

Putting Learned Treatises Under the Microscope: Assessing the Rule 803(18) Hearsay Exception and the Reproducibility Crisis in Social Science

William E. Braff *

I. Courting science: The seductiveness of scientific evidence.

Justice cannot be observed under a microscope and laws cannot be concocted in a test tube. Although science and law both employ fixed rules to help us navigate and make sense of the world,¹ the two disciplines are often at odds.² For instance, science embraces uncertainties and fluctuating interpretations of natural phenomena whereas law requires a steadfast code that can be applied to resolve human conflicts. The Supreme Court's analysis in *Daubert v. Merrell Dow Pharmaceuticals* illuminates this contrast between law and science, as “[s]cientific conclusions are subject to perpetual revision. Law, on the other hand, must resolve disputes finally and

* Law Clerk, U.S. Court of Appeals for the Sixth Circuit; J.D., The Ohio State University Moritz College of Law 2019; B.A. in English and Political Science, Kenyon College 2013. Special thanks to Judge Edmund A. Sargus, Jr. for encouraging me to pursue this topic and to my friend and colleague John Connor Dever for being a fount of knowledge on social science and the law.

1. Lee Loevinger, *Standards of Proof in Science and Law*, 32 JURIMETRICS J. 323, 328 (1992) (“[T]he basic principles of reasoning or logic are no different in the field of law than science.”).
2. Bert Black, *Science and the Law in the Wake of Daubert: A New Search for Scientific Knowledge*, 72 TEX. L. REV. 715, 716 (1994) (“Judges and lawyers usually react to science with all the enthusiasm of a child about to get a tetanus shot.”).

quickly.”³ Although science may guide jurists, jurors, and legal practitioners in untangling factual complexities, the scientific process cannot replace legal processes, such as ensuring admitted evidence will not mislead or prejudice jurors.

Considering the persuasive power of scientific evidence,⁴ courts, as well as lawyers concerned with fair play, must be cautious regarding the introduction of materials into evidence that bear the imprimatur of science. This is doubly true when expert witnesses appeal to science to persuade lay juries.⁵ As a result, the standard for admitting expert testimony flows from *Daubert*, a case centered on scientific expert witness testimony. In *Daubert*, the Supreme Court announced nonexclusive factors that establish the veracity of the science relied on by an expert.⁶ The *Daubert* decision subjected scientific expert testimony to the demands of our legal system by setting legal standards outlining when science may be legally admissible (i.e., when scientific evidence satisfies the accuracy and relevancy required by justice).⁷ Although this factors-based standard may be “a flexible one,” the Supreme Court made clear that “evidentiary relevance

3. *Daubert v. Merrell Dow Pharmaceuticals, Inc.*, 509 U.S. 579, 597 (1993).
4. See HANNAH ARENDT, *BETWEEN PAST AND FUTURE* 262 (1961) (“[M]odern science . . . has changed and reconstructed the world we live in so radically that it could be argued that the layman and the humanist, still trusting their common sense and communicating in everyday language, are out of touch with reality[.]”).
5. *Id.* at 263 (discussing the “division between the scientist and the layman.”).
6. *Daubert*, 509 U.S. at 591-95 (1993).
7. John S. Ehrett, *Antifragile Policymaking: A Strategy for Institutional Response to the Social Science Reproducibility Crisis*, 49 U. MEM. L. REV. 447, 462 (2018) (“[T]he existing *Daubert*-based legal regime governing the admission or exclusion of such expert opinion—in the *courtroom* setting—is overwhelmingly deferential to purportedly scientific findings, including those arising from the behavioral science disciplines.”).

and reliability” must be the twin foundations for admitting scientific evidence at the trial level.⁸

In *Kumho Tire Co., Ltd. v. Carmichael*, the Supreme Court extended the *Daubert* test for expert-witness admissibility to fields beyond the hard sciences.⁹ Finding classification of knowledge as either scientific or technical to be a distinction without a useful difference, the Court refused to recognize “clear legal lines” between scientific and non-scientific expert testimony.¹⁰ Under this standard, trial judges must evaluate disciplines and fields to determine whether “principles, methods, or their application are called sufficiently into question”¹¹ Accordingly, judges “ensure the reliability and relevancy of expert testimony” by deciding whether an “expert, whether basing testimony upon professional studies or personal experience, employs in the courtroom the same level of intellectual rigor that characterizes the practice of an expert in the relevant field.”¹² In short, *Kumho Tire* expanded the gatekeeping function of trial judges to determine the reliability of non-scientific expert testimony.¹³ Unfortunately,

