The *Daubert* Guidelines: Usefulness, Utilization, and Suggestions for Improving Quality Control

Lisa R. Fournier*

Washington State University, United States

Federal courts and many state courts have adopted the *Daubert* standard for the admissibility of expert witness evidence. I briefly describe the *Daubert* guidelines, their utility, and current use. I also share some thoughts on the implications of the *Daubert* guidelines for improving peer review as a method of quality control in the courtroom, alerting the courts to factors that can threaten the internal validity of research, and increasing the frequency in which experts accurately apply valid, reliable research to support their opinions.

**Keywords:** Daubert, Expert witness, Junk science, Reliability, Internal validity, Peer-reviewed publication, Error rate, Expert shortage, Bias

Federal courts and many state courts have adopted the *Daubert* standard for the admissibility of expert witness evidence. The main goal of *Daubert* is to tighten the rules of expert testimony and to eliminate the reliance on junk science in the courtroom (Bernstein & Lasker, 2015; Faigman, 2013). The *Daubert* requires that judges make a preliminary assessment as to whether the methods and principles underlying an expert’s opinions are valid and reliable, and whether the expert is reliably applying these methods and principles to the facts of the case (Faigman, 2013; Gatowski et al., 2001). Thus, judges serve as “gatekeepers” in terms of what expert witness opinions are admissible in the courtroom and are heard by the jury (e.g., Brown & Davis, 2014; Faigman, 2013). The Supreme Court provided judges with four guidelines under *Daubert* designed to increase the scientific rigor of expert evidence admitted in the courtroom. The *Daubert* guidelines are referred to as testability, peer-review and publication, error rate, and general acceptance (Faigman, 2013; Gatowski et al., 2001). I briefly describe the *Daubert* guidelines, their utility, and current use. I also share some thoughts on the implications of the *Daubert* guidelines for improving peer review as a method of quality control in the courtroom, alerting the courts to factors that can threaten the internal validity of research, and increasing the frequency in which experts accurately apply valid, reliable research to support their opinions.

**The Four Daubert Guidelines**

The *Daubert* guidelines are intended to increase the scientific rigor of evidence used in the courtroom, and they are generally consistent with the guidelines for accuracy and objectivity in science. Testability refers to whether an expert’s claim, theory or technique has been (or can be) empirically tested, which is a core feature of scientific hypothesis testing (Faigman, 2013; Gatowski et al., 2001). Peer-review and publication addresses whether the evidence presented by an expert to support a claim, theory or technique is based on peer-reviewed publications in which internal validity and statistical reliability has been evaluated (e.g., Faigman, 2013; Gatowski et al., 2001). Error rate refers to the known or potential error rate of a particular scientific

---

**Author Note**

* Please note that this paper was handled by the current editorial team of JARMAC.

1 Under *Daubert*, judges can also restrict expert testimony that covers information thought to be already within the ken of the jury; however, this paper only focuses on the role of *Daubert* in the elimination of junk science.

---

* Correspondence concerning this article should be addressed to Lisa R. Fournier, Psychology Department, Washington State University, Pullman, WA 99164-4820, United States. Tel.: +1 5093354415. Contact: lfournier@wsu.edu
method or technique underlying the expert’s claim and whether
the method or technique is internally valid, statistically reliable,
and externally valid to the current case (Faigman, 2013; see also
Gatowski et al., 2001; Meixner & Diamond, 2014). Finally, general acceptance refers to the consensus of experts
in the field (e.g., scientific community) regarding the expert’s
 testimony (Faigman, 2013; Gatowski et al., 2001).

Are the Daubert Guidelines Useful?

The Daubert guidelines can be useful in eliminating expert
opinions that have no empirical scientific basis (Bernstein, 2013;
Bernstein & Lasker, 2015; Meixner & Diamond, 2014). Some argue that testability and error rate are the two most important
guidelines due to their direct relevance in assessing the scientific
validity of the research evidence relied upon (Gatowski et al.,
2001; Haug & Baird, 2011; Meixner & Diamond, 2014). The
two remaining guidelines, peer-review and publication and general acceptance, provide information concerning the scientific
validity of the cited research and reliability (consistency) of the
cited phenomena or theories over time, respectively. However,
considered in the absence of testability and error rate, these
guidelines are less diagnostic for at least two reasons. First,
the quality of peer review and publication varies, and hence
peer review alone cannot guarantee a proper vetting of scientific
validity. Second, general acceptance does not necessarily mean
acceptance in a formal scientific field or acceptance based on
formal scientific evidence. If it can be clearly determined that
both peer review and publication and general acceptance were
based on the scientific validity of the cited research, then the
four guidelines together represent a more thorough evaluation
of the validity and reliability of an expert’s opinion—with one
caveat: the Daubert guidelines cannot easily reveal whether an
expert is ignoring important, contrary evidence.

