC. 1980. The admissibility of novel scientific evidence: *Frye v. United States*, a half-century later. *Columbia Law Rev.* 80:1197–250
- Golan T. 2004. *Laws of Men and Laws of Nature: The History of Scientific Expert Testimony in England and America*. Cambridge, MA: Harvard Univ. Press
- Goodstein D. 2000. How science works. In *Reference Manual on Scientific Evidence, Second*, pp. 67–82. Washington, DC: Fed. Judic. Cent.
- Groscup JL, Penrod SD, Studebaker CA, Huss MT, O'Neil KM. 2002. The effects of *Daubert* on the admissibility of expert testimony in state and federal criminal cases. *Psychol. Public Policy Law* 8:339–72
- Gross S. 1991. Expert evidence. *Wis. Law Rev.* 1991:1113–232
- Gross SR, Mnookin J. 2003. Expert information and expert evidence: a preliminary taxonomy. *Seton Hall Law Rev.* 34:141–89
- Haack S. 2003a. *Defending Science—Within Reason: Between Scientism and Cynicism*. New York: Prometheus
- Haack S. 2003b. Inquiry and advocacy, fallibilism and finality: culture and inference in science and the law. *Law Probab. Risk* 2:205–14
- Haber S. 1991. *The Quest for Authority and Honor in the American Professions, 1750–1900*. Chicago: Univ. Chicago Press
- Heller v. Shaw*, 167 F.3d 146 (3d Cir. 1999)
- Horrobin DF. 1990. The philosophical basis of peer review and the suppression of innovation. *JAMA* 263:1438–41
- In re Silicone Gel Breast Implant Product Liability Litigation*, 996 F. Supp. 1110 (N.D. Ala. 1997)
- Jasanoff S. 1992. What judges should know about the sociology of science. *Jurimetr. J.* 32:345–59
- Jasanoff S. 1996. *Science at the Bar: Law, Science, and Technology in America*. Cambridge, MA: Harvard Univ. Press
- Johnson v. Commonwealth*, 12 S.W. 3d 258 (1999)
- Jonakait RN. 1997. The assessment of expertise: transcending construction. *Santa Clara Law Rev.* 37:301–47
- Kaye D, Bernstein D, Mnookin J. 2004. *The New Wigmore, A Treatise on Evidence: Expert Evidence*. New York: Aspen
- Kovera MB, McAuliff BD. 2000. The effects of peer review and evidence quality on judge evaluations of psychological science: Are judges effective gatekeepers? *J. Appl. Psychol.* 85:574–86
- Krafka C, Meghan A, Dunn MA, Johnson MT, Cecil JS, Miletich D. 2002. Judge and attorney experiences, practices, and concerns regarding expert testimony in federal civil trials. *Psychol. Public Policy Law* 8:309–32
- Kuhn T. 1996. *The Structure of Scientific Revolutions*. Chicago: Univ. Chicago Press. 3rd ed.
- Kumho Tire Ltd. v. Carmichael*, 526 U.S. 137 (1999)
- Leiter B. 1997. The epistemology of admissibility: why even good philosophy of science would not make for good philosophy of evidence. *BYU Law Rev.* 1197:803–19
- Lewontin RC. 1994. Facts and the factitious. In *Questions of Evidence: Proof, Practice, and Persuasion Across the Disciplines*, ed. J Chandler, A Davidson, H Harootunian, pp. 478–91. Chicago: Univ. Chicago Press