8. *Daubert*, 509 U.S. at 595.
9. *Kumho Tire Co., Ltd. v. Carmichael*, 526 U.S. 137, 141 (1999) (“We conclude that *Daubert*’s general holding—setting forth the trial judge’s general ‘gatekeeping’ obligation—applies not only to testimony based on ‘scientific’ knowledge, but also to testimony based on ‘technical’ and ‘other specialized’ knowledge.”).
10. *Id.* at 148.
11. *Id.* at 149.
12. *Id.* at 152.
13. Leslie Morsek, *Get On Board for the Ride of Your Life! The Ups, the Downs, the Twists, and the Turns of the Applicability of the “Gatekeeper” Function to Scientific and Non-Scientific Expert Evidence: Kumho’s Expansion of Daubert*, 34 AKRON L. REV. 689, 734-35 (2001) (“Following the *Kumho* decision, it is apparent that the trial judge’s “gatekeeper” role has expanded to include non-scientific expert testimony. With this expanded role, trial judges will be expected to entertain numerous facts and

this gatekeeping, when applied leniently, carries the cost of granting non-scientific expert testimony the same authority as scientific expert testimony under *Daubert*.¹⁴ As trial courts have been permissive when admitting experts and learned treatises under *Kumho Tire*, the current system erroneously stamps the authority of science onto social science research that may not meet the reliability and authoritativeness necessary for expert evidence. This could easily lead juries to afford inappropriate weight to social science research admitted into evidence.

By expanding the *Daubert* standard to experts in all fields, including social sciences, the Court in *Kumho Tire* embraced reproducibility and reliability as general benchmarks for expert admissibility. These changes are reflected in Federal Rule of Evidence 702, which was updated post-*Daubert* to incorporate the *Daubert* factors for evaluating expert witness testimony.¹⁵ This article argues that the reproducibility crisis in the behavioral and social sciences demands that currently overbroad admissibility of ex-

factors to determine its admissibility, a realm not previously entered by trial judges.”).

14. Julie A. Seaman, *A Tale of Two Dauberts*, 47 GA. L. REV. 889, 903 (2013) (“[S]uch a compromise gives the appearance of gatekeeping where in fact there is little to none and reinforces the impression that Daubert provides a check on unreliable evidence, suggesting that what is admitted satisfies the rigors of the scientific method.”).
15. Cynthia Lynne Pike, *The Impact of Revised MRE 702 and 703 in Response to Daubert*, 52 WAYNE L. REV. 285, 290 (2006) (“In amending FRE 702 to incorporate the *Daubert* factors, the advisory committee’s note identified additional factors including: (1) whether the expert has unjustifiably extrapolated from an accepted premise to an unfounded conclusion; (2) whether the expert has adequately accounted for obvious alternative explanations; (3) whether the expert is being as careful as he would ‘in his regular professional work outside his paid litigation consulting;’ (4) the reliability of the particular field of expertise; and (5) the existence and maintenance of standards and controls.”).

pert testimony be scrutinized, particularly when using Federal Rule of Evidence 803(18) to admit a learned treatise. First, this article explains the relationship between expert testimony and the learned treatise hearsay exception, as well as the purpose of the learned treatise hearsay exception. Next, this article explores the reproducibility crisis in the social sciences and the corresponding implications for admitting evidence. Part Four of this article then addresses the inadequacy of current admissibility standards for learned treatises relating to social science. Finally, this article concludes by proposing changes to Federal Rule of Evidence 803(18) to prevent abuse of the learned treatise exception, thereby limiting unreliable or unreproducible evidence from being introduced under the guise of scientific accuracy. For attorneys and judges seeking fair procedures of exposing evidence to jurors, reforming the learned treatise exception is a simple step that could limit the unfair prejudice of evidence being presented as scientifically accurate despite lacking the hallmark reliability of science.

II. What can be learned from a treatise: An overview of Rule 803(18).

Under Federal Rule of Evidence 803(18), statements contained in “a treatise, periodical or pamphlet” may be presented at trial if (1) “the statement is called to the attention of an expert witness on cross-examination or relied on by the expert on direct examination” and (2) “the publication is established as a reliable authority by the expert’s admission or testimony, by another expert’s testimony, or by judicial notice.”¹⁶ Statements published in learned treatises are read into evidence after these elements are met, but the entire treatise is not entered into evidence as an exhibit.¹⁷

16. Fed. R. Evid. 803(18).

17. *Id.*

Importantly, reading treatises into the record would otherwise be prohibited as inadmissible hearsay, so Rule 803(18) plays a key role in admitting otherwise inadmissible information.¹⁸

While prohibitions on admitting an entire treatise into evidence prevent jurors from taking loaded texts into the jury room, permitting experts to establish the authority of a learned treatise in their field essentially ties the admissibility of a learned treatise to the *Daubert* factors of verifying a witness as an expert. Once verified under *Daubert*, an expert witness can bless any text in their given field as admissible under the learned treatise exception.¹⁹ This article explores the ramifications of reading statements from a learned treatise into evidence and leaves cross examination and impeachment for another day. Because learned treatise exceptions vary among the state courts, this Article also limits its scope solely to Federal Rule of Evidence 803(18).

