

A TALE OF TWO *DAUBERTS*

*Julie A. Seaman**

TABLE OF CONTENTS

| | | |
|------|--|-----|
| I. | INTRODUCTION | 890 |
| II. | THE EVIDENCE IS RELIABLE, THE EVIDENCE IS UNRELIABLE: TWO <i>DAUBERTS</i> | 894 |
| III. | WE KNOW EVERYTHING, WE KNOW NOTHING: A PLEA FOR EPISTEMIC HUMILITY | 914 |
| IV. | CONCLUSION..... | 921 |

* Associate Professor, Emory University School of Law. J.D., Harvard Law School. I'm grateful to Brad Strickland and Sarah Sternlieb for their excellent and tireless research assistance. Discussion with my co-panelists, Ed Imwinkelried and Ron Carlson, as well as with many of the other participants in the Symposium, was invaluable to me in developing the ideas presented in this Article. *Georgia Law Review* Symposium Editor Clare Ellis deserves particular accolades for her outstanding work in conceiving, organizing, and bringing to fruition the Symposium and these Articles.

I. INTRODUCTION

“[I]t was the age of wisdom, it was the age of foolishness, it was the epoch of belief, it was the epoch of incredulity”¹

A notable feature of the Federal Rules of Evidence is their relatively uniform applicability across civil and criminal cases.² There is something appealing about the notion that, for the most part, relevance is relevance, hearsay is hearsay, and reliability is reliability,³ no matter whether the trial concerns a slip-and-fall, a business dispute, or a murder.

But this surface sameness obscures certain systematic differences in application. In particular, several scholars have noted that the *Daubert* standard seems to be applied differently in criminal versus civil cases.⁴ Professor David Faigman, for

¹ CHARLES DICKENS, *A TALE OF TWO CITIES* 1 (1859).

² See FED. R. EVID. 1101(b) (“These rules apply in . . . civil cases [and] criminal cases and proceedings.”); Erica Beecher-Monas, *Reality Bites: The Illusion of Science in Bite-Mark Evidence*, 30 CARDOZO L. REV. 1369, 1370 (2009) (noting that the “goal of accuracy applies to both civil and criminal cases, and the same rules governing admissibility of expert testimony apply to both contexts,” though, as she notes, Georgia is an exception).

³ See, e.g., David L. Faigman, *Admissibility Regimes: The “Opinion Rule” and Other Oddities and Exceptions to Scientific Evidence, the Scientific Revolution, and Common Sense*, 36 SW. U. L. REV. 699, 716 (2008) (“Although the *Daubert* test was first framed in a civil case amidst the swirl of controversy surrounding fears of exploding litigation, the rule applies similarly to criminal cases.”).

⁴ See, e.g., Beecher-Monas, *supra* note 2, at 1370 (“In practice, however, despite the common goal of accurate factfinding and the common threshold of relevance and reliability, judicial application of gate-keeping standards in civil and criminal trials could not be more different.”); Margaret A. Berger, *Expert Testimony in Criminal Cases: Questions Daubert Does Not Answer*, 33 SETON HALL L. REV. 1125, 1125 (2003) (observing that while courts in civil cases engage in rigorous gatekeeping, there is no sign of a parallel approach in criminal cases); Faigman, *supra* note 3, at 706 (noting that “federal courts . . . implicitly relax their restrictive/non-deferential rule of decision for certain kinds of expert evidence, in particular, forensic science”); Paul C. Giannelli, *The Supreme Court’s “Criminal” Daubert Cases*, 33 SETON HALL L. REV. 1071, 1073 (2003) (remarking upon the civil-criminal dichotomy); D. Michael Risinger, *Navigating Expert Reliability: Are Criminal Standards of Certainty Being Left on the Dock?*, 64 ALB. L. REV. 99, 110 (2000) (“[I]t seems that civil defendants win their *Daubert* dependability challenges most of the time, and that criminal defendants virtually always lose their dependability challenges.”); Joseph Sanders, *Applying Daubert Inconsistently? Proof of Individual Causation in Toxic Tort and Forensic Cases*, 75 BROOK. L. REV. 1367, 1368 (2010) (observing the lack of rigorous gatekeeping in criminal cases as compared to civil cases).

example, has pointed out that “[w]hile *Daubert* ostensibly applies in the same way in criminal and civil cases, social scientists have increasingly raised the issue whether courts, in fact, employ *Daubert* more lackadaisically in criminal trials—especially in regard to prosecution evidence.”⁵ Here in Georgia, recently the forty-fourth state to adopt the Federal Rules of Evidence,⁶ the legislature ultimately decided to depart from the Rules in several particulars; one of these departures entails retaining the former Georgia standard of scientific reliability in criminal cases.⁷ Prior to January 1, 2013, Georgia evidence law provided that “[i]n criminal cases, the opinions of experts on any question of science, skill, trade, or like questions shall always be admissible; and such opinions may be given on the facts as proved by other witnesses.”⁸

Under the new Georgia evidence code, this language is retained.⁹ While it is difficult to imagine a more lenient standard of admissibility than “shall always be admissible,” the Georgia Supreme Court held in *Harper v. State* that “it is proper for the trial judge to decide whether the [scientific] procedure or technique in question has reached a scientific stage of verifiable certainty, or . . . whether the procedure ‘rests upon the laws of nature.’”¹⁰ Despite its language, however, application of the *Harper* test is more lenient than *Daubert*, as evidenced by the long—and ultimately successful—campaign by “tort reform” advocates to persuade the legislature to replace it with the *Daubert* standard in civil cases.¹¹

⁵ Faigman, *supra* note 3, at 716.

⁶ See Ga. Legis. 52 § 1 (2011) (“It is the intent of the General Assembly in enacting this Act to adopt the Federal Rules of evidence, as interpreted by the Supreme Court of the United States and the United States circuit courts of appeal as of January 1, 2013, to the extent that such interpretation is consistent with the Constitution of Georgia.”).

⁷ See O.C.G.A. § 24-7-702(a) (2013). Other departures from the Federal Rules are detailed in Paul S. Milich, *Georgia’s New Evidence Code—An Overview*, 28 GA. ST. U. L. REV. 379 (2012).

⁸ O.C.G.A. § 24-9-67 (repealed effective January 1, 2013 by 2011 Ga. Laws Act 52 (H.B. 24)).

⁹ See O.C.G.A. § 24-7-707 (effective January 1, 2013).

¹⁰ *Harper v. State*, 292 S.E.2d 389, 395 (Ga. 1982).