- Maletskos CJ, Spielman SJ. 1967. Introduction of new scientific methods in court. In *Law Enforcement, Science & Technology*, ed. SA Yefsky, pp. 957–64. Washington, DC: Thompson
- Monahan J, Walker L. 1986. Social authority: obtaining, evaluating, and establishing social science in law. *Univ. Penn. Law Rev.* 134:477–517
- Natl. Inst. Justice. 2000. *Solicitation for Forensic Friction Ridge (Fingerprint) Examination Validation Studies*. Washington, DC: US Dep. Justice
- O'Connor S. 1995. The Supreme Court's philosophy of science: Will the real Karl Popper please stand up? *Jurimetr. J.* 35:263–76
- Pavalko RM. 1988. *Sociology of Occupations and Professions*. Itasca, IL: Peacock. 2nd ed.
- Popper K. 1989. *Conjectures and Refutations: The Growth of Scientific Knowledge*. New York: Basic Books. 5th ed.
- Risinger DM. 2000a. Defining the "task at hand": non-science forensic science after *Kumho Tire v. Carmichael*. *Wash. Lee Law Rev.* 57:767–800
- Risinger DM. 2000b. Navigating expert reliability: Are criminal standards of certainty being left on the dock? *Albany Law Rev.* 64:99–152
- Risinger DM. 2005. Preliminary thoughts on a functional taxonomy of expertise for the post-*Kumho* world. See Faigman et al. 2005, 1:69–93
- Risinger DM, Saks MJ, Rosenthal R, Thompson WC. 2002. The *Daubert/Kumho* implications of observer effects in forensic science: hidden problems of expectation and suggestion. *Calif. Law Rev.* 90:1–56
- Saks MJ. 1998. Merlin and Solomon: lessons from the law's formative encounters with forensic identification science. *Hastings Law J.* 49:1069–141
- Saks MJ. 2004. *Johnson v. Commonwealth*: How dependable is identification by microscopic hair comparison. *Advocate [J. Crim. Justice Educ. Res.]* 26:14–23
- Saks MJ, Faigman DL, Kaye D, Sanders J. 2004. *Annotated Reference Manual on Scientific Evidence, Second*. Minneapolis, MN: West
- Saks MJ, Koehler JJ. 1991. What DNA "fingerprinting" can teach the law about the rest of forensic science. *Cardozo Law Rev.* 13:361–72
- Saks MJ, Koehler JJ. 2005. The coming paradigm shift in forensic identification science. *Science.* 309:892–95
- Sanders J. 2001. *Kumho* and how we know. *Law Contemp. Probl.* 64:373–415
- Sanders J, Diamond SS, Vidmar N. 2002. Legal perceptions of science and expert knowledge. *Psychol. Public Policy Law* 2:139–53
- Seton Hall Symposium. 2003. Expert admissibility symposium: What is the question to which standards of reliability are to be applied? *Seton Hall Law Rev.* 34:1–388
- Shuman DW, Sales BD. 1999. The impact of *Daubert* and its progeny on the admissibility of behavioral and social science evidence. *Psychol. Public Policy Law* 5:3–15
- Soldo v. Sandoz Pharmaceuticals Corp.*, 244 F. Supp. 2d 434, 534–36 (W.D.Pa. 2003)
- State v. Coon*, 974 P. 2d 386 (Alaska 1999)
- Thornton J, Peterson J. 2005. The general assumptions and rationale of forensic identification. See Faigman et al. 2005, 3:157–222
- U.S. v. Crisp*, 324 F.3d 261 (4th Cir. 2003)
- U.S. v. Havvard*, 117 F. Supp. 2d 848, 849 (S.D. Ind. 2000), aff'd, 260 F.3d 597 (7th Cir. 2001)
- Walker L, Monahan J. 1987. Social frameworks: a new use of social science in law. *Va. Law Rev.* 73:559–98
- Walker L, Monahan J. 1988. Social facts: scientific methodology as legal precedent. *Calif. Law Rev.* 76:877–96
- Weisgram v. Marley Co.*, 528 U.S. 440 (2000)

CONTENTS

COMING OF AGE: LAW AND SOCIETY ENTERS AN EXCLUSIVE CLUB, <i>Lawrence M. Friedman</i>	1
THE COMPARATIVE STUDY OF CRIMINAL PUNISHMENT, <i>James Q. Whitman</i>	17
ECONOMIC THEORIES OF SETTLEMENT BARGAINING, <i>Andrew F. Daughety and Jennifer F. Reinganum</i>	35
LAW AND CORPORATE GOVERNANCE, <i>Neil Fligstein and Jennifer Choo</i>	61
TRANSNATIONAL HUMAN RIGHTS: EXPLORING THE PERSISTENCE AND GLOBALIZATION OF HUMAN RIGHTS, <i>Heinz Klug</i>	85
EXPERT EVIDENCE AFTER <i>DAUBERT</i> , <i>Michael J. Saks</i> <i>and David L. Faigman</i>	105
PLEA BARGAINING AND THE ECLIPSE OF THE JURY, <i>Bruce P. Smith</i>	131
THE DEATH PENALTY MEETS SOCIAL SCIENCE: DETERRENCE AND JURY BEHAVIOR UNDER NEW SCRUTINY, <i>Robert Weisberg</i>	151
VOICE, CONTROL, AND BELONGING: THE DOUBLE-EDGED SWORD OF PROCEDURAL FAIRNESS, <i>Robert J. MacCoun</i>	171
LAW, RACE, AND EDUCATION IN THE UNITED STATES, <i>Samuel R. Lucas</i> <i>and Marcel Paret</i>	203
LAW FACTS, <i>Arthur L. Stinchcombe</i>	233
REAL JURIES, <i>Shari Seidman Diamond and Mary R. Rose</i>	255
FEMINISM, FAIRNESS, AND WELFARE: AN INVITATION TO FEMINIST LAW AND ECONOMICS, <i>Gillian K. Hadfield</i>	285
CRIMINAL DISENFRANCHISEMENT, <i>Christopher Uggen, Angela Behrens,</i> <i>and Jeff Manza</i>	307
AFTER LEGAL CONSCIOUSNESS, <i>Susan S. Silbey</i>	323
WHY LAW, ECONOMICS, AND ORGANIZATION? <i>Oliver E. Williamson</i>	369
REVERSAL OF FORTUNE: THE RESURGENCE OF INDIVIDUAL RISK ASSESSMENT IN CRIMINAL JUSTICE, <i>Jonathan Simon</i>	397
INDEX	
Subject Index	423