Until the adoption of Federal Rule of Evidence 803(18), learned treatises played a narrow role in informing the jury. Far from tying admission of learned treatises to expert verification, the common law limited the

18. Robert F. Magill Jr., *Issues Under Federal Rule of Evidence 803(18): The 'Learned Treatise' Exception to the Hearsay Rule*, 9 ST. JOHN'S J. LEGAL COMMENT. 49, 59-60 (1993) (describing how learned treatises permit the introduction of evidence "so reliable that the rule against hearsay will not be applied to exclude it.").
19. This is especially problematic considering the difficulty of aligning an expert's field with the subject covered by the learned treatise. See Richard M. Cagen, *Dealing with the Problem of Unreliable Evidence Admitted Under a Literal Interpretation of Federal Rule of Evidence 803-18*, 14 VAL. U. L. REV. 329, 345 (1980) ("Many treatises of very narrow scope may deal with a topic that is within the realm of several fields of knowledge. For example, a text concerning the proper metal used in a hip prosthesis involves the fields of orthopedic surgery, metallurgy, and mechanical engineering, among others.").

learned treatise exception to “exact sciences” like chemistry and mathematics (a category so restrictive as to exclude medical textbooks.)²⁰ Because authors of learned treatises are typically unavailable at trial to explain the methods underlying their research or scholarship, learned treatises generally constitute hearsay under the common law when submitted as substantive evidence for the truth of the matter asserted.²¹ Rule 803(18), however, permits a learned treatise to be read into evidence when an expert witness explains the treatise to the jury and remains open for cross-examination.²² Emphasizing the inherent trustworthiness and accuracy of research and documents published for professional and academic use, the Advisory Committee Note on Rule 803(18) downplays the potential prejudice of reading learned treatises into evidence in a jury trial.²³ Because the trial transcript may be read or transmitted to the jury at the discretion of the trial court,²⁴ learned treatises under Rule 803(18) are accessible to jurors despite the technical provision against admitting them as exhibits.

While it may seem sensible to inform jurors through authoritative sources, the ability of experts to independently authorize a learned treatise raises concerns. In *Constantino v. Herzog*, the Second Circuit found that the “rationale for [the learned treatise] exception is self-evident: so long as the

20. Edward J. Imwinkelried, *Evidence Law Visits Jurassic Park: The Far-Reaching Implication of the Daubert Court's Recognition of the Uncertainty of the Scientific Enterprise*, 81 IOWA L. REV. 55, 75 (1995).

21. Ralph-Ruebner, *Why Illinois Should Adopt Federal Rule of Evidence 803(18) to Allow the Learned Treatise Exception to the Hearsay Rule*, 39 S. ILL. U. L.J. 275, 283 (2015).

22. FED R. EVID. 803 advisory committee's note to exception (18).

23. *Id.*

24. *United States v. Kuta*, 518 F.2d 947, 954 (7th Cir. 1975) (“[I]t is recognized that it is within the discretion of the trial court whether to read portions of the trial transcript back to the jury at its request.”).

authority of a treatise has been sufficiently established, the factfinder should have the benefit of expert learning on a subject, even though it is hearsay.”²⁵ Despite the sound logic of this rationale, this account of the learned treatise exception breezes over the problematic details of how an expert establishes the authority of a learned treatise. While trial judges may act as “gatekeepers” when determining whether experts view a learned treatise as trustworthy in a given field, these decisions are hardly “draconian.”²⁶

Despite their gatekeeping function, trial courts possess expansive discretion when evaluating the reliability of a learned treatise. Under *Constantino*, the Second Circuit employs a “liberal standard” for the authoritative-ness inquiry of learned treatises subject to peer review.²⁷ Most concerning, the Second Circuit described the evaluation of a learned treatise’s reliability as a “freewheeling” undertaking that can be “conducted by any means.”²⁸ Far from checking an expert’s verification of a learned treatise’s authority, trial judges tend to permit learned treatises when an expert testifies that the treatise is accepted in a given field. Under this permissive standard, the Second Circuit upheld the admission of video tape discussing techniques for treating shoulder dystocia as a learned treatise while “recognizing some periodicals—provided there is a basis for doing so—as sufficiently esteemed to justify a presumption in favor of admitting the articles accepted for publication therein.”²⁹ In other words, the reputation

25. *Constantino v. Herzog*, 203 F.3d 164, 170-71 (2d Cir. 2000).

26. *Id.* at 171.

27. *Id.* at 172.