If the Daubert guidelines are applied, admissibility of expert
opinions based on empirical scientific evidence should be
favored compared to those based only on reason or tradition
(precedent; see Bernstein & Lasker, 2015; Faigman, 2013;
Schwartz & Silverman, 2006). Eliminating expert opinions that
have no scientific basis should reduce unreliable and biased
expert testimony and eliminate “junk science” from the courtroom (Bernstein, 2013; Bernstein & Lasker, 2015; Meixner &
Diamond, 2014). This is the utility of the Daubert guidelines. For example, in Kilpatrick v. Breg, Inc., the court excluded expert
testimony indicating that a particular agent caused a disease
because the expert relied on studies that only reported an associa-
tion between the agent and disease—indicating that the expert’s
causation conclusions were not based on a reliable methodology (Rutner & Bach, 2015). In Rider v. Sandoz Pharmaceuticals Corp., an expert used several epidemiological studies with sta-
tistically insignificant results to support his testimony of a causal
link between Paralodol and stroke; as a consequence, the expert’s
testimony was rendered unreliable and inadmissible (Rutner &
Bach, 2015). For other examples of inadmissible expert testi-
mony based on lack of statistical evidence see Rutner & Bach

The courts have also restricted experts from providing
opinions that go beyond data supported by their and others’
investigations. For example in United States v. Horn, a neu-
ropsychologist was not permitted to testify about toxicological
causation of a neuropsychological illnesses because he was not
an expert in toxicology, and “qualified toxicologists” were not
permitted to approximate blood alcohol concentrations based
on the review of results from police-administered sobriety tests
(Fradella, O’Neill, & Fogarty, 2004, p. 358). In addition, several
common forensic techniques with little or no scientific evidence
supporting their validity and reliability were admissible in the
courtroom for decades due to their general acceptance
in the field (e.g., fingerprint analysis, handwriting analysis, and
comparative bullet lead analysis)—now, these techniques are
no longer accepted by some courts (e.g., State v. Rose cited in
Commonwealth cited in Iannecci, 2015, p. 198, respectively) and
often testimony utilizing these tests is restricted (e.g., United
States v. Green cited in Giannelli, 2010, p. 63; see also cases
cited by Fradella et al., 2004; but see contrary examples in
Giannelli, 2010).

The Daubert guidelines can undoubtedly be applied to expert
witnesses in scientific fields, but it is less obvious how they
apply to experts who are not scientists that have skilled or expert
knowledge based on experience (e.g., police officers, clinical
psychologists, medical doctors, social workers). Faigman, Kaye,
Saks, and Sanders (2000) suggest that expert testimony governed
by skill or experience is admissible when “no more is available
and no more can be expected” (p. 650). For example, the court
may allow a police officer to testify, based on experience, as
to whether officers may experience fatigue after a long shift
in the field, but not whether officer fatigue causes an increase
in inattentive blindness—the latter would require an opinion
backed up by scientific evidence. Importantly, because Daubert
requires that experts justify their claims based on reliable data,
testimony based on scientific evidence should trump testimony
based on experience alone, especially since scientific data shows
that reasoning from experience can lead to incorrect conclusions.
Unfortunately, as discussed below, the Daubert guidelines are
not always applied by the courts.

Are the Daubert Guidelines Actually Used?

