FROM THE EDITOR

Theodore Benditt
University of Alabama at Birmingham, AL

The topic for the current issue of the Newsletter is Science in the Law. I want to thank Professor Susan Haack (Philosophy and Law, University of Miami) for taking on the role of guest editor for this issue. As Professor Haack indicates in her introduction, expert testimony becomes more prominent in legal controversies, there has been considerable interest in sharpening the criteria for determining when expert (and especially scientific) testimony is allowable and when proffered expert testimony should be excluded. Recent Supreme Court decisions have not necessarily clarified the issue, leaving trial judges to make difficult judgments, with little guidance, for which they may not be well qualified. Yet the outcomes of important and massive lawsuits can rest on what expert testimony jurors are permitted to hear.

As always, the Newsletter also contains a number of abstracts of recent articles from law reviews and other journals of interest to philosophers. I want to thank Professors Julie van Camp (Philosophy, Cal State-Long Beach) and Sarah Holtman (Philosophy, University of Minnesota, Twin Cities) for contributing abstracts for this issue.

FROM THE GUEST EDITOR

Science in the Law: Introduction
Susan Haack
University of Miami, Miami, FL

Science and the law are often intimately entangled. In law-making, scientists are asked to gauge the effects, costs, and benefits of this or that policy; in adjudication, they are asked to testify about the causes of this disease, the identity of this semen sample, the possibility of hypnotically refreshed memory, the victim's lost earning power, etc., etc. Scientific testimony has come to play a significant role in the pre- and post-regulation of risk: in environmental law, occupational health and safety law, and torts; in cases involving equal protection and discrimination; in intellectual property, patent, and anti-trust law; in mental-health law and sexual predator law; in creation—"science" cases and in disputes over whether ancient human remains may be given to anthropologists to determine their origin, or must be given to the Native American tribes that claim them for burial; and, of course, in criminal cases, where the forensic sciences of identification and causation loom large.

According to one estimate, by 1990 around 70% of cases in the U.S. involved the testimony of experts, many of them scientists: economists and epidemiologists, toxicologists and tool-mark examiners, psychologists and serologists, experts on PCBs and experts on paternity, experts on rape trauma syndrome and experts on respiratory disorders, experts on blood, on bullets, on battered women, etc., etc. But, as the titles of books like Steven Goldberg's Culture Clash: Law and Science in America and David Faigman's Legal Alchemy: The Use and Misuse of Science in the Law remind us, the interactions of science and the law have sometimes been uneasy. In particular, while it can be a powerful tool for justice, scientific testimony can also be a powerful source of uncertainty and confusion, not to mention of opportunities for—well, for opportunism. As the role of scientific testimony has grown, so too have concerns that flimsy scientific evidence, "junk science," and pseudo-scientific gobbledegook are finding their way into the courts.

In 1993 the Supreme Court issued its landmark ruling in Daubert v. Merrell Dow Pharmaceuticals, the first case in its 204-year history directly concerned with the criteria of admissibility of scientific testimony; and two more key expert-testimony cases followed in quick succession: G.E. v. Joiner (1997) and Kumho Tire v. Carmichael (1999). In this issue of the Newsletter—serendipitously timed just ten years after Daubert—contributors take up some of the issues in epistemology, philosophy of science, and public policy raised by Daubert, and the legal factfinding process more generally.

I have put my paper first because it begins with a brief summary history of the law on the admissibility of expert,
including scientific testimony. In its Daubert ruling, I argue, the Supreme Court ran together two incompatible philosophies of science—Popper’s and Hempel’s—neither of which is adequate to the task demanded of it; and, in the process, revealed a deep misunderstanding of the sciences and their place in inquiry generally. And while in its subsequent rulings on expert testimony in Joiner and Kumho the Supreme Court seems to have backed quietly away from the muddled philosophy of science built into Daubert, it has left federal judges with large responsibilities and broad discretion in keeping the gate against unreliable expert testimony, but very little guidance about how they are to perform this demanding task.

Dale Nance’s paper, also focused on Daubert’s reliability requirement, develops a different theme. Reliability, as Nance points out, is usually thought of as a dichotomous attribute of evidence, as if expertise were either reliable or unreliable. But, he argues, to make progress in developing meaningful and appropriate restrictions on the admissibility of expert testimony requires us to abandon this idea, and understand the implications of endorsing a gradational conception of reliability, according to which evidence can be more or less reliable, and which gives a prominent place to comparative assessments of reliability. Nance recommends that, consistent with Supreme Court precedent and empirical evidence about jury decision-making, a court’s reliability inquiry should focus on whether more-reliable evidence is reasonably available to the proponent.

Carl Cranor’s paper, again focused on Daubert but now from a broadly policy-oriented angle, explores and develops an idea of Charles Nesson’s about the difference between the acceptability of verdicts that convey a liability judgment, and of verdicts that tend to focus on the evidence. Daubert, Cranor argues, makes some contribution to the acceptability of a verdict, by attempting to ensure that jury decisions, whether based on plaintiffs’ or defendants’ experts’ accounts of the science involved, will be within the bounds of the scientifically acceptable. But, he continues, there are several disadvantages to the heightened pre-trial scrutiny of proffered evidence brought by Daubert: it tends to reduce plaintiffs’ access to the law, and to deprive the community of structured public discussion of the issues; it tempts defendants to distort scientific evidence, to dismiss legitimate scientific evidence, and perhaps (given Daubert’s stress on peer-reviewed publication) to distort the scientific literature; and it increases their incentive to refrain from conducting studies that might provide warning of products’ adverse effects.

Vern Walker’s paper completing the issue offers a broad perspective on the interaction of epistemic and other desiderata in legal factfinding, the larger territory within which issues about scientific testimony in general, and Daubert in particular, are to be located. Factfinding in a governmental institution, he observes, is a highly structured process usually involving three decision-making roles. Each of these three principal roles (factfinder, presiding authority, and reviewing authority) has a significant zone of discretion in which to operate, yet decisions within that zone are constrained by decisions within the other two. The process as a whole is designed to balance epistemic and non-epistemic objectives. The strategy is to maintain a dynamic process of rule-governed decision-making, through which (it is hoped) reasonable decision-makers will come close enough to achieving the epistemic objective over time. There is, he concludes, a great deal of theoretical work to be done in deciding the proper balance for each type of decision, by each type of decision-maker, and for each type of propositional issue.

**Endnotes**

*My thanks to the contributors for supplying the abstracts of their papers that I have woven into the introduction; and to Vern Walker, additionally, for his help in turning my first, chaotic draft of the opening paragraphs into something at least a bit more orderly.

**Articles**

**Disentangling Daubert: An Epistemological Study in Theory and Practice**

Susan Haack

University of Miami

Sometimes the word [“science”] degenerates into a vague honorific, synonymous with the advertiser’s “reliable” or “guaranteed”... [Jacques Barzun]¹

In *Frye v. United States*, the D.C. Court upheld the exclusion of testimony of the results of a then-new blood-pressure deception test on the grounds that novel scientific testimony “crosses the line between the experimental and the demonstrable,” and so is admissible, only if it is “sufficiently established to have gained general acceptance in the particular field to which it belongs.” Ignored for a decade, rarely cited for a quarter-century, over time the “Frye test” became increasingly influential, until by the early 1980s it had been adopted by 29 states.

In 1975, however, newly-enacted Federal Rules of Evidence (FRE) had set a seemingly less restrictive standard: the testimony of a qualified expert, including a scientific expert, is admissible provided it is relevant (unless it is excluded, under Rule 403, on grounds of unfair prejudice, waste of time, or confusing or misleading the jury). 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 to hear psychiatric testimony of his future dangerousness at the sentencing hearing—even though an amicus brief from the American Psychiatric Association reported that 2 out of 3 such predictions are mistaken. Writing for the majority, Justice White observed that state and federal rules of evidence “anticipate that relevant, unprivileged testimony should be admitted and its weight left to the fact-finder, who would have the benefit of cross-examination and contrary evidence by the opposing party.”² Justice Blackmun wrote an angry dissent.

In 1991, amid increasing public concern that the tort system was getting out of hand, Peter Huber argued in his influential *Galileo’s Revenge* that under the Federal Rules worthless “junk science,” which would have been excluded by the *Frye* test, was flooding the courts. In 1992 proposals to tighten up the Federal Rules were before Congress. In 1993 the Supreme Court issued its ruling in *Daubert*—the first case in its 204-year history where the central questions concerned the admissibility of scientific testimony. The *Frye* rule arose in a criminal case, and had for most of its history been cited in criminal cases; but *Daubert* was a tort action in which the trial court had relied on *Frye* in excluding the plaintiffs’ experts’ testimony that the morning-sickness drug Bendectin was teratogenic. So the Supreme Court was to determine whether the FRE had superseded *Frye*, and in particular how Rule 702 was to be interpreted.

Yes, Justice Blackmun wrote for the majority, the FRE *had* superseded *Frye*; but the Rules themselves require judges to screen proffered expert testimony not only for relevance, but also for reliability. In doing this courts must look, not to an expert’s conclusions, but to his “methodology,” to determine whether proffered evidence is really “scientific...knowledge,” and hence reliable. As to what that methodology is, citing an article by law
professor Michael Green citing Karl Popper, and quoting an observation of Carl Hempel's for good measure, the *Daubert* ruling suggests four factors that courts might use in assessing reliability: “falsifiability,” i.e., whether 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 community.5

In dissent, however, pointing out that the word “reliable” nowhere occurs in the text of Rule 702, Justice Rehnquist anticipated difficulties over whether, and if so, how, *Daubert* should be applied to non-scientific expert testimony; worried aloud that federal judges were being asked to become amateur scientists; and questioned the wisdom of his colleagues’ readiness to get involved in philosophy of science. I think he was right to suspect that something was seriously amiss; in fact, what I shall have to say here might be read as an exploration, amplification, and partial defense of his reservations about that philosophical excursus.

---

Apparentely equating the question of whether expert testimony is reliable with the question of whether it is genuinely scientific, taking for granted that there is some scientific “methodology” which, faithfully followed, guarantees reliable results, and casting about for a philosophy of science to fit this demanding bill, the *Daubert* Court settled on an unstable amalgam of Popper’s and Hempel’s very different approaches—neither of which, however, is suitable to the task at hand.

Popper describes his philosophy of science as “Falsificationist,” by contrast with the Verificationism of the Logical Positivists, because his key theme is that scientific statements can never be shown conclusively to be true, but can sometimes be shown conclusively to be false. Hence his criterion of demarcation: to be genuinely scientific, a statement must be “testable”—meaning, in Popper’s mouth, “refutable” or “falsifiable,” i.e., susceptible to evidence that could potentially show it to be false (if it is false). Curiously, Popper acknowledged from the beginning that his criterion of demarcation is a “convention”; and in 1959, in his Introduction to the English edition of *The Logic of Scientific Discovery*, affirmed that scientific knowledge is continuous with common-sense knowledge.6 Nevertheless, his whole philosophy of science turns on his criterion of demarcation. Falsifiability is to discriminate real empirical science, such as Einstein’s theory of relativity, from pre-scientific myths, from non-empirical disciplines like pure mathematics or metaphysics, from non-scientific disciplines like history, and from such pseudo-sciences as Freud’s and Adler’s psychoanalytic theories and Marx’s “scientific socialism.”7 Falsifiability is also central to Popper’s account of the method of science as “conjecture and refutation”: making a bold, highly falsifiable guess, testing it as severely as possible, and, if it is found to be false, giving it up and starting over rather than protecting it by *ad hoc* or “conventionalist” modifications. (This readiness to accept falsification and eschew *ad hoc* stratagems is Popper’s “methodological criterion” of the genuinely scientific.)

Popper also describes his philosophy of science as “Deductivist,” by contrast with “Inductivism,” whether in the strong, Baconian form that posits an inductive logic for arriving at hypotheses or in the weaker, Logical Positivist form that posits an inductive logic of confirmation. According to Popper, Hume showed long ago that induction is unjustifiable. But science doesn’t need induction; the method of conjecture and refutation requires only deductive logic—specifically, *modus tollens*, the rule invoked when an observational result predicted by a theory fails.

Theories which have been tested but not yet falsified are “corroborated,” degree of corroboration at a time depending on the number and severity of the tests passed. That a theory is corroborated, to however high a degree, doesn’t show that it is true, or even probable; indeed, the degree of testability of a hypothesis is *inversely related* to its degree of logical probability.8 Corroboration is not a measure of verisimilitude, but at best an indicator of how the verisimilitude of a theory *appears*, relative to other theories, at a time;9 and that a theory is corroborated doesn’t mean that it is rational to believe it. (It *does* mean, Popper writes, that it is rational to prefer the theory as the basis for practical action; not, however, that there are good reasons for thinking the theory will be successful in future—*there can be no good reasons for believing this.*10 So it seems that all this “concession” amounts to is that in deciding how to act we can do no better than go with theories we don’t so far know to be false.)