28. *Id.*

29. *Id.*

of a publication within a given field can bestow potential evidentiary authority to articles it publishes.³⁰

Courts tend to view a learned treatise's acceptance in the expert's field as sufficient grounds for establishing reliability and accuracy appropriate for a hearsay exception. For instance, a court may assume published authors "are acutely aware that their material will be read and evaluated by others in their field, and accordingly feel a strong pressure to be accurate."³¹ Yet even when authors of learned treatises demonstrate ulterior motives aside from providing an informed view of a field, such as preparation for litigation, such treatises can be found admissible. In *Welding Fume Products*, the trial court admitted into evidence articles published by an expert medical witness "with a view towards litigation" because the articles would inform the jury on the science at hand.³² This leniency of judges in admitting learned treatises approved by experts admitted under *Daubert* raises concerns, especially after *Kumho Tire* expanded the *Daubert* test to experts outside the hard sciences.³³

Although *Kumho Tire's* expansion of the *Daubert* factors demonstrated an attempt by the Supreme Court to ensure all expert testimony satisfied

30. Indeed, Federal Rule of Evidence 803(18) and the advisory committee notes lack an "explicit requirement that the writings eligible for admission be deemed trustworthy." Cagen, *supra* note 19, at 337.

31. *In re Welding Fume Products Liability Litigation*, 534 F. Supp. 2d 761, 765 (N.D. Ohio, 2008).

32. *Id.* at 765-66.

33. Indeed, the legal community viewed *Kumho Tire* as a response and remedy to the "widespread use of expert witnesses in civil cases" testifying about "junk science" with the intent to cause "manipulation of jurors." Joan Biskupic, *Trial Judges Told to Screen Experts*, THE WASHINGTON POST (March 24, 1999), <https://www.washingtonpost.com/wp-srv/national/longterm/supcourt/stories/court032499.htm>.

the reliability and general acceptability standards set forth in *Daubert*, the expansion of trial court discretion ironically detracted from that mission. Tellingly, Justice Scalia felt compelled to briefly concur in *Kumho Tire*, stressing that the majority opinion only granted judges “discretion to choose among *reasonable* means of excluding expertise that is *fausse* and science that is junky.”³⁴ By concurring, Scalia voiced his concern that trial courts under *Kumho Tire* may fail to hold inexact science to the necessary standards for qualifying experts. So failure to reasonably “apply one or more of the *Daubert* factors” to experts outside the hard sciences “may be unreasonable, and hence an abuse of discretion.”³⁵ Considering the lenient standards for admitting learned treatises and experts demonstrated in *Welding Fume Products* and *Constantino*, Justice Scalia’s concurrence appears particularly apt. District courts should take to heart Justice Scalia’s trepidation about under-scrutinizing “junk” science, especially considering recent turbulence in social and behavioral science.³⁶ Because the test for admitting a learned treatise under Rule 803(18) looks to the qualifications of the expert witness, not the trustworthiness of the treatise,³⁷ the current system risks exposing jurors to treatises that are not properly vetted for the trustworthiness necessary for a hearsay exception.

34. *Kumho Tire Co., Ltd. v. Carmichael*, 526 U.S. 137, 159 (1999) (Scalia, J., concurring).

35. *Id.*

36. *See infra* Part III.

37. Cagen, *supra* note 19, at 346 (“Rule 803-18 puts great importance on the qualification of the text by expert testimony. Given the conspicuous absence of language requiring a judicial determination of trustworthiness, drafters of the rule apparently intended to rely on the experts rather than judges to determine the trustworthiness of the submitted text. Unfortunately, qualification by experts as prescribed by Rule 803-18 is not an adequate substitute for a judicial determination of trustworthiness.”).

III. More social than science: How the reproducibility crisis raises reliability concerns for admitting social science publications under Rule 803(18).

The reasoning in *Daubert* ties the authoritative nature of expert testimony to whether the expert's methodology: (1) can be reliably tested, (2) is subject to peer review, (3) has an acceptable error rate, (4) operates under reliable standards, and (5) is accepted by the scientific or academic community.³⁸ These factors emphasize reliability, reproducibility of results, and strict publication standards. But the recent reproducibility crisis in psychology and other social sciences should lead courts to consider whether contemporary social science research and scholarship possesses the inherent authoritative nature to justify a hearsay exception under Federal Rule of Evidence 803(18).

Unlike mathematical proofs and theorems that generate identical results no matter how many times they are performed, science experiments and studies require frequent testing to confirm their truthfulness.³⁹ While chemical reactions may produce constant results, experiments outside the hard sciences often prove difficult to recreate. Studies or experiments that produce inconsistent, arbitrary, or divergent results upon multiple attempts fail to meet the standards for authoritative nature embodied in the

38. *Daubert v. Merrell Dow Pharmaceuticals, Inc.*, 509 U.S. 579, 591—95 (1993).

39. Brian A. Nosek et al., Open Science Collaboration, *Estimating the Reproducibility of Psychological Science*, (Aug. 2015), <https://science.sciencemag.org/content/sci/349/6251/aac4716.full.pdf> (“Reproducibility is a core principle of scientific progress. Scientific claims should not gain credence because of the status or authority of their originator but by the replicability of their supporting evidence.”).

Daubert rules.⁴⁰ Thus, the lack of reproducibility and reliability in the social sciences demands reconsideration of whether social science experts deserves the same clout as hard science experts in the courtroom, particularly when offering a learned treatise for admission into the trial record.

