

## Expert Witness Standards Must Consider Peer Review Crisis

By **Jeffrey Gross and Robert LaCroix** (August 1, 2024)

—The recent amendment to Rule 702 of the Federal Rules of Evidence, effective since Dec. 1, 2023, was the rule's first substantive amendment in two decades.

It followed years of criticism that courts had not followed the rule or its intent regarding the standard of proof that applies when evaluating the reliability of expert testimony.[1]

A 2022 report from the Advisory Committee on Evidence Rules concluded that courts had been confused about the standard of proof, which led to the 2023 amendment.[2]

But the committee's narrow focus on the standard of proof ignored other inconsistencies in how courts apply Rule 702, and the U.S. Supreme Court's landmark 1993 decision in *Daubert v. Merrell Dow Pharmaceuticals Inc.*

One issue stands out: The *Daubert* decision lists, as one of several nonexhaustive factors for evaluating reliability, whether the expert's methodology "has been subjected to peer review and publication." [3] Just what to make of the peer-review factor has bedeviled courts ever since, as recent decisions following *Daubert* continue to illustrate.

For instance, the Zantac products liability litigation has yielded conflicting results in cases considering nearly identical expert testimony, as discussed in greater detail below.

And just last month, the Michigan Supreme Court, over a sharply worded dissent, held in *Danhoff v. Fahim* that the lower courts had focused too "strictly on plaintiffs' inability to support [their expert's] opinion with published literature," and reversed the decision to exclude the expert.[4]

Including peer review as a factor seemed uncontroversial at the time *Daubert* was decided. But between the *Daubert* decision in 1993 and today, the so-called replication crisis has upended how the scientific community views the reliability of studies.

The crisis began in the mid-2000s, after independent researchers discovered they could not replicate — that is, repeat — many experimental results. Researchers also could not reproduce the published results of earlier research using the source data. The crisis undermined public trust in scientific studies, including those published in peer-reviewed journals.

Many have described the crisis in stark terms: Alvaro de Menard, a participant in a Defense Advanced Research Projects Agency initiative on replication, who surveyed studies that could not be replicated, lamented that the crisis left him with "a sense of Lovecraftian awe at the sheer magnitude of it all." [5]

Given these revelations regarding the soundness of peer-reviewed studies, courts and



Jeffrey Gross



Robert LaCroix

rulemakers should reevaluate whether or when peer review should receive any weight when assessing the reliability of expert testimony.

### **The Replication Crisis and Peer Review**

Broadly speaking, "peer review" refers to the practice of many academic journals inviting critiques from other researchers in a study author's field before publication of the study.[6]

Theoretically, peer review can root out errors before publication. But that theory is questionable for many reasons. Reviewers may not have the right background to find the errors. Reviewers also may lack sufficient financial or reputational incentives to review rigorously. Or they may not want to challenge an author directly.[7]

In 2005, Stanford University professor John Ioannidis published a groundbreaking essay, "Why Most Published Research Findings Are False," which exposed what is now called the replication crisis.

He found that, "for most study designs and settings, it is more likely for a research claim to be false than true." He also concluded that, "for many current scientific fields, claimed research findings may often be simply accurate measures of the prevailing bias." [8]

Further studies and analyses concluded that peer review did not prevent serious errors.[9]

The revelations were deeply unsettling. Research on replication studies has found that fewer than 30% of studies in social psychology, and approximately 50% in cognitive psychology, were able to be replicated.[10]

Even among the studies published in some of the highest-profile American scientific publications, *Nature* and *Science*, only about 67% of studies were able to be replicated.[11]

Even if corrective measures in recent years — such as preregistration of hypotheses and study plans, more enforcement of requirements to share the data underlying experimental results, and other norms of open scholarship — may have reduced the problem, research has found that some high-profile peer-reviewed articles still contain apparently manipulated data.[12]

In Daubert parlance, we now see that there is a high known error rate for peer-reviewed studies, which should cause us to reconsider whether peer review is a trustworthy proxy for reliability.

The development of artificial intelligence could further undermine peer review if authors use it inappropriately in their published studies — for instance, by using a large language model to mass produce low quality work, or by failing to check for hallucinations in AI-generated text.[13]

Nowadays, many scientists do not automatically assume that peer-reviewed work is a bellwether of reliability. Scientists seriously consider many sources that are not peer-reviewed, such as conference papers and preprints of articles submitted to, but not yet accepted for publication by, academic journals.[14]

And many of the most celebrated scientific findings in history were not peer-reviewed. Indeed, only one of Albert Einstein's papers was peer-reviewed. The peer-review process so infuriated him that he moved the article to another journal.[15]

It is time for peer review to be seriously reconsidered by courts and the advisory committee, including whether it is an outmoded shibboleth for identifying reliable science in a courtroom.