¹¹ Furthermore, *Harper* allows for liberal judicial notice of a technique’s reliability; evidence of reliability is required only in the case of “novel” scientific techniques. “[O]nce a procedure has been utilized for a significant period of time, and expert testimony has been received thereon in case after case, the trial court does not have to keep reinventing the

Georgia thus remains unusual—perhaps unique—for its explicit acknowledgement¹² that different standards govern the admissibility of scientific evidence in criminal and civil cases.¹³ In light of this feature of Georgia’s new rules, it seems an opportune moment to reflect upon the virtues (and vices) of uniformity in the arena of expert opinion testimony and scientific evidence.

In Part II, this Article offers some tangible data that, while hardly a scientific sample, suggests that the *Daubert* standard indeed may be disparately applied to even very similar evidence when offered in criminal versus civil cases. This discussion is meant to supplement prior analyses, which were global in nature and did not try to compare like expertise to like. A narrower look at admissibility decisions in two specific areas of topical expertise that are common to civil and criminal cases suggests that courts at times do tend to treat even very similar types of scientific or

wheel; a once novel technology can and does become commonplace.” *Hawkins v. State*, 476 S.E.2d 803, 807 (Ga. Ct. App. 1996). However, the Georgia Court of Appeals has subsequently made clear that defendants may nonetheless challenge the expert’s application of the technique in the particular case. *See State v. Tousley*, 611 S.E.2d 139, 144–45 (Ga. Ct. App. 2005) (noting that “even though HGN testing generally is admissible under *Harper v. State*, we have recognized that a defendant may . . . challenge the method by which the HGN test was administered to him or her”).

¹² In the federal courts, where a uniform standard ostensibly applies, a more or less explicit acknowledgment occasionally peeks through. In *United States v. Prime*, 220 F. Supp. 2d 1203 (W.D. Wash. 2002), *aff’d*, 363 F.3d 1028 (9th Cir. 2004), “the court [came] as close to formally embracing the ‘civil plaintiff proffer vs. prosecution proffer’ double standard of acceptable reliability as one can imagine a court explicitly doing.” D. Michael Risinger, *Appendix: Cases Involving the Reliability of Handwriting Identification Expertise Since the Decision in Daubert*, 43 TULSA L. REV. 477, 525 (2007). There, the court reasoned that certain “time-tested” forensic techniques used by law enforcement should not be excluded simply because of a lack of scientific data, methods, or statistical significance. *See Prime*, 220 F. Supp. 2d at 1210.

¹³ Georgia adopted *Daubert* for civil cases by statute in 2005. Prior to that time, Georgia applied the *Harper* test to scientific evidence in both civil and criminal cases. *See Harper*, 292 S.E.2d at 395. Georgia courts have held that the somewhat heightened reliability standard set out in *Harper* applies only to scientific, and not technical or skilled, evidence. *See Salinas v. State*, 722 S.E.2d 432, 435 (Ga. Ct. App. 2012) (“Since the officers’ observations were not a matter of scientific principle or technique, the *Harper* standards did not apply.”). This is in contrast to Federal Rule 702 and *Daubert*, which apply to all expert opinion testimony. *Kumho Tire Co. v. Carmichael*, 526 U.S. 137, 138 (1999). In Georgia, the *Harper* test continues to apply to scientific evidence in criminal cases. *See Vaughn v. State*, 722 S.E.2d 212, 215 (Ga. 2007) (reaffirming that the legislature’s adoption of the *Daubert* test as part of Georgia’s Tort Reform Act did not change the standard in criminal cases, which remained the traditional *Harper* test).

forensic evidence differently depending on whether it is offered in a criminal or a civil case. On the other hand, it seems that federal courts do not always show the same disparity in treatment; the contrast between civil and criminal cases is context-dependent.

I then indulge, in Part III, in a brief detour to consider the case for “epistemic humility” and the relevance such a stance might have in the *Daubert* debate. Science, after all, is an uncertain enterprise, as Professor Imwinkelried emphasized nearly two decades ago: “The adherents to the traditional conception of science tended to think that each discovered truth was immutable In contrast, modern scientists realize not only that truth is ‘extremely mutable’ but also that progress is often discontinuous.”¹⁴ Yet this inherent uncertainty and the consequent prospect of discontinuity is in tension with the legal and judicial temperament; this is particularly so in the criminal context, where the trend is toward finality and “closure”¹⁵ even in the face of legal or factual error.¹⁶ This uncertainty, which is not only metric but also conceptual,¹⁷ is obscured by the rhetoric of the *Daubert* inquiry. A more transparent approach, as is at least made possible by Georgia’s frank adoption of a non-*Daubert* standard in criminal cases, might on balance lead to more accurate verdicts in cases involving forensic evidence.

I therefore advance the tentative, uncertain, and mutable position that Georgia has it right in being explicit about having two distinct tests. The Federal Rules of Evidence—and the states

¹⁴ 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, 61 (1995) (footnotes omitted).

¹⁵ See Susan Bandes, *Victims, “Closure,” and the Sociology of Emotion*, 72 LAW & CONTEMP. PROBS. 1, 1 (2009) (observing the recasting of various aspects of the criminal justice system, particularly in capital cases, to serve the therapeutic goal of “closure”).

¹⁶ See, e.g., *Cullen v. Pinholster*, 131 S. Ct. 1388 (2011) (holding that a federal court on habeas corpus review cannot hold an evidentiary hearing to consider evidence of ineffective assistance of counsel, but instead is limited to the factual record developed in the state court); *Dist. Att’y’s Office v. Osborne*, 557 U.S. 52 (2009) (holding that the Due Process Clause does not require that convicted defendants be permitted access to biological evidence for DNA testing at their own expense to support a claim of actual innocence).

¹⁷ Professor Imwinkelried has discussed both genres of uncertainty, see Imwinkelried, *supra* note 14, at 63, though his recent work emphasizes uncertainty of measurement, see Edward J. Imwinkelried, *Forensic Metrology: The New Honesty About the Uncertainty of Measurements in Scientific Analysis*, available at <http://ssrn.com/abstract=2186247>.

that follow them in this respect—would do well to consider the potential downside of a uniform standard in the context of expert opinion and scientific evidence.

II. THE EVIDENCE IS RELIABLE, THE EVIDENCE IS UNRELIABLE: TWO *DAUBERTS*

As noted above, in the years since the Supreme Court issued its landmark decision in *Daubert v. Merrell Dow Pharmaceuticals, Inc.*,¹⁸ several scholars have remarked on the apparent tendency of courts to be more lenient in admitting expert testimony in civil as compared to criminal cases.¹⁹ I suspect that few would argue with the proposition that many prosecutorial applications of the forensic “sciences” that are routinely admitted, that have long been admitted, and that continue to be admitted despite the serious questions raised in the National Academy of Sciences (NAS) Report²⁰ and by many commentators,²¹ would not satisfy a *Daubert* inquiry as written and generally applied, for example, to plaintiff evidence of causation in toxic tort cases.²² We know that there are serious reliability issues with latent fingerprint

¹⁸ 509 U.S. 579 (1993).