Recently, psychologists are admitting that many published studies cannot be reproduced with reliability.⁴¹ Indeed, merely “half of psychological studies can be successfully repeated” despite large, international efforts to demonstrate the reliability of the field.⁴² One such research team, Many Labs 2, sought to recreate “studies that had made a big splash and been highly cited,” but this effort still failed to replicate findings at greater than a fifty percent clip.⁴³ While skeptics of the reproducibility crisis argue that attempts to reproduce foundational psychological experiments fail due to small sample size, inexperienced researchers, and variability among test subjects, the Many Labs 2 project specifically accounted for these issues and still proved unable to reproduce key experiments.⁴⁴ Despite using large sample sizes, accounting for cultural differences, and collaborating with the researchers who published the initial experiment, Many Labs 2 failed to consistently validate widely-cited psychology studies.⁴⁵ A recent

40. See Seaman, *supra* note 14, at 918—20 (citing the failure of courts to prevent untrustworthy data, studies, and forensic evidence when performing the *Daubert* gatekeeping function).

41. Ed Yong, *Psychology’s Replication Crisis Is Running Out of Excuses*, THE ATLANTIC (Nov. 19, 2018), <https://www.theatlantic.com/science/archive/2018/11/psycholodys-replication-crisis-real/576223/>.

42. *Id.*

43. *Id.*

44. *Id.*

45. Richard A. Klein et al., *Many Labs 2: Investigating Variation in Replicability Across Samples and Settings*, 1 ADVANCES IN METHODS & PRAC. IN PSYCHOL. SCI. 443, 483—84 (2018).

study suggests researchers can only reproduce twenty-five percent of social psychology studies and fifty percent of cognitive psychology studies.⁴⁶ Unlike the Pythagorean theorem or the production of hydrochloric acid, many published psychology studies lack repeatability and reproducibility,⁴⁷ which are key indicia of authoritativeness under evidentiary jurisprudence.

This crisis of questionable research methods and publication criteria extends beyond psychology, covering many of the social, biological, and behavioral sciences.⁴⁸ Although the peer-review process employed in social science may create an aura of authoritativeness and reliability, a significant amount of scientific studies and research do not accurately reflect the world as accurately as a layperson might expect. A study by *Nature*, one of the most respected and recognizable interdisciplinary science journals,⁴⁹ found that “[m]ore than 70% of researchers have tried and failed to reproduce another scientist’s experiments” and that fifty-two percent of researchers agree there is a “significant crisis” of reproducibility.⁵⁰ Citing increased pressure to publish, selective reporting of results, growing bureaucracy, and competition for grants, the study presents a grim prospectus for the reliability of published studies.⁵¹

46. See Nosek, *supra* note 39.

47. *Id.* (“The present results suggest that there is room to improve reproducibility in psychology.”).

48. Ehrett, *supra* note 7, at 454 (“A growing body of evidence suggests that much contemporary research in the social and behavioral sciences cannot be replicated when the studies in question are conducted with new participants.”).

49. *About the Journal*, NATURE INTERNATIONAL JOURNAL OF SCIENCE, <https://www.nature.com/nature/about/>.

50. Monya Baker, *1,500 Scientists Lift the Lid on Reproducibility*, NATURE INTERNATIONAL JOURNAL OF SCIENCE (May 25, 2016), <https://www.nature.com/news/1-500-scientists-lift-the-lid-on-reproducibility-1.19970>.

51. *Id.*

The highly publicized Sokal affair of 1996 and the Sokal Squared hoax of 2018 suggest inadequate peer-review methodology among soft science publications, such as sociological or anthropological journals, undermines the authoritativeness of social science scholarship. In 1996, Alan Sokal, a New York University physics professor, penned a purposefully incoherent essay titled “Transgressing the Boundaries: Toward a Transformative Hermeneutics of Quantum Gravity,” which he submitted to peer-reviewed journals in an effort to expose the political agenda of academia.⁵² Sokal’s essay discussed “emancipatory mathematics,” methods to “demystify and democratize the production of scientific knowledge,” and the “crisis of late-capitalist production relations.”⁵³ By blending academic jargon into an essay on gravity, Sokal concocted a farcical paper that he believed appealed to postmodern and poststructuralist affinities among academics.⁵⁴ Sokal’s essay achieved publication in *Social Text*, an esteemed academic journal published by Duke University.⁵⁵ After Sokal revealed the truth about his article, publication standards of social science journals came under fire from both the academy and the general public.⁵⁶

Twenty years after Sokal’s high-profile undressing of publication standards among social science journals, the Sokal Squared hoax took up

52. George F. Will, *The Hilarious Hoax That Should Have Taught the Academy a Lesson*, THE WASHINGTON POST (Jan. 11, 2017), https://www.washingtonpost.com/opinions/the-hilarious-hoax-that-should-have-taught-the-academy-a-lesson/2017/01/11/ffbbe1fe-d758-11e6-b8b2-cb5164beba6b_story.html?utm_term=.87e3d542820b.