The first problem with the *Daubert* Court’s reliance on Popper is that applying his criterion of demarcation is no trivial matter; as Justice Rehnquist pointed out, observing wryly that, since he didn’t really know what is meant by saying that a theory is “falsifiable,” he doubted federal judges would, either.11 Indeed, Popper himself doesn’t seem quite sure how to apply his criterion. Sometimes, for example, he says that the theory of evolution is not falsifiable, and so is not science; at one point he suggests that “survival of the fittest” is a tautology, or “near-tautology,” and elsewhere that evolution is really a historical theory, or perhaps metaphysics. Then he changes his mind: evolution is science, after all.12 It’s ironic; for Popper’s criterion of demarcation had already found its way into the U.S. legal system a decade before *Daubert*, in a 1982 First Amendment case: *McLean v. Arkansas Board of Education*, where Michael Ruse’s testimony that creation science is not science, by Popper’s criterion, but the theory of evolution is, apparently persuaded Judge Overton.13

But there is an even more serious problem with the *Daubert* Court’s reliance on Popper, of which Justice Rehnquist doesn’t seem aware: Popper’s philosophy of science is signally inappropriate to the Court’s concern with reliability. When Popper describes his approach as “Critical Rationalism,” it is to emphasize that the rationality of the scientific enterprise lies in the susceptibility of scientific theories to criticism, i.e., to testing, and potentially to falsification, not in their verifiability or confirmability. True, early on, Carnap translated Popper’s word “Bewährung” by “confirmation”; and for a while, thinking the issue merely verbal, Popper let it go—even, occasionally, using “confirm” himself. But in a footnote to the English edition of *The Logic of Scientific Discovery* he comments that this had been a bad mistake on his part, conveying the false impression that a theory’s having been corroborated means that it is probably true.14 Except for the weak moments when he conditioned Carnap’s (mis)translation, Popper insisted that corroboration must not be confused with confirmation. The degree of corroboration of a theory represents its past performance only, and “says nothing what-ever about future performance, or about the ‘reliability’ of a theory”; even the best-tested theory “is not ‘reliable’”—so scornful is Popper of the concept of reliability that he refuses even to use the word without putting it in precautionary scare quotes! Reiterating that he puts the emphasis “on negative arguments, such as negative instances or counter-examples, refutations, and attempted refutations—in short, criticism—while the inductivist lays stress on ‘positive instances’, from which he draws ‘non-demonstrative inferences’, and which he hopes will guarantee the ‘reliability’ of the conclusions of these inferences,” Popper specifically identifies Hempel as representative of those inductivists with whom he disagrees.15

---
Hempel is not, perhaps, the prototypical inductivist; he describes the method of science as “hypothetico-deductive”; he affirms that scientific claims should be subject to empirical check or testing; and he doesn’t follow Reichenbach and Carnap in explaining confirmation by appeal to the calculus of probabilities. Nevertheless, Popper is surely right to see Hempel’s approach as very significantly at odds with his own: Hempel is not centrally concerned with demarcating science; he questions the supposed asymmetry between verification and falsification, and argues that Popper’s criterion “involves a very severe restriction of the possible forms of scientific hypotheses,” e.g., in ruling out purely existential statements; when he speaks of “testing” he envisages both disconfirmation and confirmation of a hypothesis; and one of his chief projects was to articulate the “logic of confirmation,” i.e., of the support of general hypotheses by positive instances.

Apparently the Supreme Court hoped, by combining Hempel’s account of confirmation with Popper’s criterion of demarcation, to craft a crisp test to identify genuine, and hence reliable, science. But, though Hempel’s philosophy of science is more positive than Popper’s, it isn’t much more help with the question of reliability. For one thing, the confirmation of generalizations by positive instances which preoccupies Hempel is just too simplified to apply to the enormously complex congeries of epidemiological, toxicological, etc., etc., evidence at stake in a case like Daubert. For another, Hempel himself seems eventually to have concluded (rightly, I believe) that the “grue” paradox shows that confirmation isn’t a purely syntactic or logical notion after all, and late in life began to think that maybe Kuhn had been on the right track.

But the most fundamental problem is that what Hempel offered was an account of supportiveness of evidence, or as he said, of “relative confirmation,” the relation between observational evidence and hypothesis, expressible as “E confirms H [to degree n],” or “H is confirmed [to degree n] by evidence E.” This, as Hempel acknowledged, falls short of an account of “absolute confirmation,” the warrant of a scientific claim, which would be expressed in non-relative terms, as “H is confirmed [to degree n], period.” To discriminate reliable testimony from unreliable, however, would require an account of the non-relative concept—which Hempel doesn’t supply.

So, the Daubert Court mixes up its Hoppers and its Pempels; but isn’t this just a slip, of merely scholarly interest? No: it is symptomatic of the serious misunderstanding of the place of the sciences within inquiry generally revealed by the Court’s equation of “scientific” and “reliable.”

So successful have the natural sciences been that the words “science,” “scientific,” and “scientifically” are often used as generic terms of epistemological praise, meaning vaguely “strong, reliable, good”—as in television advertisements, actors in white coats urge viewers to get their clothes cleaner with “strong, reliable, good”—as, in television advertisements, actors in white coats urge viewers to get their clothes cleaner with “strong, reliable, good.” Wizzo. This honorific usage is unmistakably at work in the Daubert ruling; indeed, it seems to be implicit even in the way Justice Blackmun writes of “scientific…knowledge,” strategically excising a significant phrase from the reference in FRE 702 to “scientific or other technical knowledge,” and apparently signalling an expectation that a criterion of the genuinely scientific will also discriminate reliable testimony from unreliable.

If “scientific” is used honorifically, it is a tautology that “scientific” = “reliable”; but this tautology, obviously, is of no help to a judge trying to screen proffered scientific testimony. If “scientific” is used descriptively, however, “scientific” and “reliable” come apart: for, obviously, physicists, chemists, biologists, medical scientists, etc., are sometimes incompetent, confused, self-deceived, dishonest, or simply mistaken, while historians, detectives, investigative journalists, legal and literary scholars, plumbers, auto mechanics, etc., are sometimes good investigators. In short, not all, and not only, scientists are reliable inquirers; and not all, and not only, scientific evidence is reliable. Nor is there a “scientific method” in the sense the Court assumed: no uniquely rational mode of inference or procedure of inquiry used by all scientists and only by scientists. Rather, as Einstein once put it, scientific inquiry is “nothing but a refinement of our everyday thinking.” Superimposing on the inferences, desiderata, and constraints common to all serious investigation a vast variety of constantly evolving local ways and means of stretching the imagination, amplifying reasoning power, extending evidential reach, and stiffening respect for evidence.

Every kind of empirical inquiry, from the simplest everyday puzzling over the causes of delayed buses or spoiled food to the most complex investigations of detectives, of historians, of legal and literary scholars, and of scientists, involves making an informed guess about the explanation of some event or phenomenon, figuring out the consequences of its being true, and checking how well those consequences stand up to evidence. This is the procedure of all scientists; but it is not the procedure only of scientists. Something like the “hypothetico-deductive method,” really is the core of all inquiry, scientific inquiry included. But it is not distinctive of scientific inquiry; and the fact that scientists, like inquirers of every kind, proceed in this way tells us nothing substantive about whether or when their testimony is reliable.

The sciences have extended the senses with specialized instruments; stretched the imagination with metaphors, analogies, and models; amplified reasoning power with numerals, the calculus, computers; and evolved a social organization that enables cooperation, competition, and evidence-sharing, allowing each scientist to take up his investigation where others left off. Astronomers devise ever more sophisticated telescopes, chemists ever more sophisticated techniques of analysis, medical scientists ever more sophisticated methods of imaging bodily states and processes, and so on; scientists work out what controls are needed to block a potential source of experimental error, what statistical techniques to rule out a merely coincidental correlation, and so forth. But these scientific “helps” to inquiry are local and evolving, not used by all scientists.

You may object that, since I have acknowledged that scientific inquiry is continuous with everyday empirical inquiry, I have in effect agreed with Popper that science is an extension of common sense. Indeed, I think science is well-described, in Gustav Bergmann’s wonderfully evocative phrase, as the Long Arm of Common Sense. But the continuity is not between the content of scientific and of common-sense knowledge, but between the basic ways and means of everyday and of scientific inquiry; and it is precisely because of this continuity that the Popperian preoccupation with the “problem of demarcation” is a distraction.

Or you may object that the Daubert Court’s Popperian advice that courts ask whether proffered scientific testimony “can be and has been tested” surely is potentially helpful. This is true; but it is no real objection. “Check whether proffered testimony has been tested” is very good advice when a purported expert hasn’t made even the most elementary effort to check how
well his claims stand up to evidence: such as the knife-mark examiner in Ramirez, who testified that he could infallibly identify this knife, to the exclusion of all other knives in the world, as having made the wound—though no study had established the assumed uniqueness of individual knives, and his purported ability to make such infallible identifications was untested. This is not, however, because falsifiability is the criterion of the scientific, but because any serious inquirer is required to seek out all the potentially available evidence, and to go where it leads, even if he would prefer to avoid, ignore, or play down information that pulls against what he hopes is true.

Yes, this is a requirement on scientists, as Darwin recognized when he wrote in his autobiography that he always made a point of recording recalcitrant examples and contrary arguments in a special notebook, to safeguard against his tendency conveniently to forget negative evidence. But it is no less a requirement on other inquirers, too, as we all realized a few years ago when a historian who announced that he had evidence that Marilyn Monroe had blackmailed President Kennedy turned out to have ignored the fact that the supposedly incriminating letters were typed with correction ribbon, and that the address included a zip code—when neither existed at the time the letters were purportedly written!

“Non-science” is an ample and diverse category, including the many human activities other than inquiry, the various forms of pseudo-inquiry, inquiry of a non-empirical character, and empirical inquiry of other kinds than the scientific; and of course there are plenty of mixed and borderline cases. The horrific use of “science” and its cognates tempts us—like the Daubert Court—to criticize poorly-conducted science as not really science at all; but “not scientific” is as unhelpful as generic epistemic criticism as “scientific” as is generic epistemic praise. The pejorative tone of the phrase “pseudo-science,” which presumably refers to activities which purport to be science but aren’t really, derives in part from its imputation of false pretenses, and in part from the favorable connotations of “scientific.” But rather than sneering unhelpfully that this or that work is “pseudo-scientific,” it is always better to specify what, exactly, is wrong with it: that it is not honestly or seriously conducted; that it rests on vague or flimsy assumptions—assumptions there is no way to check, or for which there is no good evidence; that it seeks to impress with decorative or distracting mathematical symbolism or elaborate-looking apparatus; that it fails to take essential precautions against experimental error; or whatever.

So, the Daubert Court’s philosophy of science was muddied; but haven’t subsequent Supreme Court rulings cleared things up? Not exactly: it would be more accurate to say that in Joiner (1997) and Kumho (1999) the Supreme Court quietly backed away from Daubert’s confused philosophy of science. At any rate, those references to Hepper, Pompel, falsifiability, etc., so prominent in Daubert, are conspicuous by their absence from Joiner and Kumho. But there are points of epistemological interest.

In Joiner there is a bit of a kerfuffle about “methodology”: Mr. Joiner’s attorneys had argued that the lower court erred in excluding their proffered expert testimony because, instead of focusing exclusively on their experts’ methodology—which, they maintain, was the very same “weight of evidence” methodology used by the other party’s (G.E.’s) experts—it improperly concerned itself with the experts’ conclusions. Apparently anxious to sidestep this argument, the Joiner Court (with the exception of Justice Stevens) flatly denies the legitimacy of the distinction between methodology and conclusions. Opining that this is No Real Distinction, the Court sounds like nothing so much as a conclave of medieval logicians; but given their citation to Paoli, it seems likely that they didn’t really intend to make a profound metaphysical pro-

— Philosophy and Law —

nouncement, only to acknowledge, as Judge Becker had, that if an expert’s conclusions are problematic enough, this alerts us to the possibility of some methodological defect.

This focus on “methodology”—an accordion concept expanded and contracted as the argument demands—obscured a much deeper epistemological question. Mr. Joiner’s attorneys proffered a collage of bits of information, none sufficient by itself to warrant the conclusion that exposure to PCBs promoted Mr. Joiner’s cancer, but which, they argued, taken together gave strong support to that conclusion; G.E.’s attorneys replied, in effect, that piling up weak evidence can’t magically transform it into strong evidence. In response, Mr. Joiner’s attorneys referred to the EPA guidelines for assessing the combined weight of epidemiological, toxicological, etc., evidence. But no one ever addressed the key question: is there a difference between a congeries of evidence so interrelated that the whole really is greater than the sum of its parts, and a collection of unrelated and insignificant bits of information, between true consilience and the “faggot fallacy”—and if so, what is it?

There is a difference. Evidence of means, motive, and opportunity may interlock to support the claim that the defendant did it much more strongly than any of these pieces of evidence alone could do. Similarly, evidence of increased incidence of a disease among people exposed to a suspected substance may interlock with evidence that animals biologically similar to humans are harmed by exposure to that substance, and evidence indicating what chemical mechanism may be responsible, to support the claim that this substance causes, promotes, or contributes to the disease, much more strongly than any of these pieces of evidence alone could do. However, the interlocking will be less robust if, e.g., the animals are unlike humans in some relevant way, or if the mechanism postulated to cause damage is also present in other chemicals not found to be associated with an increased risk of disease, or, etc.

“Interlocking” is exactly the right word; for evidence is structured like a crossword puzzle, with each claim, anchored by experiential evidence (the analogue of the clues), enmeshed in reasons (the analogue of completed intersecting entries). How reasonable a crossword entry is depends on how well it is supported by the clue and completed intersecting entries, how reasonable those other entries are, independent of this one, and how much of the crossword has been completed; similarly, how warranted a claim is depends on how supportive the evidence is, how secure the reasons are, independent of this claim itself, and how much of the relevant evidence the evidence includes. Because of the ramification of reasons, the desirable kind of interlocking of evidence gestured at in Joiner is subtle and complex, not easily captured by any mechanical weighting of epidemiological data relative to animal studies or toxicological evidence. Nor, moreover—as Justice Rehnquist already pointed out in Daubert—can its quality readily be judged by someone who lacks the necessary background knowledge.

