Daubert, Schmaubert: Criminal Defendants and the Short End of the Science Stick

Susan D. Rozelle

Follow this and additional works at: https://digitalcommons.law.utulsa.edu/tlr

Part of the Law Commons

Recommended Citation

Available at: https://digitalcommons.law.utulsa.edu/tlr/vol43/iss2/12

This Article is brought to you for free and open access by TU Law Digital Commons. It has been accepted for inclusion in Tulsa Law Review by an authorized editor of TU Law Digital Commons. For more information, please contact megan-donald@utulsa.edu.
DAUBERT, SCHMAUBERT: CRIMINAL DEFENDANTS AND THE SHORT END OF THE SCIENCE STICK

Susan D. Rozelle*

I. INTRODUCTION

I have a hypothesis. It is falsifiable, and if someone wants to do the empirical work, we could establish error rates, subject our methodology to peer review—the whole nine yards. The results would be admissible in any court of law, satisfying even the strictest Daubert1 interpretation. But the courts will not like it. I suspect that, almost2 regardless of the quality of the science involved, judges tend to admit scientific evidence when it favors the prosecution while refusing to admit it when it favors the defense.3

First, a caveat: Obviously this is a broad generalization. Of course there are defense-friendly judges out there—some notoriously so.4 And no doubt close scrutiny of many excellent judges’ records would reveal an utterly boring (or utterly astounding, depending on one’s level of cynicism) neutrality. My hypothesis is simply that, in general, more judges favor the prosecution in admitting scientific evidence than favor the defense.

Enough data points in the direction of such a prosecution bias that the study seems worth undertaking. A few examples suffice for this essay’s purposes: The judiciary admits the lousy science of the classically prosecution-friendly fingerprint evidence, while suppressing the sound science of the classically defense-friendly evidence about eyewitness identifications. Some might claim the difference between the two is that the

---

* Associate Professor of Law, Capital University Law School, visiting at the University of Oregon School of Law, 2007–2008. Special thanks to Tamara Piety for inviting my contribution to this symposium issue.

1. Daubert v. Merrell Dow Pharm., Inc., 509 U.S. 579 (1993). Daubert sets out the standards by which expert evidence is judged competent (or not) to be admitted in court, and lays the burden of making this decision at the feet of trial judges. Id. at 597.

2. Of course some pseudo-science smells so bad that not even the most results-oriented judges could hold their noses tightly enough to let it slip past. See e.g. Gen. Elec. Co. v. Joiner, 522 U.S. 136, 153 n. 6 (1997) ("An example of ‘junk science’ that should be excluded under Daubert as too unreliable would be the testimony of a phrenologist who would purport to prove a defendant’s future dangerousness based on the contours of the defendant’s skull."). I would like to think there are analogously strong sciences that not even the most results-oriented judge could refuse to admit, though I confess to being less certain on that score.


first involves "hard" (if faulty) science while the other involves the sort of "soft" science that always draws skepticism. This proposed explanation is belied, however, by courts' ready admission of the both soft and faulty behavioral science evidence of future dangerousness—the most prosecution-friendly "science" of them all.

Some might argue that DNA evidence proves my thesis wrong out of the starting gate: After all, DNA evidence is widely admitted, good science, and defense-friendly, right? Not exactly. While it is true that DNA can exonerate (and there have been some spectacular examples of that recently5), most criminal defendants' lack of resources means that DNA is much more likely to be a tool of the prosecution than of the defense.6 Similarly, "prosecutors may get clobbered for awhile on suggestive line-ups, but once they get the method down, it will bolster the government's case."7 The game of scientific evidence looks fixed.

II. ANECDATA8

It goes without saying that the sciences and the law differ in many ways. One of the most fundamental is in the importance placed on the single instance. Scientists are skeptical of single instances; they do not prove much of anything. Instead, scientists look for trends, and in order to do so, they need statistically significant sample sizes. Lawyers, on the other hand, value single instances quite highly. Like a treating physician with a patient, a lawyer with a client is more interested in the singular than in the aggregate. Even more than that, our system of precedent means that practicing lawyers must comb through multitudes of records in search of the one, single instance that most resembles their client's situation, for that is the case that will govern theirs.9

That said, what follows are a few examples of the hypothesized phenomenon. No representation is made to having used any sort of scientific method in selecting these examples; the suggestion is simply that these instances point clearly enough in the direction of bias that it is worth doing the study. Grant proposal, anyone?