¹⁹ See *supra* note 4 and accompanying text. The most rigorous study of this disparity is Risinger, *supra* note 4, at 104. Professor Risinger looked at all federal and state cases that cited *Daubert* through 1999.

²⁰ See NAT'L RES. COUNCIL, STRENGTHENING FORENSIC SCIENCE IN THE UNITED STATES: A PATH FORWARD 7 (2009) [hereinafter NAS REPORT], available at <http://www.ncjrs.gov/pdffiles1/nij/grants/228091.pdf> (noting that “[w]ith the exception of nuclear DNA analysis . . . no forensic method has been rigorously should to have the capacity to consistently, and with a high degree of certainty, demonstrate a connection between evidence and a specific individual or source”).

²¹ “For decades, legal scholarship has been stressing the lack of empirical research in some forensic disciplines.” Paul C. Giannelli, *The 2009 NAS Forensic Science Report: A Literature Review*, 48 CRIM. L. BULL. 378 (2012) (citing several examples).

²² Professor Risinger makes the point in his wonderful 2000 survey of post-*Daubert* cases by urging the reader to ask the following question:

If, after *Daubert*, substantial liability of General Motors or Microsoft were dependent on the identification of bite marks found in various non-ideal media, and on their attribution to various corporate employees, is it not clear that these issues would have been litigated differently and more thoroughly than they have been, and that the results would have often been different?

Risinger, *supra* note 4, at 143.

identification evidence,²³ handwriting identification,²⁴ ballistics,²⁵ bite marks,²⁶ and the way that these are often presented to criminal juries.²⁷ We know that other kinds of evidence that are backed up by relatively robust findings—for example, expert testimony about the dangers of unreliable eyewitness identification²⁸—have often been excluded, and that these tend to be offered primarily by criminal defendants.²⁹

²³ *United States v. Mitchell*, 365 F.3d 215, 221 (3d Cir. 2004); see also Jennifer L. Mnookin, *The Validity of Fingerprint Identification: Confessions of a Fingerprinting Moderate*, 7 J.L. PROB. & RISK 127, 127–30 (2008).

²⁴ See, e.g., Risinger, *supra* note 12, at 480–94 (summarizing and evaluating the extant studies on the reliability of forensic document examination).

²⁵ Paul C. Giannelli, *Ballistics Evidence Under Fire*, 25 CRIM. JUST. 50, 51 (2011).

²⁶ See, e.g., Beecher–Monas, *supra* note 2, at 1372; Allen P. Wilkinson & Ronald M. Gerughty, *Bite Mark Evidence: Its Admissibility Is Hard To Swallow*, 12 W. ST. U. L. REV. 519, 520 (1985) (examining the state of the art of forensic odontology and bite mark analysis in order to determine whether identification via these techniques should be admissible); Adam Deitch, Comment, *An Inconvenient Tooth: Forensic Odontology Is an Inadmissible Junk Science When It Is Used To “Match” Teeth to Bitemarks in Skin*, 2009 WIS. L. REV. 1205, 1215–16 (2009) (discussing the lack of a scientific basis for bitemark evidence).

²⁷ Many scholars have criticized the level of certainty claimed by many forensic experts in their trial testimony (and also in their professional literatures). See, e.g., Paul C. Giannelli, *The NRC Report and Its Implications for Criminal Litigation*, 50 JUR. J. 53, 57–61 (2009) (noting problems of claims of zero error rate, one hundred percent accuracy, and exclusion of all other possible sources). As elucidated recently by Professor Imwinkelried, this practice was closely entwined with decisional rules that required certainty, which rules in turn flowed from a particular view of the physical sciences. See Imwinkelried, *supra* note 17, at 7.

²⁸ Lauren Tallent, Note, *Through the Lens of Federal Evidence Rule 403: An Examination of Eyewitness Identification Expert Testimony Admissibility in the Federal Circuit Courts*, 68 WASH. & LEE L. REV. 765 *passim* (2011) (discussing the Circuit split over admissibility of eyewitness identification expert testimony and proposing a solution). Though the United States Supreme Court recently rejected a claim that due process requires heightened reliability protections in the context of eyewitness identification evidence, see *Perry v. New Hampshire*, 132 S. Ct. 716, 721 (2011), some states have moved to try to ensure greater reliability in this area, see *State v. Lawson*, 291 P.3d 673, 691–97 (Or. 2012); *State v. Henderson*, 27 A.3d 872, 920–22 (N.J. 2011).

²⁹ See, e.g., Julie A. Seaman, *Triangulating Expert Testimony: The Constitutional Boundaries of Expert Opinion Testimony*, 96 GEO. L.J. 827, 834 & n.24 (2008) [hereinafter Seaman, *Triangulating Expert Testimony*]. In addition, polygraph evidence is more often proffered by the defense, and though its reliability is highly contested, it is at least arguably as reliable as bite mark or ballistics evidence. See Julie A. Seaman, *Black Boxes*, 58 EMORY L.J. 427, 460–61 (2008) [hereinafter Seaman, *Black Boxes*] (noting that “the judicial reaction against credibility expertise, and lie detection evidence in particular, has been wholly out of proportion to its purported lack of scientific reliability”).

Some scholars approach this as a competence question and wonder why courts seem able to critically evaluate the reliability of scientific and technical evidence in civil cases but not in criminal cases.³⁰ A recent survey of scholarship generated in the wake of the 2009 NAS Report on forensic sciences pointed out that “several commentators [have] contrasted the courts’ apparent capacity to deal with complex toxic tort litigation with their hands-off policy in criminal litigation,” quoting one scholar who stated that

[i]n civil cases, courts seem quite up to the task of evaluating microbiology, teratology, and toxicology evidence . . . Yet when it comes to evaluating the shortcomings of lip prints and handwriting, courts are unable to muster the most minimal grasp of why a standardless form of comparison might lack evidentiary reliability or trustworthiness.³¹

While this lack of consistency in the application of *Daubert* is widely acknowledged among evidence scholars, this Article seeks to explore these questions by gathering data on the treatment of certain very specific types of expert testimony to further examine the nature of the difference in the application of the *Daubert* standard in criminal versus civil cases, and also (likely related) as among criminal defendants, civil defendants, prosecutors, and plaintiffs. It also considers whether the publication of the 2009 NAS Report on forensic evidence might have had any effect on the way in which judges approach forensic expert opinion testimony in criminal (or civil) cases.

With these tasks in mind, I searched for federal cases involving forensic expert testimony of uncertain reliability that also is offered more than occasionally in civil cases. Ultimately, two

³⁰ Judge Nancy Gertner recently noted: “We, the courts, can do better. In fact, we already do, albeit in civil, not criminal, cases.” Nancy Gertner, *National Academy of Sciences Report: A Challenge to the Courts*, 27 CRIM. JUST. MAG. 8 (2012), available at http://www.americanbar.org/content/dam/aba/publications/criminal_justice_magazine/sp12_report.pdf.