53. *Id.*

54. *Id.*

55. *Id.*

56. Brian Leiter, *Rethinking Legal Realism: Toward a Naturalized Jurisprudence*, 76 TEX. L. REV. 267 n.22 (1997).

the mantle of questioning peer-review standards outside the hard sciences.⁵⁷ Three scholars penned twenty hoax papers, in the style of Sokal, to expose what they called “grievance studies,” or social sciences that place politics and advocacy above the pursuit of truth.⁵⁸ Seven of these parodical studies found a home in “the best journals in the relevant fields,” including a paper titled “Our Struggle is My Struggle,” which the authors described as a “feminist rewrite of a chapter of Adolf Hitler’s *Mein Kampf*.”⁵⁹ Perhaps unsurprisingly, the response to Sokal Squared was mixed; proponents stated the experiment “showcase[d] a serious problem with big parts of academia” and critics dismissed Sokal Squared as a “bad faith” attempt to “pull a fast one” on scholarly journals.⁶⁰ In any event, the success of Allan Sokal in 1996 and Sokal Squared in 2018 in defrauding esteemed academic journals suggests the emperors of the social science community may wear no clothes.⁶¹

IV. Learned treatises, learned juries; arbitrary treatises, arbitrary juries.

Considering the increasingly apparent unreliability of social science and behavioral science publications, courts, and practitioners should reevaluate the propriety of these fields under the expert testimony schema created by *Daubert* and *Kumho Tire*. Revelations over the past twenty-five

57. Alexander C. Kafka, ‘Sokal Squared’: Is Huge Publishing Hoax ‘Hilarious and Delightful’ or an Ugly Example of Dishonesty and Bad Faith, THE CHRONICLE OF HIGHER OF EDUCATION (Oct. 3, 2018), <https://www.chronicle.com/article/Sokal-Squared-Is-Huge/244714>.

58. *Id.*

59. *Id.*

60. *Id.*

61. *Id.*

years show that illustrious social science journals can be defrauded and that well-cited psychological experiments may not be reproducible or repeatable.⁶² By granting experts, including social scientists and academics, the power to classify publications as learned treatises, the current learned treatise exception risks allowing unreliable studies or research to be read into evidence under Federal Rule of Evidence 803(18). This frustrates the logic of *Daubert*, which verifies experts and expert testimony based on scientific reliability and the inherent authority of widely accepted studies. In short, the reproducibility crisis should inspire courts to employ a heightened skepticism for admitting social science evidence under Rule 803(18).

Federal Rule of Evidence 803(18) exempts learned treatises from being excluded as hearsay because of a presumed accuracy and authoritativeness found in published treatises and expert accreditation of treatises.⁶³ Under this theory, “chaperonage” of a learned treatise by a qualified expert prevents admitting a treatise without independently establishing the treatise’s credibility.⁶⁴ This ensures that “mere publication,” even in a “highly regarded” journal, does not render a treatise an “automatically reliable authority.”⁶⁵ But recall the post-*Kumho Tire* “liberal standards” and “freewheeling” evaluation of experts and learned treatises under *Constantino* and *Welding Fume Products*.⁶⁶ Under the current system, standards for

62. *See supra* Part III.

63. Fed. R. Evid. 803 advisory committee notes.

64. *Campos v. Safety-Kleen Systems, Inc.*, 98 F. Supp. 3d 372, 384 (D. Puerto Rico 2015).

65. *Meschino v. North Am. Drager, Inc.*, 841 F.2d 429, 434 (1st Cir. 1988).

66. *See supra* Part II.

judicial scrutiny of learned treatises vary significantly across jurisdictions.⁶⁷ So the most lenient jurisdictions should be brought to heel, either by the Supreme Court or a revision of Rule 803(18).

Walker v. Yamaha Motor illuminates the need for an updated evidentiary rule governing admission of learned treatises. In *Walker*, the trial court admitted an expert's study, which "related to human cognition and perception."⁶⁸ Reviewing the reliability of the expert's study, the trial court cited the "applied flexibility" afforded by *Kumho Tire* in admitting the testimony of "non-scientist experts."⁶⁹ Giving the cursory explanation that the expert's study should be admitted as a learned treatise because the expert stated the study used widely accepted methodology in the experimental psychology community and that "he would submit his study for peer review," the court found the expert's study reliable.⁷⁰ Otherwise stated, the court allowed the expert's study to be read into evidence after

67. Magill, *supra* note 18, at 60 ("So, what standards will the courts want met before they hear reliability established? The wording in the cases varies widely. The Fifth Circuit in *United States v. Jones*, established a formula for reliability for works which 'have been subjected to widespread collegial scrutiny.' In *Burgess v. Premier Corp.*, the Ninth Circuit apparently established the standard that a book must be determined an authority 'accepted' within the industry. In *Allen v. Safeco Insurance Co.*, the Eleventh Circuit allowed articles established by an adverse witness as 'somewhat authoritative,' where the author was the director of a well-respected forensic science department, and the expert used the periodicals 'in keeping up to date.'").