7. Telephone Interview with John Mitchell, Professor of Law at Seattle University School of Law (June 5, 2007).
9. See Andre A. Moenssens et al., Scientific Evidence in Criminal & Civil Cases 1259 (5th ed., Foundation Press 2007) (describing scientists as skeptical of small samples, while lawyers seek the one case that looks most like theirs); see also Herbert M. Kritzer, The Arts of Persuasion in Science and Law: Conflicting Norms in the Courtroom, 70 L. & Contemp. Probs. ___ (forthcoming 2007) (available at http://papers.ssrn.com /sol3/papers.cfm?abstract_id=1024520) (discussing different "data sources," the specific in law and the general in science, and acknowledging that scientists are often called upon to use their knowledge of the general to make a diagnosis about a specific individual).
A. **Fingerprint and Toolmark Evidence**

Fingerprint evidence has a long and respectable history in court. What it does not have is a long and respectable history in academic literature.\(^ {10} \) Professor Jennifer Mnookin pointedly gathers the judiciary’s almost uniform acceptance of the technique,\(^ {11} \) and contrasts that with the academic literature’s almost uniform condemnation of its lack of standardization or proof of validity.\(^ {12} \) The error rate is untested, and in fact untestable, given the lack of standardization in the technique.\(^ {13} \)

What is perhaps the most famous erroneous match in history occurred when Oregon lawyer Brandon Mayfield’s fingerprints were declared to match those found at the scene of the Madrid bombing.\(^ {14} \) A lot of things went wrong in that case,\(^ {15} \) and its very mishandling led researchers, in addition, to uncover unconscionable observer effects tainting fingerprint experts’ opinions. In 2006, Itiel R. Dror published a study involving five fingerprint experts who had previously testified under oath that a given pair of prints was a certain match.\(^ {16} \) These experts were provided the very same prints they had testified about before, but this time were told the prints were the two from that infamous Brandon Mayfield case.\(^ {17} \) Three examiners now unequivocally stated that those prints came from different sources, directly contradicting their previous testimony; one stated he could not tell, whereas before he had asserted a positive match; and only one of the five consistently stated that the prints matched.\(^ {18} \) The study has its limits, notably its small sample size; however, additional research so far has only bolstered the conclusion of contamination.\(^ {19} \)

Similar controversy surrounds all of the so-called “match” evidence, such as “hair, bite mark, shoe and footprints, and handwriting comparison . . . that rely primarily upon experience and visual comparison to reach conclusions of a ‘match.’”\(^ {20} \) Tool mark evidence, wherein an expert testifies that the particular marks left by a harder object’s
impression on a softer one identify the exact harder object that left those marks (commonly firearms and crowbars), also falls into this category. Judge Nancy Gertner’s opinion in United States v. Green permitted the tool mark witness to point out areas of similarity, but refused to allow the witness to testify to any conclusion. While clearly stating that she did not believe the evidence met Daubert, she explained her decision partially to admit it nonetheless as a purely pragmatic one:

I reluctantly come to the above conclusion because of my confidence that any other decision will be rejected by appellate courts, in light of precedents across the country, regardless of the findings I have made. . . . But the standards should be higher than were met in this case, and than have been imposed across the country. The more courts admit this type of toolmark evidence without requiring documentation, proficiency testing, or evidence of reliability, the more sloppy practices will endure; we should require more.

Perhaps it is a coincidence, or perhaps my hypothesis is simply wrong, but I suspect that when the evidence favors the defense, we do require more. Much more.

B. Eyewitness Identification Evidence

Eyewitness identification surely must be the oldest form of match evidence known to civilization. As with fingerprint and tool mark evidence, eyewitnesses “rely primarily upon experience and visual comparison to reach conclusions of a ‘match.’” The limitations of our ability accurately to identify another person with whom we had not previously been acquainted have been even better documented than the limits of our ability to make other kinds of matches. Indeed, if there is scientific consensus about anything, there is scientific consensus about how bad we are at identifying each other.