³¹ Giannelli, *supra* note 25, at 385 (quoting Jane Campbell Moriarty, *Will History Be Servitude?: The NAS Report on Forensic Science and the Role of the Judiciary*, 2010 UTAH L. REV. 299, 315).

categories of evidence emerged that seemed relatively comparable across the range of civil and criminal cases. In general, the results are mixed. They support the view that there is one *Daubert* standard for civil cases and quite another for criminal cases, at least in certain realms of forensic expertise; they also suggest that, at least in some areas, criminal defendants' expert testimony, when offered, is less likely to be admitted than similar prosecution expert testimony. This disparity also exists—though flipped in the other direction to favor defense experts—in civil cases. However, it also seems that generalizations are somewhat misleading, and that courts are more inclined to scrutinize some kinds of forensic evidence than others.

For a variety of reasons that others have discussed, it is not a simple matter to compare *Daubert* rulings across civil and criminal cases systematically: Expert testimony in civil cases versus criminal cases often looks like apples and oranges. As Professor Risinger points out in his global survey of post-*Daubert* federal decisions:

[B]ecause different forms of expertise are commonly proffered in civil and criminal cases, these numbers [showing very large differences in admission and exclusion rates in civil and criminal cases and also depending on which party offered the evidence] do not directly establish disparate standards of dependability in the two contexts, but they are fairly striking in their own right. Maybe it is true that the prosecution always proffers highly dependable expertise, and that criminal defendants and civil plaintiffs usually proffer garbage, or that prosecutors and civil defendants only object to low quality proffers whereas criminal defendants (and to a lesser extent civil plaintiffs) object to demonstrably dependable evidence as a matter of course.³²

³² Risinger, *supra* note 4, at 108.

Certainly these are potentially confounding variables when comparing civil and criminal cases.³³ As a small attempt to liken apples to apples when examining this disparity in courts' application of *Daubert*, this section considers some of the relatively few types of scientific or technical expertise that tend to be offered in both criminal and civil cases. Two such potentially fruitful areas for comparison are handwriting analysis and fire cause and origin evidence. While none of the sample sizes is very large,³⁴ these two categories present enough cases to at least be suggestive of patterns that reveal a distinct difference between federal civil and criminal application³⁵ of *Daubert*.

The field of questioned document analysis, which includes handwriting identification,³⁶ has been the subject of significant critical commentary in the academic literature³⁷ (though not so

³³ Professor Joseph Sanders takes a similar approach in comparing admissibility rulings concerning alcohol testing in drunk-driving cases to forensic admissibility decisions in other criminal cases. He concludes that, under certain circumstances, judges indeed are willing to exclude what they perceive to be unreliable scientific evidence. See Joseph Sanders, "Utterly Ineffective": Do Courts Have a Role in Improving the Quality of Forensic Expert Testimony?, 38 FORDHAM URB. L.J. 547, 564 (2010) (noting that "[d]runk driving admissibility cases indicate that if courts are able to distinguish cases based on the quality of the evidence, they will exclude evidence"). This is a crucial point: Another area with strong indications that evidence may be of questionable reliability is eyewitness identification testimony (and some other eyewitness testimony more generally). See *supra* note 28 and accompanying text. However, it is nearly impossible for courts, juries, or anyone else to know whether in any particular case the eyewitness's identification suffered from the potential flaws revealed in the academic literature.

³⁴ As it turns out, in the federal courts there are many more civil than criminal cases involving arson or other fire origin testimony.

³⁵ I confined my analysis to federal cases in order to be sure that the standard of reliability being applied was at least formally consistent across the cases.

³⁶ As described in the NAS REPORT, "[q]uestioned document examination involves the comparison and analysis of documents and printing and writing instruments in order to identify or eliminate persons as the source of the handwriting; to reveal alterations, additions, or deletions or to identify or eliminate the source of typewriting or other impression marks." *Supra* note 20, at 163.

³⁷ Professor Risinger's work in this area is exhaustive and generally quite critical of the way that the vast majority of courts—both trial level and appellate—have evaluated the reliability of forensic document testimony. See generally D. Michael Risinger & Michael J. Saks, *Science and Nonsense in the Courts: Daubert Meets Handwriting Identification Expertise*, 82 IOWA L. REV. 21 (1996) (examining the history and evolution of handwriting identification expertise); D. Michael Risinger, *Goodbye to All That, or a Fool's Errand, by One of the Fools: How I Stopped Worrying About Court Responses to Handwriting Identification (and Forensic Science in General) and Learned To Love Misinterpretations of Kumho Tire v. Carmichael*, 43 TULSA L. REV. 447 (2007) (criticizing confused application of

much in the courts). In the case of forensic handwriting analysis, research revealed forty-five written opinions in which federal courts considered the admissibility of such evidence under *Daubert* in criminal cases.³⁸ In all but one of those cases, the evidence was offered by the prosecution.³⁹ It was completely excluded in five of

key precedent); Risinger, *supra* note 12 (same). The discussion in this section of handwriting cases relies heavily on Professor Risinger's work.

³⁸ A research assistant compiled the initial case list, which I then supplemented using Professor Risinger's Appendix and other materials. Following Risinger's lead, I omitted the small number of cases that involved forensic "stylistics" testimony, though these are perhaps even more striking than the more conventional handwriting identification cases in their casual treatment of their *Daubert* gatekeeping responsibilities. See *United States v. Van Wyk*, 83 F. Supp. 2d 515 (D.N.J. 2000); *United States v. Zajak*, 748 F. Supp. 2d 1340 (D. Utah 2010); see also *United States v. Salameh*, 54 F. Supp. 2d 236, 251 (S.D.N.Y. 1999) (denying a claim of ineffective assistance of counsel in World Trade bombing case for failure to retain a linguistics expert and noting that the proposed expert testimony, that defendant's writing skills were not sufficient for him to have authored a particular document, "is both immaterial and speculative and would not have been admitted into evidence"). In *Van Wyk*, the relevant question was whether the defendant had written threatening communications to several different victims. Some of these were handwritten and others were typed. The prosecution proffered the testimony of an FBI agent, who had compared stylistic elements of the questioned documents with known writings of the defendant—grammar, punctuation, idiosyncratic phrasing, and the like—that in his opinion the defendant was the author of the questioned documents. The court, after engaging in a fairly critical discussion of the field of forensic stylistic analysis, held that the agent was qualified under Rule 702 and that he would be permitted to point out relevant similarities and differences, though not to give his ultimate conclusion about defendant's authorship of the documents. 83 F. Supp. 2d at 525. As Professor Risinger has pointed out in the context of handwriting identification,

it is not clear how the rather Solomonic decision of restricting the document examiner to pointing out similarities really helps in practice, since the conclusion of common source will be clearly implied by the form of the testimony, although in fairness, such a restriction might reduce its impact and the weight given to it by the jury.