The Daubert standard is required in all federal courts and in
about one-third of state courts. However, there is no uniformity
in applying the Daubert guidelines, and some federal and
state judges have ignored them altogether (e.g., Bernstein,
2013; Bernstein & Lasker, 2015; Fradella et al., 2004; Gatowski
et al., 2001; Giannelli, 2010). Many federal and state courts
“have allowed expert witnesses to ignore the scientific method
and consider only studies or other data that support their side’s
case” (e.g., Kuhn v. Wyeth, Inc. cited in Bernstein & Lasker, 2015,
p. 34). Some federal courts have concluded that “weak expert
testimony should be admitted for jury consideration” (e.g., Mil-
ward v. Acuity Specialty Products Group, Inc. cited in Bernstein
& Lasker, 2015, p. 41) and that the court may only exclude
“nonsense opinions” (e.g., Bonner v. ISP Techs., Inc. cited in
Bernstein & Lasker, 2015, p. 23). Some federal courts have ruled that “inadequacies in expert testimony are a matter of weight, not admissibility, and admissibility should be decided by the jury” (citation from Federal Circuit opinion in Liquid Dynamics Corp. v. Vaughan Co., 2006, Co. 132. Id. at 22; referenced in Bernstein & Lasker, 2015, p. 24). Some federal courts have ruled that the court must analyze the reliability of the expert’s methodology, but the jury should decide whether the method is applied appropriately to the expert’s opinion or to the facts in the case (e.g., City of Pomona v. SQM N. Am. Corp. cited in Bernstein & Lasker, 2015, p. 7). Finally, some courts continue to permit forensic expert evidence, such as fingerprint identification, based on general acceptance in the field (e.g., United States v. Llera Plaza II, United States v. Mitchell, United States v. Crisp, and United States v. Sullivan cited in Giannellie, 2010, p. 62), even though scientific evidence shows that this technique is unreliable (see reviews by Giannellie, 2010; Epstein, 2001). These are just a handful of examples in which the Daubert guidelines are disregarded (see also Bernstein, 2013; Bernstein & Lasker, 2015; Fradella et al., 2004; Gatowski et al., 2001; Giannellie, 2010). However, some courts are ordering new trials, and some are reversing sentences due to expert testimony based on unreliable science (e.g., Iannace, 2015).

Variability in the application and weighting of the Daubert guidelines across the courts suggests that at least some guidelines are not well understood (Bernstein & Lasker, 2015; Gatowski et al., 2001). In fact, one study (Gatowski et al., 2001) showed that state judges reported more difficulty understanding the error rate and testability guidelines compared to the guidelines of peer review and publication and general acceptance. Another study (Haug & Laird, 2011) showed that state and federal judges less frequently applied the guidelines of error rate and testability to cases compared to the guidelines of peer review and publication and general acceptance. In contrast, a recent study (Meixner & Diamond, 2014), sampling only from federal court cases, found that error rate was discussed to a greater extent than other guidelines—at least when error rate was defined as “an assessment of validity” as opposed to a “quantitative error rate of a method.” These latter findings suggest that federal judges, who are mandated to use Daubert, may have a better understanding of the importance of error rate than earlier studies indicated. It also may be the case that compared to state judges, federal judges better understand the importance of error rate. Thus, while the variability in the application and weighting of the Daubert guidelines may be partly attributed to the comprehension of these guidelines, it may also be due to an outright refusal to apply some or all of these guidelines, even by federal judges who are mandated to do so (Bernstein & Lasker, 2015).

How Can We Improve Peer Review as a Method of Quality Control in the Courtroom?

Ideally, research published in peer-reviewed publications should represent work vetted for internal validity and statistical reliability by other experts in the field. However, while work published in peer-reviewed journals makes it more likely that such vetting occurred, it cannot guarantee it. First, no scientist is omniscient. Conclusions drawn in peer-reviewed publications are often challenged, and alternative hypotheses are proposed and tested—this is how scientific inquiry works. Second, peer reviewers are fallible, and factors threatening internal validity of a study can be missed. Third, fraud does exist and it can be difficult to determine whether data have been fabricated. Fourth, whether the results found in one study are replicated across other studies and laboratories cannot be evaluated based on a single publication. Fifth, there is a bias in publishing significant versus insignificant effects, and this bias in combination with a lack of direct replication studies can lead the field to false conclusions. Finally, a peer-reviewed publication does not guarantee that peer reviewers were experts in the topic under investigation or that they had sufficient knowledge to evaluate the internal validity of the study. Peer-reviewed journals vary in regard to the scientific expertise of the reviewers and the rigor dedicated to determining whether a study’s methods and conclusions are empirically valid. For example, a peer-reviewed journal with editors and reviewers who are professionals in the law enforcement field with no formal, scientific research training (e.g., police or parole officers), could not be expected to accurately evaluate the internal validity of a study on stress and decision-making in police officers. Thus, peer-reviewed publication cannot stand alone as a measure of quality control, truth, or an indicator of scientific evidence.

However, we can improve quality control of peer reviewed publications relied upon in the courtroom by helping the courts distinguish between scientific vs. non-scientific publications. In order to distinguish a scientific from a non-scientific publication, the peer review process for all journals should be published and made available to the public on journal websites. All journal websites should include the journal’s topic or research focus, the members of the editorial board (name, institution, research area), the professional organization(s) the journal is associated with (e.g., American Psychological Association; Human Factors & Ergonomics Society), the criteria for acceptance, the rejection rate, the selection process and number of peer reviewers, and whether optional anonymity of authors and peer reviewers is maintained. This information, in part or whole, is available for most scientific, peer-reviewed journals (e.g., Journal of Experimental Psychology: Human Perception & Performance); but is not necessarily available for many non-scientific, professional journals (e.g., Law Enforcement Executive Forum). For scientific journals, the scientific merit should be a main component in the criteria for acceptance, the rejection rate should be typically high (consistent with quality control), at least 2–3 peers should review the manuscript, and anonymity of the authors and reviewers should be an option (to reduce potential bias). Importantly, peer reviewers should be selected based on their