Despite the clarity of the science on this point, the “overwhelming majority of courts” have refused to allow experts on eyewitness misidentification. The court in United States v. Carter provided a prototypical—and prototypically dismissive—rationale for exclusion: “[J]urors understand that memory can be less than perfect.” To the extent that jurors might need to be reminded, the pattern jury instructions are sufficient to the task. And besides, the court reasoned, allowing such experts to testify would invade the province of the jury by commenting on another witness’s credibility.

21. Moenssens et al., supra n. 9, at 463.
22. Id. at 494.
24. Id. at 109.
25. Id.
28. Henry F. Fradella, Why Judges Should Admit Expert Testimony on the Unreliability of Eyewitness Testimony, 2 Fed. Cs. L. Rev. 1, 23 (2007) [hereinafter Fradella, Why Judges Should]; see also Moenssens et al., supra n. 9, at 1382 (citing cases); Henry F. Fradella et al., supra n. 10, at 414–15, 426–27 (stating that nine of eleven post-Daubert federal cases refused to allow pro-defense eyewitness identification evidence).
29. 410 F.3d 942 (7th Cir. 2005).
30. Id. at 950.
31. Id.
While the imperfection of memory is familiar to us all (where did I leave my keys again?), the things these experts would explain are quite foreign to most people, and a jury instruction simply is not up to the task of educating them. Take, for example, the favored instruction in *Carter*:

You have heard testimony of an identification of a person. Identification testimony is an expression of belief or impression by the witness. You should consider whether, or to what extent, the witness had the ability and the opportunity to observe the person at the time of the offense and to make a reliable identification later. You should also consider the circumstances under which the witness later made the identification. The government has the burden of proving beyond a reasonable doubt that the defendant was the person who committed the crime charged.

Nothing in this instruction, well-meaning as it is, provides concrete information about the sorts of limits to eyewitness's ability to identify people that a lay jury needs to learn.

Studies prove our startling lack of ability to identify someone of a different racial group; the influence of "weapons focus" (or how the presence of a weapon causes observers to focus on the weapon rather than on the assailant's face); and how stress impairs memory, rather than heightening it as many laypeople assume. We fail to realize how memories suffer from contamination by later events, such as when police officers inadvertently create a "memory" of a weapon by asking eyewitnesses to think about whether they saw one. And although the number one criterion jurors use to decide whether to believe a witness is the witness's level of certainty, the research demonstrates no correlation whatsoever between eyewitness certainty and accuracy. This distinctly counter-intuitive point has implications for every syllable uttered by the eyewitness. It is foolish to refuse its admission.

Of course, the far-reaching implications of juror education on the issue are precisely why, at bottom, courts have refused to admit this expert testimony: It invades the province of the jury by impeaching the other witnesses' credibility. Although this is one of the most commonly given reasons for refusing to admit eyewitness identification experts' testimony, it is nonsense. Every piece of relevant evidence necessarily either

33. Schmechel et al., *supra* n. 27, at 204 (empirical study demonstrating lay confusion); *see also* Edward Stein, *The Admissibility of Expert Testimony about Cognitive Science Research on Eyewitness Identification*, 2 L. Probability & Risk 295, 300 (2003) (calling such information "quite counterintuitive and hardly commonsensical").

34. *Carter*, 410 F.3d at 951 n. 2 (quoting 7th Cir. Pattern Jury Instr. § 3.08 (1999)).


36. Schmechel et al., *supra* n. 27, at 200 (citing studies demonstrating lay confusion).

37. *Id.* at 196-97; *see also* Moenssens et al., *supra* n. 9, at 1375.

38. Moenssens et al., *supra* n. 9, at 1375.

39. Schmechel et al., *supra* n. 27, at 195-96.


41. Moenssens et al., *supra* n. 9, at 1378 n. 24 (citing B.L. Cutler, Steven D. Penrod & Thomas E. Stuve, *Juror Decision Making in Eyewitness Identification Cases*, 12 L. & Hum. Behav. 41 (1988)).

42. *Id.* at 1377.