In Kumho, the Supreme Court made a real epistemological step forward. In this products-liability case, focused on the proffered testimony of an expert on tire failure, the Court tried to sort out the problems with non-scientific experts which, as Justice Rehnquist had anticipated, soon arose in the wake of Daubert, and ruled that judges can’t evade their gatekeeping duty on the grounds that proffered expert testimony is not science: the key word in FRE 702, after all, is “knowledge,” not “scientific.” No longer fussing over demarcation, recognizing the gap between “scientific” and “reliable,” in Kumho the Supreme Court acknowledges that what matters is whether proffered testimony is reliable, not whether it is scientific. Quite
Far from backing away from federal courts’ gate-keeping responsibilities, however, the Joiner Court had affirmed that a judge’s decision to allow or exclude scientific testimony, even though it may be outcome-determinative, is subject only to review for abuse of discretion, not to any more stringent standard; and the Kumho Court, pointing out that, depending on the nature of the expertise in question, the Daubert factors may or may not be appropriate, held that it is within judges’ discretion to use any, all, or none of them. A year later, revised Federal Rules made explicit what, according to Judge Overton’s ruling, and Ruse’s testimony, along with Larry Laudan’s properly scathing critique, can be found in Bat Is It Science? The Philosophical Question in the Creation/Evolution Controversy, ed. by Michael Ruse (Buffalo, NY: Prometheus Books, 1996).

Karl R. Popper, The Logic of Scientific Discovery (supra, note 6), 251-2, note *1, added in the English edition. When Popper uses “confirm” for “corroborate”—as he does in his 1957 Philosophy of Science: A Personal Report (supra, note 7)—the effect is powerfully confusing.

Karl R. Popper, Objective Knowledge (supra, note 9) 18, 22.

Id., 20: the reference to Hempel is in footnote 29.


For a detailed development of the conception of scientific method on which I have relied here, see Susan Haack, Defending Science — Within Reason: Between Scientism and Cynicism (Buffalo, NY: Prometheus Books, 2003), Chapter 4.

Ramirez v. State, 542 So. 2d 352 ( Fla. 1989); Ramirez v. State, 651 So. 2d 1164 (Fla. 1995); Ramirez v. State, 8120 So. 2d 836 (Fla. 2001).

Florida remains officially a Frye state, but seems to be rapidly evolving in the direction of (as Michael Saks puts it) Fryebert.


The phrase “faggot fallacy” was introduced in Petr Skrabanek and J. McCormick, Follies and Fallacies in Medicine (Buffalo, NY: Prometheus Press, 1997), and adopted by G.E.’s attorneys in Joiner.

I first introduced the analogy in “Rebuilding the Ship While Sailing on the Water,” in Roger Gibson and Robert Barrett, eds., Perspectives on Quine (Oxford, UK: Blackwell, 1990), 111-27. It was articulated in more detail in Susan Haack, Evidence and Inquiry: Towards Reconstruction in Epistemology (Oxford: Blackwell,1993), Chapter 4, and is developed further in Susan Haack, Defending Science — Within Reason (supra, note 21), Chapter 3.

My thanks to Mark Migotti for very helpful comments on a draft.
Two Concepts of Reliability

Dale A. Nance
Case Western Reserve University, Cleveland, OH

In Daubert v. Merrell Dow Pharmaceuticals, Inc. and Kumho Tire Co. v. Carmichael, the United States Supreme Court set the law of expert testimony on a quest for “reliability.” These decisions made it clear that trial judges are to perform a “gatekeeping” function, filtering out proposed testimony when the expertise on which it is based, whether scientific or otherwise, is not reliable. The new requirement has spawned a substantial literature and furious intellectual battles. Little attention has been given, however, to analyzing the relationship between the reliability requirement and the purposes of admissibility rules. In this essay, I argue that one key to making progress on the contested matters is changing the way courts and commentators conceive, or at least how they articulate, the requirement, so that this relationship becomes more transparent.

I. The Problem

Law is a blunt instrument. We require that a criminal defendant be found guilty or not guilty of any charged offense, and that a civil defendant be found liable or not liable for injuries alleged. For most cases, compromise verdicts are not allowed, at least not officially. Of course, the great range of human interactions that are the subject of law are not easily categorized in such dichotomous terms, even when there is little doubt about what actions were taken and with what mental states or purposes. Even when known precisely, a person’s actions can be appraised as more or less blameworthy, more or less irresponsible. Epistemic considerations complicate the matter further. But when brought before the courts, categorization is required. A judgment of “somewhat guilty” or “very liable” is not given. If compromise is to be had, it is instead to be found in the domain of negotiated settlement, a process that occurs primarily outside the courtroom.

The law of evidence presents many examples of this tension. The law of admissibility, for example, specifies what evidence will be “admitted,” that is, allowed to be introduced by a party and considered by the trier of fact. (The trier of fact is sometimes a jury, sometimes a judge. For simplicity, and to keep roles straight, I will refer to the trier of fact as “the jury.”) There is no compromise: a given item of evidence, offered to support a given inference, is either admissible or inadmissible; it is not somewhat admissible. This, despite the fact that evidentiary proffers come with widely varying degrees of epistemic strength, what lawyers call “probative value.”

Because admissibility is not about what weight to accord evidence, this would not seem to be a serious problem. Admissibility being about what we allow a party to introduce and the jury to consider, as a first approximation trial judges need only exclude any evidence that is irrelevant to the issues to be decided. Such a framework is, indeed, the law’s starting point. Under the Federal Rules of Evidence, and similar state rules, all irrelevant evidence is inadmissible, and all relevant evidence is admissible except as otherwise specifically provided. “Relevant evidence” in turn is defined very weakly as “evidence having any tendency to make the existence of any fact of consequence to the determination of the action more probable or less probable than it would be without the evidence.” Ordinarily, therefore, evidence need not have any particular degree of probative value in order to be admissible, so long as its probative value is not nil, a simple dichotomous characterization.

If that were all there were to it, the law of admissibility would be much simpler than it is. In fact, there are numerous contexts in which relevant evidence is nonetheless excluded. Many of these rules, though not all, are based on concerns about reliability. The most familiar is the hearsay rule, which excludes evidence of an out-of-court assertion—usually not made under oath or subject to cross-examination—when offered to prove that the assertion is true, unless the assertion falls within one of the many exceptions. Behind this and almost all other specific admissibility rules is an important provision—Rule 403—granting the trial court a residual discretion to exclude any evidence the probative value of which is “substantially outweighed by the danger of unfair prejudice, confusion of the issues, or misleading the jury, or by considerations of undue delay, waste of time, or needless presentation of cumulative evidence.”

Aside from such constraints on what is admissible, there is an importantly different set of rules governing the sufficiency of evidence to support a verdict. Under these rules, upon motion of a party, a court may terminate the action prior to, or in opposition to, a jury verdict, on the ground that a reasonable jury could not find in favor of the non-moving party. Such rulings assess the weight of the evidence and depend on the burden of proof applicable in the case. They are also subject to an important limitation: a court may not terminate a criminal case in this manner against the accused.

Until 1993, the admissibility of expert testimony was largely governed by the indicated norms of relevance and Rule 403 “balancing,” although scientific evidence considered “novel” was sometimes subjected to a test of “general acceptance” among experts in the field. Then, amid publicly expressed concerns over the use of “junk science” in tort cases, the Court decided Daubert (1993) and Kumho Tire (1999), and in 2000 Congress passed a corresponding amendment to Rule 702, adding conditions (1) through (3):

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.

These authoritative sources fairly consistently express the reliability requirement in dichotomous terms: expertise is considered either reliable or unreliable. For example, the language of amended Rule 702 encourages us to believe that testimony either is or is not the product of reliable methods reliably applied. The alternative way of expressing the requirement would be to say that there must be a showing of “sufficient reliability” for admission. This employs a concept of reliability that is gradational, reliability being a matter of degree. While less commonly encountered in cases and commentary, this conception does occasionally surface, as in the drafters’ comment on the revised rule that various factors should be considered “in determining whether expert testimony is sufficiently reliable to be considered by the trier of fact.”

The obvious problem with expressing the requirement the latter way is that, without some reasonably determinate algorithm based on appropriate legal norms that would guide courts in determining what degree of reliability is “sufficient” for this purpose, the requirement is almost vacuous. It tells us nothing except that some reliability-based test should be employed. Using the dichotomous form of expression, on the other hand, may only conceal this gap by suggesting either that reliability is a binary factual question, akin to the preliminary
question of whether an out-of-court statement is hearsay, or that there is some criterion of reliability that can be articulated and applied without invoking such an algorithm. If such suggestions are false or unworkable, then we should embrace the implications of working with a gradational conception.

In *Daubert*, the Court identified reliability, in the context of scientific evidence, with *scientific validity* and suggested various factors to be considered in assessing validity. The first factor is “whether [the theory or technique] can be (and has been) tested.” That does seem to suggest a binary factual determination: tested or not. But how much testing does it take before a principle or technique can be said to have been “tested” in a way that would assure scientific validity? Scientists know that replication of results is an important part of the scientific method. Can a principle or technique be considered “tested” if only one test has ever been conducted? Two? Three? Does it not depend on the quality of the tests as well as their quantity? Such queries point to a concept of testing that is a matter of degree, with more testing being better than less, testing that is more closely related to the proposition of interest being better, and testing with better controls being better, all other things being equal.

Essentially the same observation can be made about the other *Daubert* factors and similar factors suggested by subsequent courts and commentators, including the existence of a known or potential error rate, the existence and maintenance of standards for a technique’s operation, and the more indirect factors such as peer review and acceptance among the scientific community. All these can be understood as gradational parameters. Indeed, scientists, at least sophisticated ones, surely understand that scientific validity itself is not an all-or-nothing characteristic; it is a matter of degree. Which poses the question of how one gets from a consideration of the degree of scientific validity to the undeniably binary decision of whether or not to admit proffered testimony. The same question is posed by the requirement of reliability for expertise not regarded as “scientific.” Thus, the fundamental problem is how to map from a gradational epistemic conception of reliability to a dichotomous legal choice on admissibility.

In articulating an answer to this problem, it is important to avoid the pitfall of conflating admissibility with sufficiency. That is, evidence should not be ruled inadmissible just because it is insufficient to support a verdict in favor of its proponent. “Sufficiently reliable to be considered” is not the same as “sufficiently reliable to warrant a verdict.” Some suggestions about how to fill out the reliability requirement fall victim to this error. To be sure, it may be important to facilitate preemptive determinations by judges in some cases for which the expert evidence is overwhelmingly one-sided, but a reliability-based admissibility regime is not the right way to achieve this goal. The distinction between admissibility and sufficiency must be kept clear in order to avoid conceptual confusion and mistaken rulings.

**II. Reliability Determinations Without Balancing: A Critique**

Conventional post-*Daubert* admissibility analysis seems to presuppose that reliability can be understood, and that reliability determinations can be made, without the kind of policy-informed balancing relative to legal goals that is explicit in rules like Rule 403. For example, a well-regarded treatise explains the conventional understanding of the residual role of Rule 403 in screening expert scientific testimony in the following terms:

Under Rules 702 and 104(a), judges must decide whether the proponent of scientific evidence has demonstrated the validity of the scientific basis for the testimony by a preponderance of the evidence. In many cases, however, while judges might find scientific evidence to be “valid,” they might believe that it is not valid enough, in light of the dangers associated with its use and thus exclude the evidence under Rule 403.

The obvious slide in this passage from a dichotomous conception to a gradational conception of validity (and thus reliability) points to an important question: Is there any usable, non-balancing criterion of validity with respect to which Rule 403 can serve as a residual exclusionary principle? If not, then it makes no sense to say that the validity determination under Rule 702 can be meaningfully separated from a consideration of the “dangers associated with [the evidence’s] use.” In this section, I identify and reject the most plausible lines of argument that might sustain the conventional analysis.

**Dichotomy by Deference.** Perhaps a reliability determination could be made by reference to factual propositions about the standards of non-legal institutions and their proper application. Rather than a decision based directly on legal goals and policies, the courts would assess reliability using non-legal norms. This would involve deference to the norms of the expert community. In the context of scientific evidence, *Daubert* seems to suggest just such an approach by instructing courts to assess scientific validity and directing them to look at factors most of which scientists would think pertinent in making their own assessment.

The obvious and frequently noted problem is that this methodology requires judges to become quasi-scientists in order to try to apply the standards of the scientific community. That is a task that some judges are, or believe they are, incapable of performing adequately. It is complicated by the existence of disagreements among scientists and philosophers regarding the norms of scientific disciplines, disagreements of which there is barely a hint in the Supreme Court’s opinions.

The more fundamental question, however, is this: Why should the final determination, reliable or not in court, be determined by the norms of the scientific or other expert community instead of those of the legal community? If the scientific community recognizes scientific validity as a gradational concept, then that community might employ normatively informed algorithms of its own for making any categorical determinations of validity that it is called upon to make, norms developed relative to the values and interests of the scientific community. If the scientific community’s standards sometimes regard scientific validity as an all-or-nothing concept, it is only because that community has developed dichotomous rules of thumb that, while over- and under-inclusive, roughly serve to further those values or interests. Why then should the scientific community’s balancing of such considerations, whether directly as to the particular issue or indirectly by the application of such rules of thumb, be taken as controlling for the conduct of litigation?

Various reasons can be imagined, but they are ultimately unconvincing. Most prominently, it might be thought that the purposes and goals of the scientific community are the same as those of the law of adjudication, or at least so nearly the same that the difference is not worth worrying about. If so, then some economy can be gained by piggy-backing on the established standards of the non-legal discipline, which may also avoid the controversy of having to work out standards of reliability in contested court cases.