### **Courts' Inconsistent Treatment of Peer Review After Daubert**

The Supreme Court's decision in Daubert explained that whether a particular technique or methodology had been published in a peer-reviewed journal is a "relevant, but not dispositive consideration."<sup>[16]</sup> The court noted that some propositions were too new or of limited interest, such that publication was unlikely.

But on remand to the U.S. Court of Appeals for the Ninth Circuit, in a passage that would become influential, the court ruled that "peer review [was] a significant indication" that research "meets at least the minimal criteria of good science."<sup>[17]</sup>

Courts proceeded inconsistently in how they applied Rule 702 after Daubert. Some courts followed the Ninth Circuit's view in Daubert that peer review was a strong signal of reliability, but other courts gave it little weight.<sup>[18]</sup>

The inconsistent treatment persisted, even when some courts noted the shifting views in the scientific community.<sup>[19]</sup>

Because courts rarely explained why they decided peer review deserved either substantial importance or minimal importance for a particular expert's work, many decisions appeared to reach a tentative conclusion about admissibility first, and then addressed peer review as an afterthought.

The disparate weight courts give peer review is illustrated by a recent pair of cases involving identical expert testimony in two different courts, each purporting to apply Daubert.

The cases were part of multidistrict litigation — In re: Zantac (Ranitidine) Product Liability Litigation, and related litigation in Delaware state court, In re: Zantac (Ranitidine) Litigation — involving claims that the antacid drug Zantac and its generic counterparts caused cancer.

The first court to rule, the U.S. District Court for the Southern District of Florida, excluded one expert's testimony on causation, finding in 2022 that the expert had "designed and conducted novel experiments for this litigation that did not follow any preexisting, peer-reviewed experimental designs, much less designs established to assess drug stability."<sup>[20]</sup>

But the lack of peer review was just one of the factors considered by the court, and the decision did not state how strongly it weighed that factor.

Yet in a separate lawsuit involving the same expert's proposed testimony, the Delaware Superior Court concluded in late May that the "lack of peer review ... is fodder for cross-examination, not exclusion."<sup>[21]</sup>

The different results in these nearly identical cases support the view that the overall gestalt has mattered more than careful analysis of the significance of peer review.

Further, no matter what courts have said about peer review, data suggests that peer review has not been a strong predictor of whether courts will admit expert testimony.

In the years after the Supreme Court's Daubert decision, many states adopted Daubert — including its recognition of peer review as a relevant indicator of reliability — but many did not.[22]

A large 2022 study of appellate decisions in criminal cases, published in the journal *Psychology, Public Policy and Law*, found no statistically significant difference in exclusion rates between federal court decisions before and after Daubert, nor between courts in states that chose to follow Daubert and those that did not.[23]

### **Role of Peer Review After December 2023 Rule 702 Amendment Still Uncertain**

The December 2023 amendment to Rule 702 clarified that the "preponderance of the evidence" standard applies to establishing the reliability of expert testimony. But the amendment did not address peer review.

Today, courts approach peer review the same as they did before the amendment: Some courts treat a lack of peer review as a serious shortcoming, but others do not.[24] And because courts rarely explain why they gave some or substantial weight to peer review, these decisions cannot be explained simply by the fact that Daubert lists it as a possible, but not essential, factor for assessing reliability.

This unresolved issue is important, because Rule 702 requires courts to act as gatekeepers for expert testimony. But the rules for determining what expert testimony passes muster for admission at trial should be clear.

More broadly, when courts rely on illustrative lists of factors that can support passing a legal test, reasoned opinions should give guidance to future courts about how and why the factors support particular decisions.

The jurisprudence under Rule 702 over the last few decades suggests that more guidance would be useful, including updated analysis of whether the replication crisis and current views in the scientific community should affect whether peer review belongs as a factor that helps identify reliable expert testimony.

---

*Jeffrey Gross is a partner and Robert LaCroix is an associate at Reid Collins & Tsai LLP.*

*The opinions expressed are those of the author(s) and do not necessarily reflect the views of their employer, its clients, or Portfolio Media Inc., or any of its or their respective affiliates. This article is for general information purposes and is not intended to be and should not be taken as legal advice.*

[1] See, e.g., David E. Bernstein & Eric G. Lasker, *Defending Daubert: It's Time to Amend Federal Rule of Evidence 702*, 57 *Wm. & Mary L. Rev.* 1, 8-12 (2015).

[2] Fed. R. Evid. 702 advisory committee's note to 2023 amendment.

[3] 509 U.S. 579, 593-94 (1993).

