

Two Concepts of Reliability¹

Dale A. Nance, J.D.*

* Professor of Law, Case Western Reserve University

In *Daubert v. Merrell Dow Pharmaceuticals, Inc.*² and *Kumho Tire Co. v. Carmichael*,³ 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.

1. 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.⁴ "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."⁵ 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.⁶ 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.”⁷

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.”⁸ 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.⁹

2. 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].¹⁰

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.¹¹

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. This 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 must 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*.¹²

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.¹³

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 understood 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 some types of forensic science, for which 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?”

3. 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.¹⁴

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 first entails a much larger degree of epistemic paternalism than the second. 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, 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.¹⁵ In contrast, it is generally agreed that the burden is on the proponent of expertise to establish (sufficient) reliability under 702.¹⁶ 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."¹⁷ 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.¹⁸ If anything, jurors may excessively *discount* complex expertise.¹⁹

Until we have more evidence of jury credulity, 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*.²⁰ 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.²¹ 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.

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.

¹ This article was originally published in *APA Newsletter on Philosophy and Law*, Fall 2003. Copyright 2003 Dale A. Nance; all rights reserved. The present essay is a modified and condensed version of another paper, *Reliability and the Admissibility of Experts*, 34 *Seton Hall L. Rev.* 191 (2003).

² 509 U.S. 579 (1993).

³ 526 U.S. 137 (1999).

⁴ Fed. R. Evid. 402.

⁵ Fed. R. Evid. 401.

⁶ Fed. R. Evid., Art. VIII.

⁷ Fed. R. Evid. 702, advisory committee’s note to 2000 amendment (emphasis added).

⁸ 509 U.S. at 593.

⁹ See Samuel R. Gross, *Substance & Form in Scientific Evidence: What Daubert Didn’t Do*, in *Reforming the Civil Justice System* 234 (Larry Kramer ed. 1996) (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 David L. Faigman, David H. Kaye, Michael J. Saks, & Joseph Sanders, *Science in the Law: Standards, Statistics and Research Issues*, 1B3.8 at 59-60 (2002).

¹¹ 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.

¹² *Id.* at 596-97.

¹³ For elaboration on this point, see Neil B. Cohen, *The Gatekeeping Role in Civil Litigation and the Abdication of Legal Values in Favor of Scientific Values*, 33 *Seton Hall L. Rev.* (forthcoming, 2003).

¹⁴ See Dale A. Nance, *Naturalized Epistemology and the Critique of Evidence Theory*, 87 *Va. L. Rev.* 1551, 1554-63 (2001).

¹⁵ This is the practical effect of the rules 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?

¹⁶ See Faigman et al., *supra* note 9, ' 1B3.1.2.

¹⁷ 509 U.S. 595-96.

¹⁸ See, e.g., Neil Vidmar & Shari Seidman Diamond, *Juries and Expert Evidence*, 66 *Brook. L. Rev.* 1121, 1180 (2001) (reviewing the literature).

¹⁹ See, e.g., Dale A. Nance & Scott B. Morris, *An Empirical Assessment of Presentation Formats for Trace Evidence with a Relatively Large and Quantifiable Random Match Probability*, 42 *Jurimetrics J.* 403 (2002) (reporting jurors' systematic undervaluation of the probability of guilt relative to Bayesian norms).

²⁰ 509 U.S. at 595.

²¹ See Dale A. Nance, *The Best Evidence Principle*, 73 *Iowa L. Rev.* 227 (1988) (arguing that such preferences provide better explanations of many admissibility rules than the jury distrust explanations so commonly encountered).