The indicated premise, however, is unlikely to hold true in the scientific context. To be sure, both disciplines place a high priority on the ascertainment of truth by the application of relatively formal procedures. But the parameters that give shape to these disciplines differ considerably. Those of good
adjudication include, for example, an emphasis on promptness and finality of decision that is incompatible with the norms of science, as Justice Blackmun himself observed in Daubert.\(^{11}\)

Moreover, the law’s concession to unavoidable doubt often entails an aversion to false negatives (e.g., refusing to impose liability on defendant for want of proof of causation when its product did in fact cause plaintiff’s injury) that is almost as strong as its aversion to false positives (e.g., imposing liability when defendant’s product did not cause the plaintiff’s injury), a situation not reflected in the typical statistical standards of hypothesis testing employed by scientists. These conservative scientific norms consistently privilege false negatives by saying, in effect, that it is much, much better to conclude “no causation shown” and await further testing, than to conclude erroneously that causation is present.\(^{12}\)

The argument from identity of goals is thus profoundly incomplete without an explanation of how legal admissibility norms, which concern the helpfulness of information in reaching a prompt and final verdict on particular events in accordance with the applicable burden of persuasion, map to scientific validity norms, which relate to the permissibility of scientists deriving perpetually revisable conclusions about the general patterns of causation in a context of inquiry that does not even have rules of “admissibility.” While it is fairly safe to assume that such scientifically well-grounded conclusions are going to be helpful in lawsuits to which they are relevant, the converse cannot be safely assumed. Scientific validity, as understoed by scientists, should not be considered necessary in all cases for adjudicative helpfulness. To do so risks conflating admissibility with sufficiency and applying an inappropriately conservative sufficiency standard.

Is the situation different for non-scientific expertise, what Rule 702 refers to as “technical or other specialized knowledge”? There are serious, practical disciplines, such as medicine, for which it can be said that the parameters of decision, such as the priority given to accuracy, promptness, and even finality, are often—though not always—more akin to those of adjudication than are the pure sciences. On the other hand, there are countless forms of specialized “knowledge,” ranging from astrology to forensic science, which for this is not true. Some non-scientific disciplines, even some that have been regularly used in the courts for decades, have little or no extra-disciplinary checks on reliability. As recent debates have made clear, this may be especially true for disciplines, such as handwriting identification, that have been developed for, and find little application other than, forensic uses. To defer to the normative standards of reliability in such disciplines would be to abdicate the basic gatekeeping function.

As a general strategy, therefore, deference to the reliability norms of non-legal institutions is no more plausible in the context of non-scientific expertise than it is with regard to expertise regarded as scientific, even if the border between the scientific and the non-scientific could be effectively policed, a premise explicitly rejected by the Court in Kumho Tire. Selective deference to particular disciplines is no more practical; nor is there any hint of such an approach in the authoritative sources.

Non-deferential Dichotomy. If deference will not work, it might still be possible to implement a dichotomous reliability decision procedure that does not require the judge to invoke (without saying so) a legal-policy based balancing formula. Many rules of admissibility that do not concern experts, but are concerned with reliability, provide potential models.

Take the rule excluding hearsay. That rule does not provide that, when confronted with evidence of a statement not made on the stand in the current trial, the trial judge should consider designated factors, such as the absence of cross-examination of the out-of-court declarant, and weigh these against the probative value of such evidence in order to decide admissibility. Nor does it say, more elliptically, that the trial judge should simply weigh all competing factors affecting reliability. Instead, it requires categorically the exclusion of such a statement when offered to prove the truth of the matter asserted, if it does not come within one of the (for the most part equally) factually binary exceptions. In other words, these dichotomous tests encapsulate, or so it is hoped, the goals of the adjudicative law, the most important of which in this context is verdict accuracy, but without requiring, at least in the ordinary case, any advertence to those goals or the balance of competing considerations that relate thereto.

Analogously, one might select one or more dichotomous proxies for reliability, each of which tests for one particular way in which reliability can be undermined. If the expertise is not undermined in one of these specific ways, the expertise satisfies the reliability requirement. Such a proxy test, if well or fortuitously designed or evolved, might not be too over- or under-inclusive for practical use. Indeed, this might look like the scheme that the Supreme Court initiated with Daubert, its “factors” constituting the proxies.

Such a scheme, however, cannot create an intelligible dichotomous standard of reliability unless either (a) only a single, dichotomous proxy is used, so that the expertise would be deemed “reliable” relative to the present requirement provided the single proxy test is passed, or (b) in the multiple proxy context, one has a determinate way of synthesizing the results of those proxy tests or factor applications into a single, binary judgment without performing an independent assessment of reliability. Pretty clearly, Daubert does neither. It rejects any single factor as determinative, it leaves open the possibility of additional pertinent factors, and a fortiori it disclaims providing any reasonably determinate means of synthesizing the various factors to be considered. Daubert does not contemplate the use of a set of proxies for reliability; rather, it requires a case specific, direct assessment of reliability itself by use of all appropriate factors.

Beyond that, in order to formulate a set of such tests, if we are to rely on more than luck, we must take into account those considerations that would justify the exclusion of evidence of a given degree of reliability. Authoritative expressions of the reliability requirement are surprisingly silent on this matter. Almost all factors courts and commentators identify relate only to the assessment of the degree of reliability, as if one could know that expertise is sufficiently reliable by determining only that degree. That’s like imposing a job qualification that employees must be “tall enough” and providing decision makers only guidance about how to determine an applicant’s height. It invites the obvious question, “Tall enough for what?”

III. Toward a Policy-Based Theory of Ad Hoc Balancing

To articulate an analysis that gives explicit attention to the considerations that determine how reliable expertise needs to be, we can draw on a rich body of theory about the rationale of other admissibility rules. Many such rules attempt to answer, in particular contexts, the question of whether evidence is sufficiently reliable to be admitted. Historically, these rules have responded to one or more of three main concerns: (1) that the offered item of evidence will be of too little probative value to warrant the necessary expenditure of time and resources, including the cognitive resources of the trier of fact, necessary to incorporate it into the decision-making process (generally of concern in both jury and bench trials); (2) that the trier of fact is
prone to error in its assessment of the probative value of certain types of evidence (a concern expressed primarily in regard to jury trials); and (3) that the trier of fact (judge or jury) should not be forced to assess the case based on the offered evidence when superior evidence is likely to be available.\(^{13}\)

The first two concerns lead to an admissibility structure that prefers the exclusion of the challenged evidence. all other evidence in the case being (hypothetically) the same. They differ in that the second entails a much larger degree of epistemic paternalism than the first. The third concern, however, leads to an exclusionary structure that prefers the replacement of the offered evidence with something else, better evidence that might not otherwise be presented to the tribunal. It reflects more the problem of advocate control than the problem of jury control that tends to animate the second concern. Although much academic and professional commentary over the last century has tended to focus on it, jury credulity is not the only potential source of inaccuracy in adjudication. Even in a bench trial, or a trial before a panel of experts, the court will want the best evidence reasonably available.

All these considerations may be taken into account in the exercise of the judge’s discretion under Rule 403, the most conspicuous “balancing” test in the federal rules. The reliability requirement of Rule 702, however, is clearly intended to impose stricter standards of admissibility for experts, at least in some respect, or it would be unnecessary. The interpretation that we seek should, therefore, would provide for appropriate balancing of these considerations in a way that does not simply replicate Rule 403. In the space remaining, I outline one such interpretation.

One point clearly serves to distinguish the contemplated balancing test under Rule 702 from balancing under Rule 403. Under 403, the burden is on the objecting party to convince the trial judge that the testimony’s probative value is outweighed by the indicated risks.\(^{14}\) In contrast, it is generally agreed that the burden is on the proponent of expertise to establish (sufficient) reliability under 702.\(^{15}\) One might leave it at that, switching the allocation of the burden but otherwise understanding the 702 reliability decision as a replication of the 403 balancing test. That approach is reasonable in the context of the first listed concern, avoiding waste of time and resources, but more needs to be said about the other two.

**Offsetting Jury Credulity.** Both before and after Daubert, commentators have identified — as the primary concern warranting the exclusion of proffered expertise of relatively weak reliability — the idea that a lay jury will be misled, giving such evidence greater credence than it deserves. This idea, a staple of tort-reform advocates, coheres with suspicions about the credibility of experts for hire, if it is thought that jurors are unable to discount for the effects of bias. It also reflects a more general tendency, prevalent since the end of the nineteenth century, to explain veritistic exclusionary rules in terms of distrust of the lay jury.

Strikingly, however, Justice Blackmun’s opinion in Daubert was at pains to disavow any serious skepticism about jury competence, opining that it is “overly pessimistic about the capabilities of the jury” to fear litigation “in which befuddled juries are confounded by absurd and irrational pseudoscientific assertions.”\(^{16}\) And neither the opinion in Kumho Tire nor the advisory committee’s explanation of the 2000 amendment to Rule 702 speaks to the contrary.

The available empirical evidence tends to support Blackmun’s opinion. That evidence points to jurors being remarkably conscientious in their work and not demonstrably less accurate in their inferences than judges. Specifically, we have no empirical basis to conclude that jury credulity in over-crediting expert testimony is a serious problem.\(^{17}\) If anything, jurors may excessively discount complex expertise.\(^{18}\)

Until we have more evidence of jury credibility, as well as more evidence that judges are in fact capable of offsetting such incompetence by their decisions to exclude expertise, respect for the political significance of the jury and for those citizens who are called to service on juries cautions against aggressively excluding expertise on this ground. Instead, we should leave this kind of problem to be handled under the pro-admissibility standard articulated in Rule 403, as indeed suggested by the Court in Daubert.\(^{19}\) This will help to counteract the inertial tendency of many in the legal community to employ jury distrust as the preferred mode of admissibility analysis.

**Securing Better Evidence.** Much more important in articulating a meaningful content for the reliability requirement of Rule 702 is the idea that evidence may be excluded to encourage the presentation of better evidence, evidence that is more probative or less costly for the tribunal, or otherwise presenting a more favorable balance between the two.\(^{20}\) This idea can be traced to eighteenth century theorizing, set within a context of greater concern about parties’ lack of pre-trial access to evidence in the possession of an opponent. Despite the improved pre-trial access resulting from modern rules of discovery, especially in civil cases, the “produced” nature of modern expert testimony—the ability of parties to select among experts and develop expertise, sometimes at great cost—arguably requires closer monitoring of a proponent’s choice of expert testimony than would otherwise occur under the pro-admissibility standard of Rule 403.

The suggestion is that the judge’s task is to inquire whether a reasonable jury, sensitive to the delays and costs associated with a demand for more reliable expertise, would express such a preference, given the nature of the controversy and what is at stake. The judge would act to facilitate good inferences by speaking on behalf of the mostly silent jury, not to channel the jury’s inferences out of suspicion of jury credulity. On behalf of the jury, the judge will want the best evidence that is reasonably available, with due regard to the adversarial structure of the trial and the economics of litigation, in knowledge of which the trial judge has a distinct comparative advantage relative to the jury.

Thus, upon a challenge to proffered expertise that identifies the potentially more reliable expertise that might be brought to bear, a court would first require—in view of the challenger’s superior knowledge of the matter—that the challenger convince the trial judge that such evidence is not reasonably available to the challenger. If that case is made, the proponent would bear the burden of convincing the trial judge either (i) that such expertise would not be discernibly more reliable in the present context, or (ii) that it is not reasonably available to the proponent. Again, placing the latter burden on the proponent, coupled with the consistently jury-supportive judicial viewpoint explained above, would serve to distinguish the reliability inquiry of Rule 702 from the similar balancing test of Rule 403.

This approach does not require cardinal measures of reliability, only the ability to determine whether one form of expertise is discernibly more reliable than another and to assess the indicated cost-benefit questions. Such questions, though they may be difficult at times, are more tractable for judges than trying to determine the probable effects of proffered testimony on jurors. And it poses less risk of the judge trying to engineer a jury verdict to the result that the judge thinks is correct, in part because it clearly addresses the question of what is helpful to the jury rather than what is sufficient to warrant a jury verdict in favor of the proponent.
Very generally speaking, the suggested approach will place greater demands on the prosecution than on the accused, and it will place greater demands on powerful civil defendants than on impecunious civil plaintiffs. Moreover, greater reliability might be unavailable to a party within the context of a particular case, yet reasonably available to that party within the context of repeated litigation of the same or similar issue. At the outer reaches of the better evidence idea, repeat players, such as the state in regard to forensic science techniques, may plausibly be considered in regard to the long run of cases, provided that party has a significant degree of control over the reliability of expertise that is generated.

In contrast, for the litigant who must take what he or she finds in terms of the reliability of available expertise, Rule 702 should not demand more. To exclude such a party’s expertise just to force the community of experts to improve reliability in future cases treats that party as merely a means to the end of more accurate later adjudication. Of course, if either party offers expertise that is of such weak reliability as to be simply a waste of time, it should be excluded on that basis even if more reliable evidence is not reasonably available.

Occasionally, such a structure will mean that the party with less control over reliability of particular expertise will be allowed to introduce evidence with a degree of reliability that would not be accepted from the opponent. Among the ways to address this seemingly unfair result, I favor a rule that would allow the latter party to use such evidence in rebuttal, only if the former has introduced comparable expertise at a previous stage of the trial. This structure, which gives the party with less control an option to “open the door” to a particular form of expertise, retains considerable pressure on the other party without the prospect of skewed interim verdicts.

IV. Conclusion
If sense is to be made of the post-Daubert reliability requirement, we must abandon the dichotomous conception of reliability, as well as its near cousin, the idea that judges can know that evidence is sufficiently reliable without seriously addressing the reasons that evidence of given reliability might need to be excluded. When we look at those reasons, we find that jury distrust is of little help as a framework. Instead, we should encourage judges to assure that expert testimony is as reliable as the circumstances permit, and then allow the jury to do its job.