Risinger, *supra* note 12, at 501. In *Zajak*, the court summarized at length the defense expert's *Daubert* hearing testimony regarding the lack of reliability of the prosecution witness's conclusions as to common authorship of three bomb threat letters. It then held, following *Van Wyk*, that the witness would be permitted to point out certain similarities and differences and to testify about their relevance based on his experience, but that he could not offer an opinion about common authorship of the letters. 748 F. Supp. 2d at 1351.

³⁹ In the single case in which the defendant sought to introduce handwriting identification testimony, the evidence was excluded. See *United States v. Garza*, 448 F.3d 294, 300 (5th Cir. 2006). There was also a handful of cases in which the defense sought to introduce expert testimony to the effect that handwriting identification fails to meet the *Daubert* standard by calling academic researchers, such as Professor Michael J. Saks, to testify to the state of the literature on the reliability of such evidence. Also in line with Professor Risinger's reasoning, I did not include these rulings on defense experts who are not themselves forensic document examiners but rather are academics who would offer

the forty-five cases, and thus admitted in some form in the other forty—an admissibility rate of nearly *eighty-nine percent*. Moreover, this group of cases includes many in which the “task at hand” was that most difficult and problematic one of assigning authorship to a small sample of writing that the author was possibly attempting to disguise in some way.⁴⁰ It also includes several cases in which courts expressed some skepticism about the qualifications of the expert witness, but nonetheless determined that the witness was qualified under Rule 702 and that such questions went to weight rather than admissibility. In the single case⁴¹ in which the defense attempted to introduce testimony by a forensic document examiner, the evidence was excluded as unreliable under Rule 702.⁴²

In civil cases, courts’ treatment of handwriting expertise under *Daubert* tells quite a different story. Whereas in criminal cases, for the most part, the global field of questioned document analysis is one with a long history, tested in the crucible of the adversarial process and relied upon by law enforcement and overwhelmingly approved by courts, in civil cases the field is peopled by unqualified charlatans who use untested methodologies to offer unreliable opinions that are not helpful to juries, which are perfectly capable of comparing handwriting samples on their own.⁴³ Of the eleven⁴⁴ federal civil cases that resulted in written

opinions about the reliability of handwriting identification.

⁴⁰ See, e.g., Risinger, *supra* note 12, at 499 (discussing document examiner’s testimony that a forged signature was written by the defendant).

⁴¹ Of course, this set only includes decisions that resulted in written opinions, reported, or (as many are) “unreported.” Where the number of cases is relatively small, references are provided for all of them; for larger sets, representative examples are referenced in the notes.

⁴² *Garza*, 448 F.3d at 300. Thus, the prosecution’s success record is 40/44, or nearly 91%. The defense is 0 for 1: 0.0%.

⁴³ If this is an exaggeration, it is only a slight one.

⁴⁴ A district court opinion and an appellate court opinion in the same case are counted separately here. See *Deputy v. Lehman Bros.*, No. 2:02-CV-00718-RTR, 2003 WL 24305008 (E.D. Wis. Jan. 7, 2003); *rev’d*, *Deputy v. Lehman Bros.*, 345 F.3d 494 (7th Cir. 2003). Likewise, in the criminal case set, there are two opinions generated by the same case. See *United States v. Prime*, 220 F. Supp. 2d 1203 (W.D. Wash. 2002), *aff’d*, 363 F.3d 1028 (9th Cir. 2004). This is in line with Professor Risinger’s practice in his Appendix, see Risinger, *supra* note 12, at 479, though arguably it is more justifiable for a comprehensive listing of cases than for purposes of generating statistics on admissibility rulings. If the lower court decision were excluded from the civil set, the exclusion rate would be six of ten, or 60%.

decisions on the admissibility of handwriting expertise, the evidence was *excluded* in seven, or 64%.⁴⁵ And of the remaining cases in which the evidence was not excluded, one reversed the district court's ruling excluding the testimony because the lower court erred by considering the credibility of the witness in its Rule 702 analysis and remanded the case for a proper Rule 702 determination.⁴⁶ Thus, it is not clear that this should be counted as a ruling in favor of admissibility. Of the remaining three cases that held the expert testimony admissible under *Daubert*,⁴⁷ one restricted the witnesses to pointing out relevant similarities and differences.⁴⁸

One wrinkle in the data is the question of how to treat this "compromise position" of permitting the expert to testify to relevant similarities and differences between the questioned handwriting and the known samples but not to give his or her conclusion as to authorship. Of the forty federal criminal cases in which the handwriting evidence was admitted, nine courts imposed this restriction—often referred to as the "*Hines/McVeigh*" approach—on the testimony of the expert witness.⁴⁹ And, as noted

⁴⁵ In one of these, the plaintiff offered the testimony of two forensic document examiners. One was rejected as unqualified, and the other was restricted to pointing out similarities and differences, with the result that the defendant's motion for summary judgment was granted. Because the *Daubert* rulings directly caused the plaintiff's case to fail, I counted this case as an exclusion. See *Wolf v. Ramsey*, 253 F. Supp. 2d 1323, 1344, 1347–48 (N.D. Ga. 2003).

⁴⁶ See *Deputy v. Lehman Bros.*, 345 F.3d 494, 511 (7th Cir. 2003). Later reported opinions revealed that the jury determined that one of the disputed signatures was a forgery, but there was no further mention of the admissibility of the forensic document examiner's testimony. See *Deputy v. Lehman Bros.*, 374 F. Supp. 2d 695 (E.D. Wis. 2005).

⁴⁷ See *Truman Arnold Cos. v. Green*, No. 5:03-CV-45, 2006 WL 5153151, at *1 (E.D. Tex. Sept. 28, 2006); *A.V. by Versace v. Gianni Versace S.p.A.*, 446 F. Supp. 2d 252, 267–68 (S.D.N.Y. 2006); *Legacy Vision, LLC v. Yeaman*, No. CIV-04-1320-M, 2005 WL 6227149, at *7 (W.D. Okla. June 6, 2005).

⁴⁸ *Legacy Vision*, 2005 WL 6227149, at *7. This restriction was also imposed on one of the two proffered experts in *Wolf*, 253 F. Supp. 2d at 1347–48. The other was excluded as unqualified under Rule 702. *Id.* at 1344.