---

2 I purposely excluded the journal’s impact factor as one of the criteria to judge the quality of peer review for the journal. The impact factor is greatly influenced by the scope of topics covered in a journal and does not necessarily reflect the rigor of peer review; a journal with a narrow scope tends to have a lower impact factor than a journal with a broader scope.
expert knowledge of the construct (e.g., specific cognitive process) under investigation, the method or technique employed, and the measures and analyses used to draw specific conclusions about the construct under investigation. These are minimal criteria for peer-reviewed, scientific journals, and reliance on published evidence in journals that do not have these criteria should be excluded as scientific evidence. Unfortunately, there has been a recent proliferation of “for profit” journals posing as scientific journals that do not meet the criteria outlined above, and hence it is very important for the courts to be alert to these criteria.

How Can the Courts be Alerted to Factors That Threaten Internal Validity of Research?

The courts can be alerted to these factors in a number of ways. To educate judges on various scientific topics including the scientific method, programs and conferences are available (e.g., see Haack, 2015, p. 57). For example, the National Courts and Science Institute has a Center for Basic and Continuing Judicial Science Education that is designed to promote annual conferences and a judicial certification in “scientific method, tools and measures.” In addition to these efforts, general principles of the scientific method could be included as part of the educational curriculum in law schools (Merlino, Richardson, Chamberlain, & Springer, 2008; e.g., see Harvard Law School curriculum). Another option is for the court to appoint a “nonpartisan” expert witness (with the prerogative of soliciting nominations from each party) who has expert knowledge in the research domain, research methods, techniques, and statistics to evaluate the facts in the case and whether the evidence relied upon by other experts involved in the case is internally valid (and statistically reliable). This expert could advise the court and both parties of his or her findings (e.g., potential confounds, lack of baseline measurements, etc.), and could be called to testify and be cross-examined by each party (Charles, 2015, p. 947). While the idea of a community of experts needing to police itself may sound distasteful, it may be necessary to reduce junk science in the courtroom. Importantly, when confronted with a “battle of the experts,” testimony that can be supported by scientific evidence and can be shown to be reliable over time should be included, and testimony that cannot should be excluded (e.g., Bernstein, 2013).

How Can We Increase the Frequency with Which Experts Accurately Apply Valid, Reliable Research to Support Their Opinions?

As research scientists, we can continue to write review articles on various empirical phenomena, measures, and theories proposed to explain these phenomena. These reviews can identify unresolved questions or controversies and address issues of validity, reliability, and generalizability in the context of whether the field currently accepts or rejects a particular empirical phenomenon, theory or technique (e.g., see Chun, Golomb, & Turk-Browne, 2011). Issues concerning the strength of an effect, whether sufficient power is present across studies, the possibility of publication bias, and other factors that can enhance or reduce (or even eliminate) the effect could be included in these reviews. We can also directly address whether the empirical phenomenon, theory, or technique is ready for use in the courtroom and whether there is a potential for misuse in applying it (e.g., whether we can discriminate between intentional blindness and lying when one may be motivated to lie?). This is the purpose of this forum in JARMAC. [For another example, see Kassin et al.’s (Kassin, Tubb, Hosch, & Memon, 2001) survey on the general acceptance of eyewitness testimony research]. Ideally, such reviews would be updated at least every 5–10 years, particularly in research areas in which theoretical controversies exist and new empirical measures are being applied. Experts could refer to these review articles and commentaries on “general acceptability and readiness for the courtroom” in addition to relying on specific empirical articles to support their opinions.

In the courtroom, the reputation of the expert witness should be considered. That is, whether he or she has a record of using specious science or has a bias toward supporting a particular organization or industry (e.g., Friedman, Daynard, & Banthin, 2005). In principle, an expert witness should be an advocate for the truth and not an advocate for a particular organization or industry (Friedman et al., 2005). Giannelli (2010) noted that industry-sponsored studies were significantly more likely to reach conclusions in favor of the sponsor than non-industry studies. This suggests that industry-sponsored studies are often biased, and findings by the industry that are contrary to the industries’ agenda may not be reported or published (e.g., see Friedman et al., 2005). In addition, the training of the expert witnesses and whether they have a history (and a current history) of publications in peer-reviewed, scientific journals should be considered. Unfortunately, the courts often give equal weight to testimonies coming from expert witnesses with strong and weak scientific training and scientific publication histories. The risk here is that giving equal weight to an “expert” who is really not qualified as an expert can skew the evidence. This problem is exacerbated when the “expert” belongs to an industry or organization that may be biased to support a particular agenda.