Endnotes
5. Fed. R. Evid., Art. VIII.
7. 509 U.S. at 593.
8. See Samuel R. Gross, “Substance and Form in Scientific Evidence: What Daubert Didn’t Do,” in Reforming the Civil Justice System, edited by Larry Kramer (New York: New York University Press, 1996), 234 (arguing that what some courts have wanted to do in toxic tort litigation is to rule for the defense on the merits when the defendant’s expert testimony is much stronger than plaintiff’s, and that for want of suitable doctrinal authority for doing so, they struck upon the idea of excluding the plaintiff’s expertise and then terminating the case for want of sufficient evidence to support a verdict for the plaintiff).
10. This form of deference is distinct from a deference to the collective opinion of the expert community regarding reliability or validity; the latter (under the name of the “Frye test”) was specifically rejected by the Daubert Court as an interpretation of the federal rules. 509 U.S. at 585-89.
11. Id. at 596-97.
14. This is the practical effect of the rule’s otherwise mysterious exclusion of evidence only when its probative value is “substantially” outweighed by such risks. What sense would it make to admit evidence, as the rule seems to require, when its probative value is outweighed by the risks, but not substantially so?
15. See Faigman et al., supra note 9, § 1-3.1.2.
19. 509 U.S. at 595.

Daubert and the Acceptability of Legal Decisions
Carl F. Cranor*
University of California, Riverside, CA

One who is absolutely committed to the process of ascertaining and testing the truth, and who would thus shun any concession of the search for truth to the production of acceptable verdicts, may find that he does so at the expense of other important values. He may discover that extremes in the pursuit of truth can impair the system’s capacity to generate acceptable verdicts and thus undercut its ability to project the norms embodied in the substantive law.¹

I.
In a series of cases beginning with Daubert v. Merrell-Dow Pharmaceuticals Inc., the U.S. Supreme Court gave federal judges a heightened duty to review scientific evidence and expert testimony which are proposed for admission into civil and criminal litigation.² Two of these cases were toxic tort cases, the focus of this essay.

A review of the admissibility of scientific evidence or expert testimony is not (or in principle should not be) aimed at adjudicating between the factual claims submitted by the parties; that is an issue that constitutionally should be left to juries. Rather, according to Daubert judges’ admissibility decisions should ensure that scientific evidence and expert testimony are “reliable,” that is, “ground[ed] in the methods and procedures of science,” and they must fit the facts of the case.³ Moreover, admissibility decisions on scientific evidence are made before a jury is empanelled and before a trial proper.
Moreover, it will often be difficult for persons exposed to substances that only exacerbates problems of identifying toxicants. Or biochemical threat poses a new scientific detective problem, should be utilized in the law, especially in the tort law. 4 Proof “beyond a reasonable doubt” before drawing

considerable uncertainty even though their personal discoveries, scientists ordinarily assert their views with considerable uncertainty even though their personal beliefs may be stronger. Some scientists even demand proof “beyond a reasonable doubt” before drawing conclusions because alternative explanations will slay “a beautiful [but mistaken] hypothesis.”7

These scientific conventions would further handicap the party with the burden of proof, almost always the plaintiff.

Beyond the above problems and prior to the Daubert decision, some legal scholars argued that there is a procedural bias in the tort law against plaintiffs resulting from problems of latency, proof of causation, free-rider problems, and so on. However, they also argued that this procedural bias was counterbalanced by a substantive bias for plaintiffs once they were in court, despite plaintiffs having the burden of proof. Thus, the institution struck a reasonable balance between those who create, and those who are exposed to, risks. In their view, there was no need to reform the law as many were advocating.8 Nonetheless, the Daubert decision in effect reformed the law concerning expert testimony and exacerbated some of the above problems in the process.

In implementing Daubert some judges appear to have imposed more stringent requirements on scientific evidence than scientists themselves would—by demanding evidence scientists would not require or excluding evidence scientists would routinely utilize. Such decisions deprive plaintiffs of some reasonable evidence that scientists utilize in their own fields for inferring causation, further exacerbating the barriers faced by plaintiffs.9

There are reasons for thinking that even if Daubert had been implemented to admit precisely the right kinds of expert testimony, there are likely to remain problems with toxic tort law. There is profound ignorance of the universe of substances,10 which together with the causation requirement creates incentives “on the part of corporations not to know and not to disclose” information about the harms their products may cause.11 To address this the substantive law may need to be reformed to encourage greater toxicity testing, better data-gathering and better reporting of adverse effects in order to better protect the public.12 Ignorance of the chemical universe adversely affects litigants who have the burden to remove the ignorance, typically plaintiffs. The higher the barriers for removing ignorance, the greater the burdens on plaintiffs. Almost certainly the Daubert decision has increased plaintiffs’ barriers.

The above concerns suggest that when substantial scientific evidence is required in torts, the law is likely being altered by Daubert and its progeny. However, the remainder of the essay addresses a less obvious, more subtle issue related to Daubert’s focus on evidence, namely, the difference between the acceptability of verdicts that convey a liability judgment and verdicts that tend to focus on the evidence.

II.

Even before Daubert there were a variety of reasons to be concerned about how science and scientific evidence should be utilized in the law, especially in the tort law.4

There are “institutional” tensions between the law and science concerning the standards of proof required in each “institution”. Consequently, unless legal requirements placed on scientific evidence are not tailored sensitively for the law, implicit burdens and standards of proof from scientific research can easily distort the law. For example, the scientific community strongly and asymmetrically protects against false positive mistakes (i.e., inferential procedures that show a substance is toxic when it is not), while in the tort law there is an institutional commitment to the view that legal false positives (i.e., mistakenly deciding for plaintiffs) should be approximately equal to legal false negatives (i.e., mistakenly deciding for defendants).2

Features of substances and of scientific research exacerbate these generic tensions. In research, scientists tend to presume substances have no particular properties, including toxicity, until they have been established by appropriate research. Moreover, it can be quite difficult to show toxicity for substances that cause diseases with long latency periods, cause common diseases, or have subtle effects. In many cases each substance or biochemical threat poses a new scientific detective problem, which only exacerbates problems of identifying toxicants. Moreover, it will often be difficult for persons exposed to toxicants to discover that their diseases are the result of exposure.5

In addition, there will often be considerable uncertainty attending scientific claims about suspected toxic substances, but in the law uncertainty can be a substantial barrier for the party with the burden of proof. Since science is open-ended, and since even comparatively settled conclusions are open to revision upon the presentation of new data, theories, or discoveries, scientists ordinarily assert their views with considerable uncertainty even though their personal beliefs may be stronger. Some scientists even demand proof “beyond a reasonable doubt” before drawing...
of the judicial process, then, is to project to society the legal rules that underlie judicial verdicts.13

One way to make these points is to consider some of the “gaps” between comparatively raw evidence (I use scientific evidence as the example) and an ultimate legal conclusion about liability. Consider Allen v. Pennsylvania Engineering, Inc.14

Walter Allen worked for a hospital as a maintenance man, sometimes changing bottles of ETO in hospital sterilization chambers. Mr. Allen contracted brain cancer, a quite rare disease. He and his wife alleged that his brain cancer was caused by ETO and that the hospital and ETO manufacturer should be held liable. The hospital settled before trial.

How might the Allens establish their case against the manufacturer of ETO? They need some scientific evidence showing that ETO could cause brain cancer (so-called “general causation”) and that ETO did cause Mr. Allen’s brain cancer (“specific causation”). They also need to show, given the causal claims, under the applicable law, that Pennsylvania Engineering should be held liable for his brain cancer.

There are at least two major “gaps” in arguments from basic scientific studies to liability judgments. There is an “evidence-causation” gap from basic scientific studies to the two causal judgments.15 There is also a “causation-liability” gap from any causation established to showing liability. Both inferences are typically undetermined.

The raw scientific studies in this case were good studies in rats showing that ETO could and did cause brain cancer in rats exposed to ETO compared with rats not so exposed. ETO does not appear to cause brain cancer in mice. ETO is a small molecule that is a direct acting alkylating agent; that is, without needing metabolic reduction in the body it can attach itself to DNA anywhere in the body and disrupt DNA function.16

There was also some suggestive, but not statistically significant, evidence that ETO caused brain cancers in some humans occupationally exposed. A large meta-analysis of these small studies published by a very controversial scientist did not show that there was a higher rate of brain cancer in ETO-exposed persons than in non-exposed persons.17 These scientific studies are not transparently interpretable; do they show causation or not?

Plaintiffs argued that because ETO was a direct acting alkylating agent and because it caused brain cancer in rats, this provides a mammalian model indicating that it could cross the blood-brain barrier and cause cancer in humans. Rats appear to be better models for the adverse effects of ETO in humans, because their breathing rate is much nearer that of humans than the very rapid inhalation rate of mice.18 Moreover, there were some small human studies suggesting an association between ETO exposure and brain cancer.

Defendants argued that because ETO did not cause brain cancer in mice, which are phylogenetically more similar to rats than rats are to humans, rat studies are not evidence that ETO can cause cancer in humans. Moreover, the defense greatly emphasized the significance of the meta-analysis of human studies and claimed that it showed that ETO did not cause cancer in humans. One defense expert opined, “It is impossible to believe that” ETO caused Mr. Allen’s brain cancer. The judge excluded plaintiffs’ experts and issued a summary judgment for defendants.

Ultimately, if a case goes to trial, a jury must assimilate all of the causal and pertinent historical evidence of the events leading to the case with the applicable law to come to a liability verdict (or not).

Nesson would argue that there is a difference between Pennsylvania Engineering being found liable for causing Mr. Allen’s brain cancer and the claim that the toxicological, epidemiological and exposure evidence tended to favor the view that Pennsylvania Engineering was liable for Mr. Allen’s brain cancer. A verdict of liable in civil litigation calls attention to the human and social events that transpired and the corresponding legal rules that were violated, and invites people to act according to what is lawful and avoid unlawful behavior. An evidentiary claim calls attention to the evidence available and the extent to which a verdict might be proven, in turn inviting citizens to act “according to what they think can be proved against them.”19 The evidence-only claim ultimately leaves us in doubt about causation, much less secure in a liability claim, and in doubt about whether justice has been done between the parties.

Moreover, a citizen, by viewing the “verdict as a determination of what actually happened, [can] assimilate the applicable legal rule and absorb its behavioral message”; it also reminds a person of “what constitutes proper legal conduct in the circumstances.”20 If a citizen regards a verdict “as merely a statement about the evidence, he will assimilate only the proof rule, whose deterrent power [and behavioral message] is far less pronounced.”

Nesson appears to support his distinction by means of the following considerations. “The projection [and receipt] of the verdict as a statement about what happened is the key to conveying the legal rule and its behavioral message. Projecting the verdict as such forges a link between the judicial account of the defendant’s transgression and our own behavior.”21 The idea seems to be that those receiving the verdict about the people involved and what happened tend to identify with or perhaps recognize the significance of the behavior of those who were violating the law toward others. Such recognition can have a significant impact on our and others’ behavior to the extent the liability judgment is understood and internalized; we would tend to take the legal message embedded in the verdict seriously and be disposed to modify our conduct accordingly. To the extent that a court decision conveys a message about the persons involved and their violation of the law, this creates the possibility that “[w]hen similar situations arise in the lives of those who have accepted the verdict as a determination about what happened, these citizens will govern their conduct in accordance with the behavioral rules that they have absorbed.”22

By contrast, if a legal dispute is primarily about the evidence involved, this may undermine a behavioral or deterrence message. Citizens typically would not recognize the significance of generic evidence claims for their own behavior in the same way they identify with statements about the behavior of others being held liable. And, a focus on the evidence keeps the attention on the evidence and the extent to which, say, legal violation can or cannot be established. Thus, instead of conveying that the defendant negligently injured a plaintiff, it conveys and asks those receiving the evidence message to “assimilate the rule that one should not negligently injure others in a manner that allows them to prove it.”23 A recent internal memo from Bayer Pharmaceutical concerning product testing seems to exhibit this view: “If the F.D.A. asks for bad news, we have to give, but if we don’t have it, then we can’t give it to them.”24

IV.

If Nesson’s account is plausible, it suggests some ways in which the Daubert decision fosters the acceptance of a verdict about what happened, but also shows how, by an undue focus on the evidence, it can tend to undermine the acceptance of judicial decisions and the outcomes of the judicial process. 25
How Daubert Could Foster Acceptable Verdicts. Begin with a point Nesson makes about directed verdicts; they “prevent the legal system from generating unacceptable verdicts. The directed verdict permits the court to withhold from the jury those cases in which a finding of guilt or liability would be patently untenable in light of the case presented by the plaintiff. The trial judge allows a case to go to the jury only if the evidence suffices to support a verdict either way.”272 A directed verdict is not equivalent to an admissibility decision, but it resembles a ruling on the sufficiency of the evidence that typically follows a ruling excluding critical evidence (recall the introduction).

The Daubert decision seeks to ensure that expert testimony is based on “valid science” or on reasonable “scientific reasoning and methodology.” Thus its aim seems to be to winnow expert testimony so that a jury decision based in part on either plaintiffs’ or defendants’ experts’ accounts of the science will be within the bounds of respectable scientific views about the issue involved. Consequently, whatever the jury decides will not be beyond respectable scientific reasoning on that issue and will be (broadly) scientifically acceptable. This does not ensure that the overall verdict will be acceptable, but an important aspect of it will be.

The test of admissible evidence has not been given a definitive formulation (and a guide that is too specific should probably be avoided). However, there are several guidelines that have been suggested. In extra-legal venues, Justice Breyer has suggested that the test should be analogous to one he credits to the physicist Wolfgang Pauli—expert testimony should not be permitted, if it is “so bad it cannot even be wrong,” as Pauli once said of some articles in physics.27 A second test might be that the expert testimony should be permitted as long as it does not fall “outside the range where experts might reasonably differ.”28 A third guide might be that expert testimony is permitted as long as, in Nesson’s words, it would not be “patently untenable” scientific testimony.