43. Fradella, *Why Judges Should*, *supra* n. 28, at 23 ("The reason most frequently cited by courts for
bolsters or impeaches every other piece of evidence in the case. More to the point, the evidence offered by experts in this field does not merely vouch for (or impugn) others' truthfulness. If it did, of course, exclusion would be proper. Testimony amounting to nothing more than whether one witness believes another witness is telling the truth truly would be supremely unhelpful. After all, deciding whether to believe the witnesses is the jury's function; jurors are not aided in making up their own minds about whom to believe by looking to see what others think.

But the excluded eyewitness identification experts would not be simply "choosing up sides."44 Indeed, they would not be choosing up sides at all. By offering jurors information about what sorts of situations aid accuracy in identifications and which hinder it, and by teaching them, for example, that their natural tendency to consider witnesses' certainty is a red herring, experts are providing jurors with the sorts of tools they need to make the credibility determination for themselves.45 If an eyewitness were shown to need glasses, establishing that would not usurp the function of the jury in evaluating the witness's credibility; it would aid it. The information provided by eyewitness identification experts is no different, and yet it is routinely excluded.

C. Future Dangerousness Evidence

It might be suggested that the hostility to eyewitness identification experts is a product of its nature: Society seems more skeptical of the behavioral sciences in general than it is of the hard sciences.46 Anything having to do with human nature is soft, squishy, subject to interpretation, and therefore suspect. Inquiries relating to the physical world, on the other hand, like chemistry, biology, and physics, for example, are indisputably, objectively real. Leaving aside discussions about the validity of that presumptive dichotomy, one counterexample may suffice to illustrate that the difference is unlikely to be one of a hard versus a soft science: Courts almost always admit evidence of future dangerousness.

Future dangerousness figures into an extraordinarily wide variety of situations, such as "bail, sentencing, the death penalty, parole, sexually dangerous persons, protective services to children and adults, involuntary commitments and deinstitutionalization and hospitalization of insanity acquittors."47 Balanced against our
natural desire for safety is our equally natural desire for freedom from wrongful detention. Of the two, of course, it is sometimes harder to bring the desire for freedom from wrongful detention into focus. We readily identify ourselves as in danger from bad people; we are less inclined to identify with the wrongfully detained. 48

Perhaps because of this usually one-sided empathy, courts have allowed evidence of future dangerousness though it blinks even the most liberal interpretation of Daubert. In Barefoot v. Estelle 49 (a pre-Daubert case), the U.S. Supreme Court frankly acknowledged the official position of the American Psychiatric Association (APA), which urges that psychiatrists not be permitted to offer such predictions. 50 The APA explains that “even under the best of conditions, psychiatric predictions of long-term future dangerousness are wrong in at least two out of every three cases.” 51 The Court’s opinion explicitly admitted this, saying that “the ‘best’ clinical research currently in existence indicates that psychiatrists and psychologists are accurate in no more than one out of three predictions of violent behavior.” 52 Nevertheless, the Court explained, such testimony should remain admissible, since “[n]either petitioner nor the Association suggests that psychiatrists are always wrong with respect to future dangerousness, only most of the time.” 53

Even pre-Daubert, the Court’s willingness to put such a sentence in writing is inconceivable. 54 “Under no possible standard for the admissibility of expert opinion testimony can the admission of psychiatric testimony on future dangerousness pass the test.” 55 Under Frye’s general acceptance test, 56 evidence of future dangerousness clearly fails. As the APA wrote in its brief, “[t]he unreliability of psychiatric predictions of long-term future dangerousness is by now an established fact within the profession.” 57 Far from being generally accepted science, predictions of future dangerousness are generally shunned. Post-Daubert, the evidence should be found equally inadmissible. Daubert purports to require “reliability.” 58 Even under the loosest definition of the term, evidence more often wrong than it is right simply cannot qualify.