⁴⁹ See *United States v. Hernandez*, 42 F. App'x 173, 175 (10th Cir. 2002); *United States v. Oskowitz*, 294 F. Supp. 2d 379, 384 (E.D.N.Y. 2003); *United States v. Hidalgo*, 229 F. Supp. 2d 961, 968 (D. Ariz. 2002); *United States v. Rutherford*, 104 F. Supp. 2d 1190, 1193–94 (D. Neb. 2000); *United States v. Brown*, No. CR 99-184 (C.D. Cal. Dec. 1999), as described in *Risinger*, *supra* note 12, at 512 & nn.202–03; *United States v. Santillan*, No. CR-96-40169, 1999 WL 1201765, at *5 (N.D. Cal. 1999); *United States v. Hines*, 55 F. Supp. 2d 62, 70–71 (D. Mass. 1999); *United States v. McVeigh*, 1997 WL 47724 (D. Colo. Feb. 5, 1997), as

above, a similar restriction was imposed in one of the three civil cases in which courts admitted handwriting evidence under *Daubert*.

Assuming, as I do here, that judges are quite reluctant in criminal cases to exclude prosecution evidence that carries a long historical pedigree even where they have some concern about its reliability,⁵⁰ it is arguably better than nothing to admit it with some reasonable limitation.⁵¹ Thus, we might think that the *Hines/McVeigh* approach is a move in the right direction, perhaps even a signal of a trend toward stricter gatekeeping in the area of forensic testimony. On the other hand, it is questionable whether this limitation provides very much protection from the potentially unreliable conclusions against which it is directed:

[T]here is a serious problem with this, especially if the document examiner is allowed to recite his or her credentials, titles, and job descriptions. By identifying the similarity or difference, the examiner is inevitably perceived as asserting the significance of those similarities or differences in regard to assigning authorship, so that the conclusions which are barred are easily inferred. In practice, this is profoundly true, since document examiners who believe they have identified the author of a writing by comparison will normally point out only similarities, and if differences are called to their attention they will dismiss them as

described in Risinger, *supra* note 12, at 500 & nn.127–33; *United States v. Starzecpyzel*, 880 F. Supp. 1027, 1046–47 (S.D.N.Y. 1995). Though it seemed for a time that this approach was gaining momentum, more recently it appears to have fallen out of favor. According to Professor Risinger's assessment of the current state of federal case law on handwriting analysis, "for better or worse, the game seems to be pretty much over, unless the Supreme Court steps in to deal a new hand." Risinger, *supra* note 12, at 561.

⁵⁰ That is the strong sense one is left with after reading those criminal cases in which courts canvas the criticisms and questions but in the end admit the evidence, sometimes with restrictions on the manner in which the expert may state his or her conclusions.

⁵¹ There has also recently been some general movement toward limiting the certainty with which forensic experts are permitted to express their conclusions. See Imwinkelried, *supra* note 14, at 13–15.

not being significant or “real” differences, but merely manifestations of “individual variation.”⁵²

Thus the question becomes whether it is indeed “an improvement over surrendering the gatekeeping function entirely to the guild,”⁵³ or whether instead it might actually be even worse than admitting an unreliable opinion wholesale.⁵⁴ As I argue in the next Part, it may indeed be worse because such 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.

In addition to handwriting analysis and identification, another forensic field that has applications in both civil and criminal cases is fire cause and origin. Experts in this field purport to use various techniques to determine how and where a fire originated, including in many instances whether it was “incendiary,” meaning intentionally set. Contemporary research has cast serious doubt on some of the longstanding assumptions in the field; in recent years, several time-honored techniques have been revealed to have little basis in science.⁵⁵

⁵² Risinger, *supra* note 12, at 510.

⁵³ *Id.*

⁵⁴ In a footnote, Professor Risinger suggests one reason that it could end up being worse: Some court operating under the Hines/McVeigh approach is likely to disallow cross examination of a document examiner on known error rates, such as the 8% document examiner error rate shown on one task in Kam IV, on the ground that the witness is not giving an opinion, even though the implied opinion is clear to the whole courtroom.

Id. at 510 n.186.

⁵⁵ See DAVID L. FAIGMAN ET AL., MODERN SCIENTIFIC EVIDENCE § 39:1 (2012) (noting that “the late 2000s occasioned increasing scrutiny of arson prosecutions and fire investigations,” with greater awareness in the scientific community and the courts of “some persistent misconceptions in fire investigation”). Much of the impetus for this examination stemmed from the case of Cameron Todd Willingham, widely believed to have been wrongfully convicted in the fire deaths of his three young children, for which he was executed in 2004. Although significant questions were subsequently raised as to the evidence supporting his conviction, Texas Governor Rick Perry and the Texas Board of Pardons and Paroles each declined to block the execution. See David Grann, *Trial by Fire: Did Texas Execute an Innocent Man?*, NEW YORKER, Sept. 7, 2009, at 42, 58, 59, available at http://www.newyorker.com/reporting/2009/09/07/090907fa_fact_grann. For example, the fire origin experts testified in the Willingham case that “puddle patterns” on the floor as well as charring of the wood under an aluminum door threshold were clear evidence that liquid

The federal criminal prosecutions that involve such experts are essentially arson prosecutions,⁵⁶ whereas the civil cases are more varied. Some of these, in which insurance companies defend nonpayment by claiming that the insured caused the fire that resulted in the loss, closely resemble arson prosecutions in their factual issues.⁵⁷ In many—probably most—others, the plaintiff or plaintiff's insurer asserts a claim against the manufacturer of a product alleged to have caused the fire at issue in the case.⁵⁸

accelerants were used. *Id.* at 46. As summarized later in a report commissioned by The Innocence Project,

It was widely taught that 'puddle shapes' and 'liquid-type' patterns were unequivocal evidence of accelerants in 1992 when [National Fire Protection Association's *Guide for Fire and Explosions Investigations* 921 (hereinafter NFFPA 921)] was first issued. By 2004, it was well known and generally accepted in the fire investigation community that such patterns were subject to misinterpretation in fully involved compartments, and that the only way to credibly identify a flammable liquid induced fire pattern was to obtain a positive laboratory result.

John J. Lentini et al., *Report on the Peer Review of the Expert Testimony in the Cases of State of Texas v. Cameron Todd Willingham and State of Texas v. Ernest Ray Willis, THE INNOCENCE PROJECT 16* (2006), available at <http://www.innocenceproject.org/docs/ArsonReviewReport.pdf>. The NAS Report's discussion of arson evidence is scant, comprising only a single paragraph. See NAS REPORT, *supra* note 20, at 173.

⁵⁶ Some of the cases involve other crimes, such as mail and wire fraud, that are based on allegations that the defendant intentionally set the fire. It should be noted that there are relatively few opinions in federal criminal cases involving arson expertise, and several of those that do exist are not direct appeals but collateral habeas challenges. I confined this survey to federal cases primarily to ensure that the applicable reliability standard was nominally identical in all of the cases but also so as to keep the civil fire cases to a manageable number. Unfortunately, this strategy resulted in a very small *n* on the criminal side.