I do not adjudicate between these, but only note how they could place constraints on scientific testimony to help ensure that jury verdicts are not scientifically unacceptable.

How Daubert Can Undermine the Acceptability of Judicial Decisions. Poor implementation of Daubert can undermine admissibility decisions. As already noted, some courts are requiring or excluding scientific evidence for reasons that are at odds with scientists’ conceptions of reasonable evidence. To the extent that courts err and their decisions become known, this invites criticism of the legal system analogous to that directed at courts when they were alleged to be too permissive in allowing experts to testify. As such criticisms become known, they tend to challenge the legitimacy of the decisions.

Moreover, when courts make mistakes in excluding plaintiffs’ evidence and their legal actions end with a summary judgment, this denies plaintiffs a trial and the possibility of justice. To the extent that such results become known, this too affects the perceived fairness of the legal system.

However, even if courts properly implement a reasonable version of Daubert, there are reasons, following Nesson, for being concerned about the heightened attention scientific evidence is receiving in admissibility decisions.

It seems plausible that plaintiffs have already been hurt by a focus on the evidence—analogous to Nesson’s point about the difference between a focus on evidence and a focus on liability. A pre-trial focus only on the evidence does not obviously require any consideration of past events, including discussions about defendant’s improper treatment of plaintiff. Instead, as in Allen, the discussion is about the merits of rodent studies, small versus large epidemiological studies, structure-activity tests, and case studies—not on what happened and the relations between people that led to litigation. Plaintiffs are handicapped because they can present only part of their account—the scientific evidence and expert testimony needed for causation. (They will have some account of past events in their complaints, but this is likely to be minimal.) It is difficult for them to offer a full account of defendant’s behavior, whether it was negligent or not, and how the law was violated. Thus, plaintiffs’ focus must be only on what the scientific evidence and expert testimony would support about causation, only on the evidence-causation gap.

Daubert debates more closely resemble scientific debates or seminars about the quality of studies and the inferences from them about causation, than a legal debate about defendant’s treatment of plaintiff, and whether and to what extent legal rules have been violated. Some may think this is desirable since it forces a clarification of the quality of the science, free from contamination by stories about sympathetic plaintiffs, to see if there is a reasonable basis for a plaintiff’s complaint, and it promises court efficiencies (although these tend to be overrated). Others may find this undesirable.

Moreover, plaintiffs increasingly must invest substantial monetary and expert resources in preparation for an admissibility hearing simply in order to have a trial on what happened. By some estimates, as much as 60% of costs of bringing a case now occur before trial begins.29 This increases lawyers’ screening of plaintiffs and decreases plaintiffs’ access. Often plaintiffs must invest these resources because defendants did not test their products adequately in the first place.

The larger community also loses. First, there is likely to be little discussion of the relationships between people and the applicable law before a public trial. Second, there is no structured public discussion in a trial about what happened, the relationships between plaintiffs and defendants, and whether the law had been properly followed. Third, to the extent there is a legitimate dispute about proper legal behavior, there is no public resolution of this. Thus, the community is deprived of important kinds of information about its institutions and how persons are affected by them;30 at most the community would learn that there was insufficient expert testimony for the case to proceed.

With Daubert, the Supreme Court itself has elevated the importance of the quality of scientific studies and expert opinions based upon them above other values in the system. The Court has seemed so concerned to secure these values that it permits judges to err on the side of mistakenly excluding evidence, which precludes adjudication of what happened between the parties, rather than permitting in the evidence and letting the trial and appellate process adjudicate what happened between parties. Adjudicating whether such relationships were legal is just as significant as whether every “i” is dotted and “t” crossed concerning scientific evidence. Thus, the concern about scientific evidence and expert testimony appear to have taken precedence over public adjudication of past events and relations between people, norms of behavior evidenced therein, and the rules, norms and ideals of the substantive law. In short, it seems more important to the law to get the science right by means of Daubert rulings than to adjudicate more fully the past events and lawful norms of behavior.31

There may be worse effects on potential defendants and their behavior, the main focus of Nesson’s concerns, with adverse consequences for the community and plaintiffs.

Defendants need only play “defense,” an easy thing to do where science is concerned. By definition this is their role, but it has acquired added significance concerning scientific evidence. Implicit scientific burdens and standards of proof reinforce the defense position (as it does in the criminal law, but with a much different social effect). Within science the
burden of proof is typically on a scientist who would argue against the received scientific view, e.g., that a substance is toxic when toxicity had not previously been established. As already noted, the standards of proof can be quite high in the scientific community, and typically are more demanding than the tort law’s preponderance of the evidence standard. Thus, the standard of proof to clear the admissibility barrier can be much higher than the plaintiff’s ultimate standard of proof (although the original Daubert decision suggested this should not be a significant issue). Moreover, because scientists typically demand removal of considerable uncertainty before they come to scientific conclusions, any uncertainty (which can be considerable) assists defendants. Typically, defendants argue that too little is known to draw conclusions about toxicity, a view it is easy to find some scientists to support. It may be easier to persuade judges that there are too many uncertainties about the scientific evidence than that there is enough evidence to survive a Daubert review.

These structural issues, plus the possibility of winning on evidentiary grounds alone before trial, may heighten defendants’ temptations to distort the science needed for toxicity assessments because, if they are successful, the case is over. For example, it is reasonably common for defendants to insist that before scientific testimony is admissible it must be based upon statistically significant epidemiological studies, a kind of evidence that can be quite helpful, but is rarely available to evaluate the toxicity of a substance and not necessary according to most toxicologists for a toxicity judgment. Moreover, even though defendants in their own product research typically rely upon rodent studies, structure-activity relationships, mutagenicity studies, and even case studies to assess the toxicity of their products, they typically dismiss, downplay, or denigrate such evidence in admissibility hearings, simply because it is easy to raise various kinds of doubts about them that might appear persuasive to judges (and they appear to have had some success with such tactics).

There is even the temptation for defendants to act in ways that can corrupt the science. There has always been the temptation for litigants to fund studies that favor their own view of causation. There have been systematic efforts to mislead the public and regulatory agencies about what scientific evidence shows.29 Daubert has likely heightened this effect because it encourages judges to use publication in peer-reviewed journals as one consideration to assess expert testimony. Studies can be misleadingly designed to find the desired outcome. If such studies are then published, this distorts the scientific literature. Consequently, this literature is at some risk from misleading studies created for admissibility reviews.

To the extent that defendants are successful with some of the above strategies, this only reinforces their use in the future. One of Nesson’s concerns about a focus on the evidence is that it can convey a message about “crude risk calculation (‘estimate what you can do without getting caught’).”30 The Daubert admissibility requirements may well exacerbate these temptations in toxic tort cases.

Endnotes

21. Daubert at 588. The Court has said little about “reliable” evidence, but the decision suggests that the U.S. legal system is predicated upon witnesses having “firsthand evidence” about what happened, and that by analogy any exception to firsthand testimony must be as “reliable” as firsthand testimony would be. This analogy could pose problems if firsthand evidence refers to eyewitness testimony, which can be quite unreliable.
23. Id. at 60-82. Legal false positives are not identical to factual false positives, but there can be a comparatively direct relationship between scientific errors and legal mistakes.
28. Id. at 10-13.
30. Berger, id. at 1135-1140, 1147-1152, and Cranor and Eastmond, supra note 9 at 13-14, 45-48.
31. Nesson, supra note 1 at 1357.
32. Id. at 102 F.3d 194, 197 (5th Cir. 1996).
33. Some states do not require both general and specific causation, e.g., Illinois.
34. These plaintiffs’ and defendants’ scientific arguments are taken from expert reports filed with the court and acquired by the author from plaintiffs’ counsel.
35. Gerald Markowitz and David Rosner, Deceit and Denial: The Deadly Politics of Industrial Pollution (Berkeley: University of California Press, 2002) (the author, Otto Wong, in 1991, under pressure from industry, publicly retracted some of his own scientific findings contrary to industry interests (id. at 230)).
36. David Eastmond, Professor of Toxicology, University of California, Riverside, personal communication.
37. Nesson, supra note 1 at 1357.
38. Nesson, supra note 1 at 1361. Occasionally, a high-profile case, such as the criminal case involving O.J. Simpson, exhibits how an extreme focus on the evidence can distort the acceptability of a decision.
39. Nesson, supra note 1 at 1367.
40. Id.
41. Id.
43. Nesson also shows how the jury, the judge and procedural devices can help convey a message about the verdict instead of the evidence (points not considered here).
44. Nesson, supra note 1 at 1369-1370.
47. Personal communication from a major plaintiffs’ expert.
49. Daubert may also have encouraged an attitude toward risks very much favored by the technical community (vs. the public’s conception of acceptable risk), which would further separate the public from its institutions. (Gillette and Krier, supra note 6.)
50. Markowitz and Rosner, Deceit and Denial, supra note 17, at 195-233.
51. Nesson, supra, note 1 at 1362.
Legislators, regulators, and judges attempt to create factfinding processes that integrate both epistemic and non-epistemic goals. Moreover, the rule of law requires that those factfinding processes be principled, equitable, and reasonably transparent. This complex endeavor therefore produces some of the best-documented examples of societal factfinding. This essay analyzes the major decisional elements in such factfinding processes, with attention to common sub-tasks, distinct decision makers, and points of divergence for institutional procedures and safeguards.

The fundamental point of this essay is that any factfinding process in a governmental institution is designed to balance the epistemic objective against relevant non-epistemic objectives. The epistemic objective is to produce findings of fact that are as accurate as possible and that are warranted by the evidence legally available to the factfinder. The non-epistemic objectives include many that are common across governmental institutions (such as procedural fairness to parties and administrative efficiency), as well as many that vary by institution and by area of law (such as public health, safety in the space shuttle program, low inflation, protection of labor unions, and corrective justice). Whenever participants in any factfinding proceeding use scientific witnesses and scientific information to help warrant conclusions, they do so within a process designed to achieve some particular mixture of epistemic and non-epistemic objectives. And that mixed design goal exerts an influence within subparts of the overall process—for example, in judicial “gatekeeping” over scientific testimony after the Daubert case, which I discuss below. My background hypothesis, which is not argued in the essay, is that all truth-seeking processes of human beings exhibit a similar blending of epistemic and non-epistemic aspects, although few collective factfinding endeavors are as carefully engineered as the factfinding processes in law.

Factfinding in law is always pragmatic, in the sense that it always occurs in a context in which governmental action is at stake. First, the range of possible actions that any particular governmental institution is empowered to take is usually well settled. Such an action might be issuing a final regulation, sentencing someone to prison, ordering compensation or payment of a fine, issuing a report, or making recommendations. Furthermore, the positive substantive law generally identifies which findings of fact are required in order to justify particular actions. For example, statutes and implementing rules identify the factual issues that the Environmental Protection Agency must establish before that agency can lawfully suspend or cancel an existing approval (“registration”) of a pesticide that leaves a residue in food. Judicial precedents establish the factual elements that a torts plaintiff must prove before the court can lawfully order a defendant to pay compensation. Legislative bodies, regulatory agencies engaged in rulemaking, and courts applying common law decide which factual predicates justify which actions, and those engaged in adjudication or enforcement know which factual findings are legally significant. The propositions at issue in any given proceeding determine the relevance of any evidence submitted for the record in that proceeding.

The epistemic objective in factfinding is to ensure that the factfinder will be accurate in declaring (“finding”) propositions to be true for purposes of the legal proceeding, and that the evidence that is legally available warrants or adequately supports those findings of fact. But the factfinding process cannot be divorced from the pragmatic context in which it occurs, and which justifies expending resources on the factfinding process. The non-epistemic objectives therefore influence what evidence is ruled to be “legally available,” when the factfinder should be allowed to find that the available evidence “adequately supports” a finding, what standard of proof the factfinder should use in selecting a finding, what kinds and levels of uncertainty are acceptable in factfinding, and so forth. Rules governing such decisions should always balance substantive policies, procedural fairness, and administrative efficiency against the epistemic objective. Some non-epistemic policies trump the epistemic objective (for example, the minimal demands of due process), while others only weigh in the balance, or only apply when the epistemic objective is unachievable. This essay will discuss a few of the major decisional nodes in judicial and administrative factfinding.

It is important at the outset to clarify the phrases “issue of fact” and “issue of law.” These are terms of art that indicate who the decision is, rather than the nature of what is to be decided. “Issues of fact” present decisions for the factfinder—which in a given proceeding may be a jury, a judge, an administrative hearing examiner, or a regulator. The factfinder is supervised in each case by a presiding judge or other official. “Issues of law” are decided in the first instance by the judge or regulatory official who presides over the factfinding proceeding, with an appropriate opportunity for further review. Further review of issues of law is almost always available, and incorporates various degrees of deference to the presiding judge’s or official’s ruling. In this essay, the neutral phrase “proposition at issue” refers to the content of a contested issue, which might be decided in a particular case through a finding by the factfinder (as an issue of fact) or through a ruling by the presiding or reviewing authority (as an issue of law). Participants in a legal proceeding can usually examine the applicable law and agree at an early stage on the list of propositions at issue. The entire process in which various decision makers resolve those issues, for the pragmatic purposes of the relevant legal context, is the factfinding process.

The remainder of this essay examines the structure of the factfinding process by discussing three major decisional roles, which I will call the province of the factfinder, the province of the presiding authority, and the province of the reviewing authority.