The dissenting Justice Blackmun put it eloquently in Barefoot: “[W]hen the Court

48. The occasional reminder of our own vulnerability is helpful: “[N]ever send to know for whom the bell tolls; it tolls for thee.” John Donne, Devotions upon Emergent Occasions 87 (Oxford U. Press 1987) (originally published 1624).
51. Id.
53. Id. at 901.
55. Moenssens et al., supra n. 9, at 1363.
56. Frye v. U.S., 293 F. 1013, 1014 (D.C. App. 1923) (“[T]he thing from which the deduction is made must be sufficiently established to have gained general acceptance in the particular field in which it belongs.”).
58. Daubert, 509 U.S. at 589 (“[T]he trial judge must ensure that any and all scientific testimony or evidence admitted is not only relevant, but reliable.”).
knows full well that psychiatrists' predictions of dangerousness are specious, there can be no excuse for imposing on the defendant, on pain of his life, the heavy burden of convincing a jury of laymen of the fraud."\(^{59}\) Despite the clear failure of future dangerousness testimony to pass any established test for admissibility, such evidence continues routinely to be admitted.\(^{60}\)

### III. Analysis

Truthfully, continued admission of such evidence is unsurprising. Commentators have observed since the beginning of the Daubert era that, despite the appearance of a sea change in the requirements for admission of expert testimony, courtroom evidence looks very much like it did pre-Daubert.\(^{61}\) And it has been suggested before that the articulated test for scientific evidence has less to do with admission decisions than other, unarticulated factors. As Professor Mark Brodin put it, "the frequent exclusion of expert testimony on eyewitness identification despite its scientific reliability, contrasting sharply with the widespread admission of evidence of such dubious reliability as 'future dangerousness' . . . strongly suggests that factors other than reliability are playing the determinative role."\(^{62}\)

Doubtless inertia has something to do with it. Our system of precedent literally teaches us to follow past practice, and—Daubert's instruction to scrutinize old and new science evidence alike\(^{63}\) notwithstanding—that must be a hard habit to break.\(^{64}\) But inertia does not explain how good, but pro-defendant, science came to be excluded in the first place, while lousy, but pro-prosecution, science came to be admitted.

As for the hostility to good, defense-friendly evidence, the explanation I propose is an easy (if unpalatable) one: Politics. Personal politics, certainly: Judges' personal views may be more in line with the government to begin with, since judges tend to come to the bench from prior experience as prosecutors, rather than as defense attorneys.\(^{65}\) More than that, though, judges are political animals. Although we tell our seventh-grade civics students that the judiciary is the independent branch of government—the one branch that is free to exercise its own judgment without navigating the shifting plates of political tectonics that move the executive and legislative branches—the truth is that whether judges are appointed or elected, becoming a judge "takes political support."\(^{66}\) Most judges are elected,\(^{67}\) and even in jurisdictions that appoint their judges, those who wish

\(^{59}\) Cf. 463 U.S. at 935–36 (Blackmun, Brennan & Marshall, JJ., dissenting).

\(^{60}\) Mark S. Brodin, Behavioral Science Evidence in the Age of Daubert: Reflections of a Skeptic, 73 U. Cin. L. Rev. 867, 889 (2005).

\(^{61}\) Id. at 869–70.

\(^{62}\) Id. at 891 (emphasis omitted).

\(^{63}\) Daubert differed from the old Frye test in this respect, too, since Frye concerned only novel science. Daubert, 509 U.S. at 592 n. 11.

\(^{64}\) Brodin, supra n. 60, at 892 (describing the "grandfathering in" of evidence previously admitted in the jurisdiction).


\(^{67}\) Id. ("Some State judges are appointed, but the remainders are elected in partisan or nonpartisan State
to wear the robe someday have to come to the (favorable) attention of those members of the nominating commission “composed of members of the bar and the public” who do the appointing. And the public is leery of criminal defendants and their so-called constitutional rights. Studies consistently show a punitive public, convinced that “criminals unfairly benefit from legal procedures,” and therefore keen “to deny or limit procedural rights.” In one revealing study of respondents in the California bay area, surveying a population in which over 60% were college graduates and half were self-described “liberals” (with one-third describing themselves as “moderates” and the remainder—less than 20%—as “conservatives”), the percentage of people agreeing with the famous principle that “[i]t is better to let 10 guilty people go free than to convict one innocent person by mistake” was only 56%. Now imagine that data if the survey had been conducted in the South or Midwest. Public sentiment on these issues is well known and frequently exploited. No one ever lost an election by being too tough on crime.