⁵⁷ See, e.g., *Newman v. State Farm Fire and Cas. Co.*, 290 F. App'x 106, 114 (10th Cir. 2008) (admitting expert evidence offered by the insurance company, though calling the question "close"); *Nelson v. Safeco Ins. Co. of N. Am.*, 396 F. Supp. 2d 1274, 1279 (D. Utah 2005) (excluding expert testimony offered by the insurance company); *Thompson v. State Farm Fire & Cas. Co.*, 548 F. Supp. 2d 588, 596 (W.D. Tenn. 2008) (admitting expert evidence offered by the insurance company); *Nationwide Ins. Co. v. Johnson*, No. 2:05cv94-KS-JMR, 2006 WL 1666254, at *1 (S.D. Miss. June 12, 2006) (admitting expert evidence offered by the insurance company); *Abu-Hashish v. Scottsdale Ins. Co.*, 88 F. Supp. 2d 906, 907 (N.D. Ill. 2000) (admitting expert evidence offered by the insurance company).

⁵⁸ See, e.g., *Presley v. Lakewood Eng'g and Mfg. Co.*, 553 F.3d 638, 647 (8th Cir. 2009) (excluding plaintiff's expert evidence); *Talkington v. Atria Reclamelucifers Fabrieken BV*, 152 F.3d 254, 265–66 (4th Cir. 1998) (admitting plaintiff's expert evidence in case against cigarette lighter manufacturer, though declining to apply *Daubert* as the suit pre-dated *Kumho Tire*); *Ledbetter v. Blair Corp.*, No. 3:09–CV–843–WKW, 2012 WL 2464000, at *12 (M.D. Ala. June 27, 2012) (admitting plaintiff's expert evidence against seller of bathrobe that ignited); *CIC Partners v. Sunbeam Prods., Inc.*, No. 09–3274, 2012 WL 124982, at 5–6

On the criminal side, a total of six federal cases with written opinions directly considered the reliability of fire origin and cause expertise.⁵⁹ In five of these—83%—the prosecution’s fire expert’s testimony was admitted without restriction.⁶⁰ In the sixth case, the government and the defendant each offered the testimony of fire experts and the court very carefully parsed the proposed testimony, allowing some opinions but excluding others.⁶¹ For example, the government’s expert was permitted to testify that the defendant’s burn injuries were inconsistent with her claim that she was on the stairwell when she was burned,⁶² but the court held that his testimony that the fire would have progressed in a certain manner was not supported by any scientific analysis, experiments, or other studies and was therefore not admissible.⁶³ Similarly, the

(D. Minn. Jan. 17, 2012) (admitting plaintiff’s expert evidence in case against heating pad manufacturer). Though the expert testimony in the latter cases is often similar to that in the arson-type cases, in some cases it tends more toward electrical engineering and knowledge about particular types of appliances, vehicles, or other products. In compiling the list of civil cases, I exercised some judgment and excluded cases that seemed to involve fire expertise only marginally as opposed to products or engineering expertise. For example, I excluded several cases stemming from airplane crashes caused by on-board fires.

⁵⁹ These were: *United States v. Santiago*, 202 F. App’x 399 (11th Cir. 2006); *United States v. Diaz*, 300 F.3d 66 (1st Cir. 2002); *United States v. Gardner*, 211 F.3d 1049 (7th Cir. 2000); *United States v. Markum*, 4 F.3d 891 (10th Cir. 1993); *United States v. Smallwood*, No. 5:08–CR–38, 2010 WL 4168823 (W.D. Ky. Oct. 12, 2010); *United States v. Aman*, 748 F. Supp. 2d 531 (E.D. Va. 2010). In addition, there are a handful of pre-*Daubert* federal arson cases that consider the reliability of the expert testimony under Rule 702. In the pre-*Daubert* case *United States v. Kladouris*, the Seventh Circuit affirmed the district court’s ruling excluding the testimony of the arson defendant’s proffered fire expert. 964 F.2d 658, 669–70 (7th Cir. 1992). In *United States v. Lundy*, the court affirmed the district court’s ruling admitting expert opinion that the fire at issue was purposefully set. 809 F.2d 392, 396 (7th Cir. 1987).

⁶⁰ These were *Aman*, *Santiago*, *Diaz*, *Gardner*, and *Markum*. Both *Diaz* and *Markum* involved plain error review because the appellate courts found that the defendants had not preserved their *Daubert* objections below.

⁶¹ *Smallwood*, 2010 WL 4168823, at *3–12. The defendant apparently did not dispute that the fire was the result of arson but claimed that she did not start the fire. *Id.* at *2.

⁶² *Id.* at *3–4.

⁶³ *Id.* at *4. The defendant claimed that she was “chased up the stairs by a fire rapidly advancing across the floor and up the stairs” and that she extended her legs over a patio roof while awaiting rescue. *Id.* She also claimed that “she attempted re-entry through the bedroom window from which she escaped but was driven back by the smoke and heat.” *Id.* The prosecution’s expert wished to testify that her burns and other injuries were not consistent with this scenario. The court excluded this opinion, reasoning that it was “too speculative in regards to Defendant’s actions and the progression of the fire” because the expert “presented no scientific analysis as to how the fire would have progressed after

court found that his testimony that certain types of burns (suffered by the defendant) were typically observed when that person ignited the fire were unsupported and inadmissible.⁶⁴

In addition to these direct *Daubert* rulings, the case set also included three habeas appeals from federal arson convictions.⁶⁵ The petitioners' claims were denied in two of these⁶⁶ and granted in the third.⁶⁷ Two of these habeas cases—one each granting and denying the petition—very explicitly considered the reliability of the fire origin and cause evidence.⁶⁸ In these two cases, the trial judges considering the petitions made clear how they would have ruled on a *Daubert* challenge to the evidence, had one been advanced by counsel below.⁶⁹ Thus, these two might plausibly be

ignition in the living room as far as traveling up the stairs, down a hall, and into the bedrooms” and “failed to point out any of his own experiments or other studies regarding heat levels, smoke inhalation, and the timing of fires.” *Id.*

⁶⁴ *Id.* at *5–6. With respect to the defense expert, the court likewise permitted him to testify as to the defendant's probable location when the fire started, though as to both experts the court considered this “a close decision.” *Id.* at *8. Other opinions regarding “how Defendant ‘would have carried the baby’ or that Defendant ‘would naturally hold her breath’” were excluded as speculative and unreliable. *Id.* at *9.