The Province of the Factfinder

The province of the factfinder has as its main task the evaluation of all the evidence produced and the declaration of which propositions will constitute the findings of the proceeding (called “verdicts” in the case of jury findings). In one sense, the factfinder makes these decisions with considerable freedom, and there is little “law” about how to accomplish this task. There are few if any rules about how to assess the credibility of a fact witness’s testimony; how much probative value to assign to a document or to an expert’s theory, or how to integrate (for example) the mathematical probabilities derived from DNA evidence with the likelihood that the investigating police department contaminated or planted that DNA evidence. But to say that there is little law about how the factfinder should reach decisions within the zone of factfinding discretion is largely a truism, if by “law” is meant decisional rules adopted by presiding and reviewing authorities. Such rules do channel the factfinder’s decisions, but those rules operate outside the province of the factfinder, and therefore outside the zone of factfinding discretion. The process of factfinding is a dynamic interaction between factfinder decision-making and constraining rules, with the general trend in many areas of law being in the direction of adding more constraining rules.

---

**Epistemic and Non-epistemic Aspects of the Factfinding Process in Law**

**Vern R. Walker**

Hofstra University School of Law, Hempstead, NY

---
Probably the largest category of constraining rules consists of substantive rules of law, although such rules are not always recognized as constraints on the factfinder. Consider a torts case in which the plaintiff alleges that the defendant acted negligently. A basic proposition at issue (and presumptively for the jury) is whether the defendant engaged in any conduct that is considered negligent. There are many judicial rules, however, that create necessary conditions under this factual issue. For example, many courts hold that unless the defendant knew or reasonably should have known about the risks created by a course of conduct, then the defendant could not have been negligent in engaging in that conduct. This rule therefore generates a disjunctive proposition about whether the defendant had “notice” (“knew or should have known”), which any plaintiff is required to prove in order to establish that negligence occurred. Substantive judicial rules can therefore generate new propositional issues to be resolved by the factfinder, yet at the same time impose new constraints on the discretion of the factfinder.

Another category of constraining rules, closely related to substantive rules, consists of definitional rules. In a civil or criminal trial, for example, the presiding judge instructs the jury concerning the issues of fact about which the jury must make findings. The presiding judge leaves undefined most of the words employed in those instructions, for definitions always lead to more words, and defining must end somewhere. The meanings of most words, therefore, are left to the factfinder to determine, on the basis of the factfinder’s background knowledge. With respect to certain critical terms, however, courts may adopt rules of definition. Using the example of negligence in tort law, judges routinely tell juries that the law defines “negligence” as “lack of ordinary care” and “a failure to use that degree of care that a reasonably prudent person would have used under the same circumstances.”4 The judge, however, probably will not define further such terms as “ordinary,” “degree of care,” and “prudent.” Taken together, the substantive rules and definitional rules identify for the factfinder those propositions that need to be resolved. And at the same time those rules place boundaries on the factfinder’s role, and to some extent may influence the factfinder’s decisions within those boundaries.

A third category of constraining rules addresses process aspects of the factfinder’s task. Chief among these are rules concerning the standard of proof and burden of persuasion that the factfinder is to employ. For every issue of fact to be decided, there is a standard of proof to be met before a finding should be made. Standards of proof describe for the factfinder the quality of support required between the available evidence and the finding. For most issues in civil cases, the standard of proof is a “preponderance of the evidence”: the factfinder is to make a finding that p if, but only if, the evidence supports p more than it supports p’s negation, not-p. For some issues of fact, the law imposes a “clear-and-convincing-evidence” standard of proof, while criminal cases employ the familiar “beyond-a-reasonable-doubt” standard of proof. In contrast to the standard of proof, the burden of persuasion instructs the factfinder as to which party loses if the evidence does not satisfy the standard of proof. For example, if an issue of fact is to be decided by a preponderance of the evidence, but the relevant evidence in the case is equally weighted as between p and not-p, then, if the plaintiff has the burden of persuasion on p, the jury must find not-p, against the plaintiff. Typically, rules of law allocate the burden of persuasion to some particular party, for every factual issue in the case.

A fourth category of constraining rule on factfinders consists of default inference rules. For example, a legal presumption is a rule of law that describes what inference the factfinder either must or may draw, once the factfinder finds some specified proposition to be true. An example is the presumption that a person missing for some fixed period of time is dead, absent evidence to the contrary. Presumptions may be either mandatory (“If you (the factfinder) find p to be true, then you must find q to be true”) or permissive (“If you find p to be true, then you may find q to be true”).

Yet a fifth category of constraining rules consists of relevant-factor rules – rules of law that prescribe which factors the factfinder either must or may consider in arriving at a finding. For example, in determining whether the conduct of an actor was negligent, the factfinder should take into account the magnitude of the risk involved and the utility of the act or of the manner in which the act was done.

There are undoubtedly additional types of rules that constrain factfinder discretion, but the categories mentioned above supply enough examples to illustrate the following point. Substantive rules of law, legal definitions, process rules, presumptions, and relevant-factor rules are all devices to structure the factfinder’s role in deciding the propositions at issue. Some rules provide the grounds for taking a propositional issue away from the factfinder altogether (as discussed in the next section). Some are best understood as explicit or implicit commands (“… you must …”) or permissions (“… you may …”) addressed to the factfinder. There is of course no guarantee that the factfinder will understand those instructions, let alone follow them. The surest safeguard, other than the good-faith efforts of the factfinder, is the system of oversight provided by the presiding and reviewing authorities. As discussed below, those authorities are additional decision makers, whose provinces of decision interact with the province of the factfinder in complicated ways.

The Province of the Presiding Authority

In a judicial proceeding, the presiding authority is a trial judge. In an administrative adjudication, it may an administrative law judge or a hearing examiner. In an administrative rulemaking, it is the presiding regulatory official, who may chair a commission or a board. In any case, the presiding authority is the decision maker who presides over the creation of the official evidentiary record and who oversees the participation of the factfinder. The presiding authority also decides in the first instance which legal rules are applicable in the particular proceeding, and decides whether and how to enforce the applicable legal rules. This may involve instructing the factfinder about the legal rules as it is appropriate to do so, or granting or denying the motions of participating parties.

In many factfinding proceedings, a single individual functions as both the factfinder and the presiding authority. In judicial proceedings, examples are “bench trials” (in which the judge hears and decides the case without a jury) and hearings on preliminary matters (such as hearings on motions to exclude expert testimony). In administrative adjudications, the administrative law judge generally decides the case without a jury. In all of these proceedings, however, the distinction between the functions of factfinder and presider is still vitally important. When the judge in a bench trial makes a finding as the factfinder, her decision is entitled to the deference due to any factfinder’s decision. When that same judge makes a ruling as a matter of law, however, her ruling is subject to the same scope of appellate review as any trial judge’s ruling on an issue of law. The fact that the same person often plays two roles does not cause the line between those roles to vanish. This duality of role is possible in large part because the reviewing authority enforces the distinction between issues of fact
and issues of law. The different scopes of appellate review for issues of fact and issues of law are discussed below, under the province of the reviewing authority.

Because the presiding authority decides all issues of law in the first instance (subject to appellate review), he or she can make rulings on every category of legal issue that was discussed above under the province of the factfinder. On the motions and arguments of participating parties, and occasionally on its own initiative, a presiding authority will decide which substantive rules of law, definitional rules, process rules, default inference rules, and relevant-factor rules are in force within the jurisdiction and applicable in the proceeding. Even in proceedings where the same person acts as both factfinder and presiding official, a party may move for rulings of law concerning propositional issues on which the same judge will ultimately make findings of fact. Just as in a jury trial, such rulings on issues of law may constrain the judge’s discretion as factfinder.

In addition to these types of rulings, a presiding authority must decide important issues of law concerning the evidence proffered by the parties. Some evidentiary rules are process rules—such as the rules governing the discovery of evidence (e.g., rules governing the taking of depositions, the production of documents, and the conduct of physical examinations). Other evidentiary rules are exclusionary rules, concerning which proffered items of evidence are inadmissible in the proceeding. A judicial example is Federal Rule of Evidence 702, which states the conditions under which expert testimony may be admitted in federal courts, and which the U.S. Supreme Court interpreted in its decision in Daubert. In deciding whether to admit an expert’s opinion that (for example) the plaintiff’s exposure in utero to a certain drug caused the plaintiff’s injury, a federal trial judge must first decide whether the empirical and theoretical basis for that opinion is sufficiently “reliable” and “relevant” to the case at hand, and may decide whether its probative value is “substantially outweighed” by the risk of misleading the factfinder. If the trial judge decides to exclude the testimony, then the jury will not hear that opinion and the judge as factfinder may not rely on it as evidence. In administrative proceedings, the rules governing discovery and admissibility of evidence generally are quite different than those in judicial proceedings. These differences are often due to differences in institutional structure and in the blend of epistemic and non-epistemic goals.

Another important type of ruling on evidence assesses (as a matter of law) the “legal sufficiency” of the totality of admissible evidence introduced into the case, and allocates to particular parties the burden of producing that evidence. On every proposition at issue in a judicial proceeding, there is a rule of law assigning to some party the burden of producing evidence sufficient for a reasonable factfinder to find the issue in that party’s favor. The party that has the burden of producing or coming forward with evidence (for short, the “burden of production”) must lose the contest over the proposition if that party fails to produce enough evidence—in the form of real evidence, documents, and testimony. The totality of admitted evidence that is relevant to any proposition at issue must meet minimal sufficiency requirements before the presiding judge will present the proposition to the factfinder. Rulings on sufficiency therefore create a threshold of reasonableness before the exercise of the factfinder’s discretion. For example, in a tort complaint alleging that the defendant negligently caused the plaintiff’s injury, the factfinder has the burden of production on a number of propositions, including the propositions that the defendant in fact engaged in negligent conduct and that the defendant’s negligence in fact caused the plaintiff to suffer some injury. If the plaintiff fails to produce what the courts consider legally sufficient evidence to support findings for the plaintiff on these issues, then the trial judge should decide those issues against the plaintiff “as a matter of law.” Although this example is from a judicial proceeding, administrative adjudications may have similar rules about sufficiency.

In judicial civil cases, parties can raise the issue of the legal sufficiency of the evidence at various times, through a number of different motions. First, a party can obtain such a ruling in a motion for summary judgment, before the trial begins. Second, a moving party can obtain a ruling at trial, after the non-moving party has had an opportunity to produce the evidence. The issue may be raised then by a motion for directed verdict or for judgment as a matter of law. Third, the motion can be renewed if the factfinder reaches an unfavorable finding, in a motion for judgment notwithstanding the verdict or a renewed motion for judgment as a matter of law. Each motion, brought at a different time in the proceeding, may require the moving party to establish slightly different factual predicates, but the essence of the argument is the same: that as a matter of law the non-moving party will fail, or has failed, to satisfy its burden of producing evidence that the law considers minimally sufficient.

It is also possible for a court to decide an issue as a matter of law for the party who has the burden of production, if the party successfully argues that the evidence produced is so overwhelming that any reasonable jury would have to find for that party. It is far more common, however, for a court to enter judgment as a matter of law against the party with the burden of production, because of a deficiency in the evidence. This is primarily because it is easier to devise and apply rules of law about when evidence is missing or deficient, than about when produced evidence compels an inference.

The Province of the Reviewing Authority

Just as the presiding authority oversees the work of the factfinder, the reviewing authority oversees the work of the presiding authority. In the case of judicial proceedings, the reviewing authority is an appellate court. In the case of an administrative proceeding, the reviewing authority may be either an administrative body or a court conducting judicial review of the final administrative action. The general rules governing the relative authority of the reviewer over the decisions of the presider are called the “scope” or standard of review.

When it comes to reviewing issues of fact, the reviewing authority is required by law to be extremely deferential. When the factfinder is a jury, strictly speaking, the appellate court does not review the verdict directly. Rather, any challenge to a jury verdict generally must arise first as a motion to the trial judge, and then any appeal challenges the trial judge’s ruling on that motion. This is not a matter of mere semantics. The scope of review that is due to the trial court’s ruling depends upon the nature of the motion made to the trial judge—which may be reviewed for “abuse of discretion” or reviewed “de novo,” as discussed below.

Direct appellate review of a finding of fact does occur when the presiding authority is also the factfinder. For example, in civil bench trials in federal court, the trial judge must make findings of fact on the basis of the evidentiary record. Appellate review of those findings of fact uses the “clearly erroneous” standard. That is, the appellate court is not permitted to substitute its own assessment of the evidence for that of the trier of fact, so long as it is not clear that the trier of fact has committed a mistake.

Judicial review of administrative factfinding is similarly deferential. In general, the reviewing court must respect the findings of fact of the administrative agency unless those findings are “arbitrary [or] capricious” or “unsupported by substantial evidence” in the record. In practice, these two standards may come to much the same thing.
The scope or standard of review for many other rulings of the presider is the “abuse-of-discretion” standard, which is also very deferential. As long as the presiding authority’s ruling is not “manifestly erroneous,” the reviewing authority has the obligation to let the ruling stand. Often the type of ruling at issue involves the appropriateness of applying a correctly stated rule of law to the particular circumstances of a specific case, and the reviewing authority, which does not observe those circumstances first-hand, is not the best decision. Abuse-of-discretion review is applicable, for example, to trial court decisions to exclude proffered evidence because it is potentially prejudicial or needlessly cumulative.23

Not all abuse-of-discretion review, however, can be justified on the rationale that the presider is in a better position to assess the circumstances than the reviewer is. The abuse-of-discretion standard of review applies to federal trial court rulings to exclude expert testimony for failure to satisfy Federal Rule of Evidence 702.24 In Daubert matters, however, appellate courts are often in as good a position as the trial court to decide questions about the reliability and relevance of proffered scientific testimony. In such situations, a more complicated justification, in terms of non-epistemic as well as epistemic policies, would be needed to justify abuse-of-discretion review instead of de novo review. Arguing the merits of such a justification, however, must remain beyond the scope of this essay.