68. *Walker v. Yamaha Motor Co., Ltd.*, No. 6:13-cv-1546-Orl-37GJK, 2016 WL 7325518 at *8, (M.D. Fla, Feb. 16, 2016).

69. *Id.*

70. *Id.*

examining the standards of experimental psychology and social science, and not the contents of the actual study.⁷¹

Considering the reproducibility crisis in the social and behavioral sciences, the reasoning displayed in *Walker* may permit unacceptably unreliable studies to be presented to jurors as published scientific findings. And the lax gatekeeping function of federal judges across multiple jurisdictions amplifies this problem because trial judges typically refrain from imposing any additional scrutiny aside from whether the expert was properly qualified under *Daubert* and if the learned treatise was properly read into evidence.⁷² To prevent juries from hearing potentially unreliably learned treatises read into the record, which bear the court's approval as inherently authoritative, the current learned treatise jurisprudence demands change.

71. See Maxine D. Goodman, *A Hedgehog on the Witness Stand—What's the Big Idea?: The Challenges of Using Daubert to Assess Social Science and Nonscientific Testimony*, 59 AM. U. L. REV. 635, 648 (2010) (“Ultimately, the reliability of an expert’s methodology is critical to the trustworthiness of her conclusions; yet, courts often fail to scrutinize it, particularly when the nature of the expertise is foreign to the court. The obvious irony is that the more elusive the nature of expertise (thus suggesting the need for an exacting analysis), the more lax the courts’ scrutiny of methodology.”).
72. See, e.g., *U.S. v. Norman*, 415 F.3d 466, 473 (5th Cir. 2005) (“The treatise was not established as a reliable authority because Norman did not offer an expert in this area”); *Whitford v. Gill*, 218 F.Supp. 3d 837, n.342 (W.D. Wisc. 2016) (plaintiffs merely “highlighted” twenty-three social science articles regarding gerrymandering and failed to read these treatises into evidence); *Smith v. U.S.*, No. 3:95cv445, 2012 WL 1453570 at *13, (S.D. Ohio Apr. 26, 2012) (plaintiff failed to properly read learned treatises into the record); *Barraford v. T&N Ltd.*, 988 F. Supp. 2d 81, 87 (D. Mass. 2013) (expert report failed to include the articles being submitted as learned treatises); *Glowczenski v. Taser Intern, Inc.*, 928 F. Supp. 2d 564, 573 (E.D. N.Y. 2013) (plaintiff failed to introduce expert necessary for submitting medical journals as learned treatises).

V. What can be learned: Corraling the learned treatise hearsay exception.

Justice Scalia’s concerns about the reliability and authoritativeness of “junky” science have been borne out by developments in social science in the years after *Kumho Tire*.⁷³ Just as Justice Scalia expressed a desire to ensure reasonable vetting of experts, who serve the function of verifying learned treatises under Rule 803(18), contemporary jurists and practitioners should fret over the admission of learned treatises bearing the imprimatur of science but lacking the authoritativeness and reliability necessary for a hearsay exception. Given the reproducibility crisis in the social and behavioral sciences, legal authorities should impose enhanced scrutiny for social science studies read into evidence as learned treatises under Rule 803(18).

Adding language to Rule 803(18) that explicitly calls for judicial scrutiny would prevent juries from being exposed to inaccurate or unreliable social science. For instance, potential language for a proposed Rule 803(18)(C) could require courts to hold learned treatises to the same *Daubert* factors as experts. Because *Daubert* operates as the Supreme Court’s preferred mechanism for ensuring the reliability and authoritativeness of documents presented by experts, an additional step of verifying learned treatises as reproducible meshes well with current expert evidence jurisprudence. Under this revised rule, experts qualified under *Kumho Tire* could not simply establish a learned treatise as reliable, but judges would play an active role in evaluating learned treatises as opposed to the passive

73. *Kumho Tire Co., Inc. v. Carmichael*, 526 U.S. 137, 159 (1999) (Scalia, J., concurring).

role outlined in *Constantino* and *Welding Fume Products*.⁷⁴ While this may expend additional judicial resources, mitigating the deep-seated issues in contemporary social science revealed by the reproducibility crisis would be well worth the cost.

But such language might still not cure a core underlying issue: judges and clerks are not always well-equipped to unpack complicated scientific evidence. Or they cannot separate the wheat from the chaff when dealing with potentially “junky” science. Judge Posner, although not directly in the learned treatise context, has suggested that courts should rely more heavily on court-appointed expert witnesses.⁷⁵ And federal judges have that power under Federal Rule of Evidence 706.⁷⁶ But, as Judge Posner points out, the adversarial American legal system is loath to permit judges to act like pseudo-parties in calling witnesses and undertaking research.⁷⁷ So urging courts to appoint neutral experts is not a likely quick fix.