[4] *Danhoff v. Fahim*, No. 163120, --- N.W.3d ----, 2024 WL 3333321, at \*13-14 (Mich. July 8, 2024).

[5] Alvaro de Menard, *What's Wrong with Social Science and How to Fix It: Reflections After Reading 2578 Papers*, *Fantastic Anachronism* (Sept. 11, 2020), <https://fantasticanachronism.com/2020/09/11/whats-wrong-with-social-science-and-how-to-fix-it/>.

[6] See, e.g., Melinda Baldwin, *In Referees We Trust?*, 70 *Physics Today* 44, 44 (2017).

[7] *Id.* at 49.

[8] John P.A. Ioannidis, *Why Most Published Research Findings Are False*, *PLoS Med* (2005).

[9] Sara Schroter et al., *What errors to peer reviewers detect, and does training improve their ability to detect them?*, 101 *J. Royal Soc. Med.* 507 (2008).

[10] Stephan Lewandowsky & Klaus Oberauer, *Low Replicability Can Support Robust and Efficient Science*, 11 *Nature Communications* 358, at 2 (2020) (citing Rolf A. Zwaan et al., *Making Replication Mainstream*, 41 *Behavioral & Brain Scis.* e120 (2018)).

[11] *Id.* (citing Open Science Collaboration, *Estimating the Reproducibility of Psychological Science*, 349 *Science* 1, 1-8 (2015)).

[12] See, e.g., Uri Simonsohn et al., [109] *Data Falsificada (Part 1): "Clusterfake"*, *DataColada* (June 17, 2023), <https://datacolada.org/109>.

[13] Some, including the editors of the journal *Science*, claim that the advent of AI raises the risk of shoddy papers passing peer review. See *LLMs Now Write Lots of Science. Good*, *The Economist* (June 27, 2024), <https://www.economist.com/leaders/2024/06/27/llms-now-write-lots-of-science-good>.

[14] Cassandra L. Ettinger et al., *A Guide to Preprinting for Early-Career Researchers*, 11 *Biol. Open* 7 (2022).

[15] Andre Spicer & Thomas Roulet, *Hate the peer review process? Einstein did too*, *The Conversation* (June 2, 2014), <https://theconversation.com/hate-the-peer-review-process-einstein-did-too-27405>.

[16] 509 U.S. at 593.

[17] *Daubert v. Merrell Dow Pharms., Inc.*, 43 F.3d 1311, 1318 (9th Cir. 1995).

[18] See, e.g., *Baugh v. Cuprum S.A. de C.V.*, 845 F.3d 838, 845 (7th Cir. 2017) (lack of peer review no bar to admissibility because expert applied techniques well-established in industry); *Primiano v. Cook*, 598 F.3d 558, 566 (9th Cir. 2010) (reversing exclusion of non-peer reviewed opinion based on expert's "background and experience, and his explanation of his opinion"); *Bitler v. A.O. Smith Corp.*, 400 F.3d 1227, 1235 (10th Cir. 2005) (affirming admissibility of opinion that was "not susceptible to testing or peer review"); *Danhoff*, 2024 WL 3333321, at \*13-14.

[19] *Valentine v. Pioneer Chlor Alkali Co.*, 921 F. Supp. 666, 674 (D. Nev. 1996) (noting "some serious debate in the scientific community itself over the significance of publication per se and the adequacy of pre-publication evaluation of scientific writings").

[20] In re Zantac (Ranitidine) Prod. Liab. Litig., 644 F. Supp. 3d 1075, 1140 (S.D. Fla. 2022).

[21] In re Zantac (Ranitidine) Litig., C.A. No. N22C-09-101 ZAN, 2024 WL 2812168, at \*34 (Del. Super. May 31, 2024).

[22] See Jennifer L. Groscup et al., The Effects of Daubert on the Admissibility of Expert Testimony in State and Federal Criminal Cases, 8 Psych., Pub. Pol. & L. 339, 342, 344 (2002).

[23] Id.

[24] Compare United States v. Ortiz, --- F. Supp. 3d ----, 2024 WL 2889873, at \*9 (S.D. Cal. June 10, 2024) (excluding expert opinion on DNA testing methodology that had not been peer reviewed) with Teresko v. The 3M Co., No. 22-cv-1532-JPS, 2024 WL 2864402, at \*19-20 (E.D. Wis. June 16, 2024) (excluding novel expert opinion with passing reference to lack of peer review) and Garcia v. Singh, Civ. No. 2:23-545-WJ-GJF, 2024 WL 2939126, at \*3 (D.N.M. June 11, 2024) (excluding expert based on lack of record about Daubert factors, including peer review).