As for the admission of lousy, pro-prosecution evidence, that seems to be structural—an unintended consequence of, of all things, our intentionally defense-friendly burden of proof. All persons are presumed innocent until proven guilty. This means that the burden of proof in criminal trials rests with the government. And of course, the business of evidence in general belongs to the party with the burden of proof. In the absence of evidence, the case goes to the defendant. Hence, the one with the burden of proof wants more evidence in, and the one defending opposes its admission. This means that most evidence of all kinds is being offered by the prosecution.

Now, combine that with the fact that post-Daubert, most judges are in over their heads. When the U.S. Supreme Court sent Daubert back on remand to determine the reliability of the evidence, Judge Kozinski put it bluntly:

As we read the Supreme Court’s teaching in Daubert, . . . though we [judges] are largely untrained in science and certainly no match for any of the witnesses whose testimony we are reviewing, it is our responsibility to determine whether those experts’ proposed testimony amounts to “scientific knowledge,” constitutes “good science,” and was “derived by the scientific method.”

Our responsibility, then, unless we badly misread the Supreme Court’s opinion, is to resolve disputes among respected, well-credentialed scientists about matters squarely within their expertise, in areas where there is no scientific consensus as to what is and what is not “good science,” and occasionally to reject such expert testimony because it was not “derived by the scientific method.” Mindful of our position in the hierarchy of the federal judiciary, we take a deep breath and proceed with this heady task.

Being out of their element creates a human factor that urges admission over
exclusion in general. When unsure of themselves, judges are just as likely as anyone else to defer to apparent authority. 72 Assuming that judges have no training or experience in whatever the highly specialized subjects at issue might be, all they will know is that someone avers that the evidence at issue is real science. Especially if that someone has an advanced degree—well, the judge figures, maybe it is. This is related to another human factor urging admission: The desire to be agreeable.73 Across cultures, when people who do not speak the language are asked a question they do not understand, they smile and say "yes."74

And if there should be a battle of the experts, such that the judge must pick to whom to defer, the structure of the system also presses for admission over exclusion. Whether tried to the bench or to a jury, erroneous admission of evidence is simply less likely to trigger a reversal than erroneous exclusion. One can always argue that wrongly admitted evidence was ignored; there is no way to argue that wrongly excluded evidence was considered. Because the burden of proof rests on the government, admitting more evidence in general means admitting more prosecution-friendly evidence. Because judges are not scientists, the human element ensures that this tendency applies regardless of the validity of the science. Thus, while politics likely explain the hostility to even good, defense-friendly science, both human and structural elements likely explain the warmth to even bad, prosecution-friendly evidence.

IV. Conclusion

This is merely conjecture, but a quick glance at a few examples tends to argue that, almost regardless of the quality of the science, judges are more likely to admit scientific evidence when it is pro-prosecution, and to exclude it when it is pro-defense. The science behind fingerprint and tool mark evidence is notoriously bad, and yet it is routinely admitted. The science behind eyewitness identification limitations is beyond reproach, and yet it is routinely excluded. And the obvious alternative explanation—that the meaningful variable is hard versus soft science, rather than pro-prosecution versus pro-defense—is belied by courts' eagerness to admit evidence of the soft, lousy science of future dangerousness.

"Daubert or no Daubert," the sorts of evidence that are being admitted or excluded seem unrelated to any articulated test for admissibility. The reasons include inertia, politics, and an unintended consequence of the defense-friendly burden of proof, combined with a healthy dose of the kind of human fallibility that results from asking

72. This tendency is so common that pattern jury instructions caution against it. For example:

Remember, you jurors are the sole judges of the credibility and weight of all testimony. The fact that the lawyers and I may have referred to certain witnesses as "experts," and that the witnesses may have special knowledge or skill, does not mean that their testimony and opinions are right.


74. Children do this as well. See e.g. David L. Faigman et al., Modern Scientific Evidence: The Law and Science of Expert Testimony § 14-2.3.7, 183 (2d ed. 2002) (explaining children's "strategies to cover up their limitations, such as... providing affirmative answers to yes/no questions even if they do not understand
judges to play scientist.

On reflection, a preference for pro-prosecution scientific evidence even when it is bad science, and a hostility to pro-defense scientific evidence even when it is good, make perfect sense. Let me know about that grant proposal.