With respect to “pure” issues of law, appellate review of trial court decisions or judicial review of administrative decisions is generally “de novo”: that is, the presiding authority’s rulings are subject to review without any deference to the presiding authority’s ruling.25 If the presiding authority stated a rule of law incorrectly, or applied it in such a way that it is clear that the presiding authority misunderstood the law, then the reviewing authority simply reverses such a ruling and orders an appropriate remedy (which may include sending the case back to the trial court for re-trial or remanding the matter to the administrative agency). An example is a ruling that the evidence produced by a party who has the burden of production is legally insufficient. The argument that a party who has failed to meet its burden of production raises an issue of law on which the moving party is entitled to de novo review by the appellate court. One justification for this in a judicial setting is that otherwise a trial judge would have undue power to deprive a plaintiff of her right to a trial by jury. A broader justification is to enable appellate courts to create rules about minimal sufficiency of evidence that they can then enforce uniformly across all trial courts in the jurisdiction.

In general, therefore, reviewing authorities do not have the power to make decisions that are reserved to factfinders, and have precisely defined powers to set aside decisions made by presiding officers. Unless a proposition at issue can be resolved as a matter of law, if an appellate court decides to set aside a verdict or finding that is essential to the case, then the case must be sent back to the trial court to be re-tried. Similarly, a reviewing authority generally does not have the power to make administrative decisions in the place of the agency. If the reviewing court vacates an administrative agency’s action because it was arbitrary or capricious, then the matter must be remanded to the agency for further action.

Conclusion
Factfinding in a legal context is a highly structured process in which there are distinct decision-making roles. Each of the three principal roles (factfinder, presiding authority, and reviewing authority) has a significant zone of discretion in which to operate, yet decisions within that zone are constrained by decisions within the other two provinces. The factfinder’s zone of discretion centers on making findings about issues of fact, but rules of law often take propositional issues away from the factfinder and give them to the presiding or reviewing authority. The reviewing authority enjoys the greatest discretion in deciding what the rules of law are, but this discretion is constrained by rules identifying issues of fact reserved for the factfinder to decide. Of course, the discretion of the reviewing authority is also constrained by the legislature (through statutes), by administrative rulemaking, and by higher reviewing authority. Between the factfinder and the reviewing authority lies the zone of discretion of the presiding authority, which is least constrained in decisions about how to apply the established legal rules to the particular circumstances of the individual case.

The strategy behind factfinding in law is not so much to establish a methodology or logical model of inference to use in assessing the quality of factfinding outcomes, as to maintain a dynamic process of rule-governed decision-making, through which (it is hoped) reasonable decision makers will come close enough to achieving the epistemic objective over time. In deciding either to establish or to apply each of these rules, there is a persistent need to balance epistemic and non-epistemic aspects of the factfinding process, so that the process adequately serves its pragmatic function within the governmental institution. There is a great deal of theoretical work to be done in deciding the proper balance for each type of decision, by each type of decision maker, and for each type of propositional issue.

Endnotes
*Professor Walker received a Ph.D. in Philosophy from the University of Notre Dame (1975) and a J.D. from Yale University (1980). He wishes to acknowledge the research support of Hofstra University in preparing this essay.

1. Science is perhaps the leading contender to law in this regard, although scientists might spend fewer resources than government does in overseeing the communal enterprise of scientific factfinding. Indeed, much of the oversight of scientific factfinding is performed by government – through such institutions as the National Institutes of Health, the Centers for Disease Control and Prevention, the Environmental Protection Agency, the Food and Drug Administration, and so forth.


4. E.g., supra note 3.
17. This essay does not discuss the distinctions among different layers of appellate review, between appeal as of right and by writ of certiorari, or between appeals and certified questions. While such procedural distinctions add complexity to the factfinding process, the general theme of this essay does not require examining that complexity.


19. Id.

20. James et al., supra note 6, at § 12.9, pp. 668-74.


23. Federal Rule of Evidence 403; Mueller and Kirkpatrick, supra note 9, § 4.9, pp. 174-75.


25. I leave aside here such complications as the U.S. Supreme Court’s doctrine announced in Chevron, U.S.A. Inc. v. NRDC, 467 U.S. 837 (1984) (announcing a deferential approach to an agency’s interpretation of a statute that Congress has entrusted to the agency to administer).

---

**Recent Articles of Interest**

**— Abstracts —**


It has been argued in recent years that in a democracy, which admires political equality, the burden of taxation should be spread, not equally, at least proportionally, over the population. Democracy, it is said, requires a flat tax. Cavanaugh maintains to the contrary that “shifting the tax burden to the wealthy [i.e., progressive taxation] is ...not only consistent with, but necessary for, strongly held democratic ideals.” Cavanaugh offers a largely historical argument that outlines the role of taxation in Athenian democracy and the influence of the Greek model on the Founding Fathers.

In order to enable most (free and male) Athenians to participate in government, they were freed from property qualifications and indeed received pay for public service, which included holding office and sitting on juries. At the same time most of the burden of financing government fell to the wealthy, especially the top one percent of the population. These mechanisms promoted democracy, Cavanaugh argues, by minimizing the overt political advantages of the wealthy, though through their contributions the wealthy did achieve preeminence and influence. Indeed, according to the author, it is important to note that in the only example of a functioning democracy in Western history the burden of taxation was not equal, whereas in all other political societies, such as the Persian Empire, Egypt, and the Roman Empire, taxation has been more nearly equal or proportional and yet government has been decidedly non-democratic. The Greeks, in contrast, “associated uniform assessments with tyranny.”

Cavanaugh argues that the Athenian model was very much on the minds of those who established our own country and that the design of American democracy benefits from the example of Athenian democracy in being a composite of egalitarian and aristocratic elements. In the American context this is achieved through representative democracy. Representation responds, on the one hand, to the difficulty of constructing a manageable democracy in a vast territory. But at the same time it constitutes an aristocratic outlet. The Founding Fathers rejected the idea of an aristocracy of birth, but believed in the idea of a natural aristocracy, that is, of ability, which was identified with men of property. Representation “permitted the limitation of popular participation and ultimately elevated the interests of those whose wealth derived from property, not labor.” But if the wealthy were “to achieve greater power than the equality consistent with pure democracy otherwise permitted,” the counter must be to shift to the wealthy the burden of funding government. In support Cavanaugh cites a letter of Jefferson’s in which he says “Another means of silently lessening the inequality of property is to exempt all from taxation below a certain point, and to tax the higher portions of property in geometrical progressions as they rise.”

---


The completion of the Human Genome Project threatens real harm in discriminatory treatment in employment, along with the potential benefits of predictive genetic testing. Individuals determined through testing to be likely to develop debilitating illness might be denied employment or lose current employment with employers reluctant to risk substantial health care costs. Only minimal protections currently exist in Federal regulations, along with Constitutional provisions and the Privacy Act of 1974.

Current theoretical disputes include how best to conceptualize these issues as “discrimination” and which right is paramount, whether privacy or equal opportunity. Problems with reliance on privacy rights include identification of the precise nature of these rights, administratively realistic means of providing important protections, and competing concerns which arguably override privacy.

The authors of this very comprehensive article urge instead “an equality-based protection similar to the protection that exists for race and sex discrimination.” This approach builds on well-established anti-discrimination statutes and regulations at the Equal Employment Opportunity Commission (EEOC), as well as the Americans with Disabilities Act (ADA).

The authors show that one important population group is unprotected from genetic discrimination, under either the existing ADA or proposed genetic anti-discrimination laws, viz., “presymptomatic individuals with genetic anomalies that may never be expressed or, if expressed, may not manifest as unmitigable functional impairments.” However, the authors note, although numerous court decisions have limited the groups of persons protected by the ADA, some leave open the inclusion of presymptomatic and asymptomatic persons and, thus, broader protection based on genetic testing. The authors also argue that transactional costs will not significantly increase by broadening protection against genetic discrimination.

Slippery slope arguments, Volokh maintains, are not well understood. In their simplest form (“Oppose this law because it starts us down a slippery slope”), they are too abstract to be rationally persuasive. It is always appropriate to answer “Why will a slippery slope happen here when it hasn’t happened elsewhere?” Nevertheless, slippery slope arguments cannot be ignored, for though decision A might not, as a matter of logic, lead to B, it might materially increase the probability that, for psychological or political reasons or cost/benefit reasons, B comes about. One must ask, then, how much he or she likes A, dislikes B, and in particular what the mechanisms are by which B might come to pass, and the likelihoods of B’s occurring. Volokh’s article examines such mechanisms. He is particularly interested in showing that there is no single or fundamental mechanism by which “slippage” from A to B might occur.

Volokh describes and gives examples of a number of mechanisms by which acceptance of one policy might lead to another. For example, on a given issue some voters might be opposed to A in principle, others might accept A but reject B, still others might accept B in principle, but the deciding factor in a vote might be cost. Thus, if establishing A lowers the cost of B, it increases the chance that B will be approved. Gun registration might thus make gun confiscation more likely; installing video cameras on street lamps to deter street crime lowers the cost of tracking everyone with face-recognition software; a successful government-run, nonreligious, school-choice program might satisfy enough people that government can manage such programs and lead to its extension to religious schools. As another example, consider a court’s making a decision that certain speech (for example, racial epithets) may be banned. The court must give a general application of an existing legal rule (the obligation to follow precedent).” This effect is enhanced if the justification is vague, because it starts us down a slippery slope, because B follows from A as an application of an existing legal rule (the obligation to follow precedent).” This effect is enhanced if the justification is vague, or if, though the court crafts a justification that is intended to prevent extension to B, future courts are not convinced there is a real distinction.

Slippery slopes are real, not necessarily irrational, and must be confronted, especially when A is desirable but might not be adopted because of fears that B might follow. There are, Volokh argues, some tools available for preventing this. His main theme, however, is that it is important to understand in the particular case exactly how there might be slippage from A to B, so that one can make a good decision whether to oppose A because of B, or, on the other hand, support A by devising a way to make B less likely.


Randal Dworkin’s Sovereign Virtue (Harvard University Press, 2000) brings together articles on equality published between 1981 and 1990 and supplements these with more recent elaborations and clarifications. In “Sovereign Virtue Revisited,” the final article in an issue Ethics devotes to critical reflections on his book, Dworkin responds both to these reflections and to arguments presented elsewhere.

The sovereign demand on a political community, Dworkin contends, is that it treat every citizen with equal respect and concern. This applies to the structure of its economic institutions and to the conceptions of freedom, community, and political democracy that shape its laws and policies. Dworkin’s discussion of “equality of resources,” the account of distributive justice that he endorses, receives the bulk of the critical attention in assessments of the book and thus in Dworkin’s response here.

Dworkin’s article addresses, among others, the frequent claim that equality of resources is unforgiving, refusing community aid to those who have made disastrously bad decisions. It also takes up the charge that this account of distributive justice demands leveling down so that those who suffer grievous loss or injury do not envy the resources of others and responds to the claim that it differs from Sen’s capabilities approach in important ways that ought to be acknowledged. In the process of answering these and other objections, Dworkin clarifies or emphasizes: (1) the nature and role of his hypothetical insurance device; (2) the consistency of this device with the supposition, central to equality of resources, that we should pay the true costs of our choices; (3) the appropriate version of his envy test and his view that it is to be applied before and not after certain risks have been realized; (4) more generally that equality of resources demands that people be equally situated with respect to risk and not once uncertainty regarding risk has been resolved; and (5) the importance of asking what equality of resources requires under real-world circumstances and not in the abstract.


A lawmaking system, Barnett says, is legitimate only if it creates laws that citizens have a moral duty to obey, and a constitution is legitimate only if it creates such a lawmaking system. But when, and why, is there a duty to obey? In Barnett’s view, the legitimacy of the Constitution does not rest on the consent of “We the People,” for a group can bind itself by its consent only if the consent is real and unanimous. But none of the explanations frequently given to explain how each of us has consented is credible. Neither voting, nor residence, nor the consent of the founders, nor acquiescence explains the duty to obey the law. Nor does the receipt or acceptance of benefits. Nor does hypothetical consent.

Indeed, the notion of “We the People” is misunderstood in ways that make it potentially dangerous. The Constitution is not, for the most part, binding on citizens. It is binding on government; it is the laws made under the Constitution that bind citizens. What “We the People” does, in establishing the Constitution, is limit the lawmaking officials, who are different from “We the People.” Despite this, the legislature itself is commonly identified with “We the People.” The fiction is indulged that “We the People” make the laws that bind us. But if this is the case, there is no explanation of how dissenters and nonvoters are bound, and if consent exists in such cases, then there is consent to anything the majority does. Further, constitutional limitations, on this understanding, appear to be contrary to the sovereignty of “We the People”; this is the foundation of the countermajoritarian worry about judicial review.

Nevertheless, there is constitutional legitimacy, there is a duty in conscience to obey laws properly passed by majorities.
The argument can only be summarized briefly here: In some limited circumstances (not including the most relevant—a lawmaking authority for a large geographical area) unanimous consent is possible. But consent can be binding only if there is a right to *withhold* consent. In general, then, rights (natural; human) must precede law. In a relatively large geographical area, unanimous consent is unlikely or impossible, but majority lawmaking is legitimate if the laws passed are needed to protect people’s background rights (this is why such laws are obligatory) and do not violate the background rights of those on whom they are imposed (this requirement explains why consent is not needed). Under these conditions, there is assurance that the laws are not unjust, and a constitution establishing such a lawmaking system qualifies as legitimate. There can, of course, be difference of opinion about exactly what the background rights are.

**RECENT BOOKS OF INTEREST**