74. While concerns may exist about generalist judges evaluating complex fields, improvements in technology and availability of data render this concern unavailing. See Pamela S. Katz, *Expert Robot: Using Artificial Intelligence to Assist Judges in Admitting Scientific Expert Testimony*, 24 ALB. L. J. SCI. & TECH. 1, 12—44 (2014).
75. Richard A. Posner, *What is Obviously Wrong With the Federal Judiciary, Yet Eminently Curable Part I*, 19 Green Bag 2d 187, 190 (2016). (“A big problem with jury trials is that often they involve technological or commercial issues that few jurors understand (not that many judges understand them either) and that the lawyers and witnesses are unable or unwilling to dumb down to a level that the jurors would understand. There is a solution to this problem, however, though one that few judges employ: appointment by the judge of an expert witness (thus a “neutral” expert, by virtue of not having been selected by the lawyer for one party to the litigation).”).
76. Fed. R. Evid. 706(a) (“The court may appoint any expert that the parties agree on *and any of its own choosing*.”) (emphasis added).
77. Posner, *supra* note 75, at 190 (“The authority to make such an appointment is explicitly conferred on federal judges by Rule 706 of the Federal Rules of Evidence, but is alien to the Anglo-American judicial culture, in which the witnesses

In addition to cultural inhibitions, the cost of employing independent experts with advance scientific knowledge can be significant. And we know that from experience. In the silicone breast implant litigation of the 1990s, Judge Posner appointed experts from the National Science Panel to both review the arguments of the parties and conduct independent research.⁷⁸ After two years and \$800,000, the Panel produced its report.⁷⁹ But considering the length and expense of that process, using appointed experts like the National Science Panel did not catch on in other courts.⁸⁰ Aside from the monetary and delay costs, appointing experts under Rule 706 rests on the questionable assumption that neutral experts are readily identifiable and would not usurp the judge as final decisionmaker.⁸¹ As a result, suggesting an increased reliance on Rule 706 appointed experts to reform the learned treatise exception is unlikely to persuade courts and rule makers.

In pursuit of a less costly alternative, Rule 803(18) could be also amended to include a notice requirement forcing each party to give ad-

in a case are designated by the lawyers rather than by the judge. The fault is the culture. Our legal culture, in contrast to that of most countries in the world (notably Japan and the nations of Continental Europe), is ‘adversary,’ in the sense that the judge is the arbiter of a contest—a drama, really—put on by the lawyers for the contending parties.”).

78. Izabelle Tully, *The Courtroom Turned Classroom: A Model Procedure for Educating the Gatekeepers of Expert Evidence in Complex Tort Cases*, 40 CARDOZO L. REV. 2405, 2414-16 (2019).

79. *Id.*

80. *Id.* at 2433-34.

81. See Samuel R. Gross, *Expert Evidence*, 1991 WIS. L. REV. 1113, 1193–95 (1991) (discussing criticism of appointed experts having “too much power” and “project[ing] a false aura of infallibility,” especially in “psychiatry and clinical psychology.”).

vanced notice of an intent to introduce a learned treatise. Such an amendment would permit parties to evaluate adverse learned treatises and independently determine the trustworthiness of the source.⁸² While a notice requirement would shift the burden and cost of evaluating treatises onto the parties, it would likely deter the introduction of research containing methodological flaws or other features that produce untrustworthy data.

Additionally, the Supreme Court could revisit Scalia's concurrence in *Kumho Tire* and recalibrate the gatekeeping function of district judges, should the proper case arise. Such an opinion could give direct and explicit guidance to lower courts about the proper ways to verify the reliability and reproducibility of a learned treatise or an expert's field. Even so, the well-documented failings of social sciences to reproduce test results and properly vet scholarly articles demand immediate action, which could be most expeditiously achieved by revising Rule 803(18). One juror being improperly swayed by unreliable research or data being read into evidence under the Rule 803(18) learned treatise exception is one juror too many.

Conclusion

Social science, much like the law, touches almost every imaginable facet of life. By explaining behavioral patterns, modes of cognition, and other human tendencies, social science research can reveal counterintuitive or obscure truths about our world. Indeed, in almost any case, an expert social science witness may exclaim "But lest my liking might too sudden seem, / I would have salved it with a longer treatise."⁸³ And the pervasive reach of social science shows precisely why trial courts must be

82. Magill, *supra* note 18, at 53 ("[C]ommentators suggest that a pretrial conference be used to determine which learned treatise will be presented along with any expert to testify with such treatise. Advance disclosure allows the parties opportunity to test learned treatise material by a motion in advance of trial.").

83. WILLIAM SHAKESPEARE, *MUCH ADO ABOUT NOTHING*, ACT. 1, SC. 1.

vigilant in protecting juries from unreliable social science treatises that bear the lofty authority of general science. Thus, Federal Rule of Evidence 803(18) should be revised to impose strict gatekeeping procedures among trial courts, especially considering the impact of *Kumho Tire* on expert witness jurisprudence and the deficiencies of the social sciences put on public display by the reproducibility crisis.