One way to reduce this problem is to have all expert testimony undergo peer review as part of Daubert. To facilitate this review, experts should be required to provide a list of empirically-based, peer-reviewed references as examples that support their opinions. This way, reviewers (either court-appointed or experts hired by either side) could evaluate the validity and reliability of each expert’s opinions, as well determine if evidence that may weaken or contradict his or her point of view has been overlooked. This is particularly important when an expert is making a very specific claim (e.g., providing a time estimate to execute a particular action). In addition, providing a list of references allows the court to more easily evaluate whether the research cited to support one’s testimony is published in peer-reviewed, scientific journals. Unfortunately, this peer review process of expert testimony is not always done. Also, in cases in which peer review of expert testimony is practiced, experts are not necessarily required to provide a reference list in support of their testimony, which can make peer review difficult. While it would be impossible for most expert witnesses to provide an exhaustive list of all of the relevant literature relied upon, referencing
a subset of the literature should be sufficient to indicate whether there is empirical support for the expert’s testimony.

Finally, the court may consider using a court-appointed, “nonpartisan” expert to better warrant that the expert has “an allegiance to the truth” and will not “cherry pick” evidence to support the side of the plaintiff or defense (Charles, 2015). In cases in which it may be difficult to find a genuine, nonpartisan expert because the number of experts in a particular field is small (e.g., Haack, 2015), experts could be appointed who have a science background, and knowledge of research methods and techniques utilized in the cited studies. As noted earlier, these experts could also serve as peer reviewers of the evidence offered by other experts hired by the plaintiff and defense.

Increased application of the Daubert guidelines will likely increase the demand for expert witnesses in the courtroom. Unfortunately, the number of experts in some specialty fields is limited, and hence this increased demand can lead to an “expert shortage”. There is evidence that this is already occurring in some fields in psychology, and the “Innocence Project” recently sponsored a workshop at the 2016 American Psychology-Law Society meeting explicitly advertised as an effort toward recruiting new experts. The obvious problem with the shortage of experts is that less qualified experts will be, and are already, asked to testify (e.g., some “expert witnesses” are testifying as scientific experts who do not have a history of scientific publications, a history of research, or a graduate-level degree in the area of expertise they are professing). This shortage and the resulting problem of utilizing less qualified experts can work in the opposite direction of that intended by Daubert by allowing more junk science and unreliable testimony to slip through the courtroom for purview by the jury. Thus, the issue of how to handle the shortage of experts is of some urgency.

## Conclusion

The Daubert guidelines are important because not all scientific experts are reliable. As Haack (2015) posits: “some are honestly mistaken, some incompetent, some self-deceived, and some are outright dishonest” (p. 65). The Daubert guidelines are designed to help judges identify what opinions by expert witnesses are sufficiently reliable to be admissible in the courtroom and heard by the jury. Application of the Daubert guidelines could reduce the incidence of experts misusing seemingly valid and reliable methods to reach a predetermined result, biasing or confusing the issues, and/or misleading the jury with “scientifically cloaked evidence” (Bernstein & Lasker, 2015). Unfortunately, the haphazard application of the Daubert guidelines in the courts has undermined their utility. To increase their utility and application, we can help the courts distinguish between scientific and non-scientific peer reviewed publications, educate the courts on factors that threaten internal validity, and continue to write research reviews as well as critical essays regarding the general acceptance, controversies, and potential misuse of various empirical phenomena, theories, and techniques. In addition, the courts can be wary of experts who have a record of spurious science or may be biased due to their association with a particular organization or industry, they can consider the training and publication history of the expert witness, and they can require peer-review of expert testimony and appoint nonpartisan experts to serve the court. Since the establishment of Daubert and the Daubert guidelines, there is evidence that expert testimony in general has been more strictly limited by the courts (e.g., Faigman, 2013; Jurs & DeVito, 2014; Kafka, Dunn, Johnson, Cecil, & Miletich, 2002). Although encouraging, more courts need to utilize the Daubert guidelines to improve the validity of expert testimony and eliminate reliance on junk science. The courts cannot do this without the active support and advisement from the scientific community.

## Acknowledgements

I would like to thank Brian Dyre and Rebecca Craft for their comments on the initial draft of this essay.

## References


Received 30 April 2016; received in revised form 17 June 2016; accepted 29 June 2016