⁶⁵ They were: *Thompson v. United States*, 436 F. App'x 669 (7th Cir. 2011); *Schlesinger v. United States*, No. 09–CV–4278, 2012 WL 407098 (E.D.N.Y. Feb. 6, 2012); *United States v. Hebshie*, 754 F. Supp. 2d 89 (D. Mass. 2010). In addition, there is a smattering of federal court opinions in habeas appeals from state arson convictions. See *Richey v. Bradshaw*, 498 F.3d 344 (6th Cir. 2007); *Dugas v. Coplan*, 428 F.3d 317 (1st Cir. 2005); *Jackson v. McQuiggin*, No. 10-12426, 2012 WL 5410993 (E.D. Mich. Nov. 6, 2012). The habeas appeals from federal prosecutions all involve claims of ineffective assistance of counsel based on the failure to challenge the government's arson expert or to investigate and introduce defense expertise on the fire origin and cause issue. The state habeas cases involve either claims of ineffective assistance or a due process claims based on similar defense.

⁶⁶ *Thompson*, 436 F. App'x at 670; *Schlesinger*, 2012 WL 407098, at *28.

⁶⁷ *Hebshie*, 754 F. Supp. 2d at 128.

⁶⁸ *Id.* at 114; *Schlesinger*, 2012 WL 407098, at *15–16. In the third case, petitioner claimed ineffective assistance based on defense counsel's failure to challenge expert testimony to the effect that the fire, which resulted in the death of defendant's elderly mother, was intentionally set. The Seventh Circuit held that this trial strategy was reasonable because the defendant's theory of the case was that his mother had committed suicide by intentionally setting the fire. Thus, the court did not directly address the reliability issue. See *Thompson*, 436 F. App'x at 670.

⁶⁹ Under the standard set out in *Strickland v. Washington*, a defendant claiming ineffective assistance of counsel must demonstrate both that his counsel's performance fell below an objective standard of reasonableness and that he was prejudiced as a result. 466 U.S. 668, 688, 694 (1984). A showing of prejudice entails convincing the reviewing court that the outcome would have been different if not for counsel's errors. *Id.* at 694. Thus, where the petition is considered by the original trial judge who reasons that he or she would

included in the above criminal data set, which would bring the admission rate to six of eight, or 75%.

The civil arson cases present a more complicated picture, not least because the issues are less uniform than in any of the other sets of cases thus far considered. Thus, one can assume that the “task at hand” analysis—assuming courts engage in it—is quite varied across the cases. In addition, there are a number of cases in which the result was mixed or the expert testimony was limited to a significant extent, as described in further detail below.

Of a total of 135 civil cases, fire cause and origin testimony was admitted outright in 98, or 73%, and was completely excluded in 29, or 21%.⁷⁰ In another four cases, courts admitted one or more experts but also excluded one or more.⁷¹ If these are added to each total, the percentages become 76% and 25% respectively. Finally, in four cases the evidence was admitted with a significant limitation that essentially prevented the expert from giving an opinion as to the cause of the fire, thus in effect amounting to exclusion.⁷²

Comparing the admission and exclusion percentages in criminal and civil cases, then, it is apparent that the disparity seen in the

have admitted the expert testimony as reliable even if it had been challenged, the claim will necessarily fail on the second prong of *Strickland*.

⁷⁰ The percentages do not add up to 100% because there are additional cases that do not cleanly fall into one or the other of these categories, as discussed *infra* at notes 72–73 and accompanying text.

⁷¹ See *Gaskin v. Sharp Elecs. Corp.*, No. 2:05-CV-303, 2007 WL 2572397, at *5–10 (N.D. Ind. Aug. 31, 2007) (admitting some of plaintiff’s proffered experts but excluding one); *Hackert v. First Alert, Inc.*, No. 1:03-CV-216, 2005 WL 6021858, at *4 (N.D.N.Y. Nov. 14, 2005) (defendant’s expert excluded; plaintiff’s expert permitted to testify with some limitation); *Pro Serv. Auto., L.L.C. v. Lenan Corp.*, No. 04-587-CV-GAF, 2005 WL 3371054, at *6–10 (W.D. Mo. Dec. 12, 2005) (one of plaintiff’s experts admitted, one excluded); *State Farm Fire & Cas. v. Jarden Corp.*, No. 1:08-cv-1506-SEB-DML, 2010 WL 2541249, at *1–2 (S.D. Ind. June 16, 2010) (one of plaintiff’s experts admitted, one excluded).

⁷² See *Somnis v. Country Mut. Ins. Co.*, 840 F. Supp. 2d 1166, 1172–73 & n.2 (D. Minn. 2012) (expert precluded from relying on “negative corpus” method, therefore could not state opinion that the fire was incendiary); *Amica Mut. Ins. Co. v. Willard*, No. 4:07CV1745DDN, 2009 WL 2982902, at *6 (E.D. Mo. Sept. 14, 2009) (expert precluded from offering opinion that fire was intentionally set); *Dart Indus., Inc. v. Acor*, No. 6:06-cv-1864-Orl-28DAB, 2008 WL 4539480, at *4–6 (M.D. Fla. Oct. 9, 2008) (expert precluded from testifying that “carbon black dust” caused fire); *Hickerson v. Pride Mobility Prods. Corp.*, No. 05-0035 CV W FJG, 2006 WL 290587, at *6 (W.D. Mo. Feb. 7, 2006) (expert precluded from testifying that motorized scooter chair caused fire), *rev’d*, 470 F.3d 1252 (8th Cir. 2006). If these four are considered exclusions, the rate of exclusion becomes 37 of 135, or 27%.

handwriting cases is not evident in the fire cause and origin cases. In the handwriting cases, prosecution evidence was admitted in nearly 90% of the criminal cases, whereas on the civil side it was admitted (or at least not excluded) in fewer than 40% of cases. In contrast, the admission rates for expert testimony in the fire cases hovered close to 75% for both criminal and civil cases. Given the number of cases, these results might very well imply nothing but statistical anomaly and randomness. On the other hand, it might be that there is some explanation for the more uniform treatment of fire expertise under *Daubert*.

Though the two fields of expertise—handwriting identification and fire investigation—have certain parallels, it is worth noting as well the differences both within the fields and in their treatment by courts. Though the changes did not come quickly enough for Cameron Todd Willingham and others who were convicted based on the earlier, flawed assumptions about the “indicators” of arson, fire investigation training and methodology has indeed moved toward a greater reliance on scientific methodology as opposed to intuition and common sense. Similarly, some recent opinions in criminal cases demonstrate a much more nuanced and sophisticated gatekeeping analysis than was seen in the earlier cases.⁷³ Whether this is a trend remains to be seen; in the handwriting cases, what appeared to be a trend toward more serious gatekeeping turned out to be a brief detour leading to a dead end.

One factor that may prove significant, though, is that evidence offered by forensic arson investigators will continue to be relevant even as their methods improve. If testing suggests that fewer fires are the result of arson,⁷⁴ this is not an indictment of the field per se but merely an evolution in its knowledge about how fires behave.⁷⁵ With respect to forensic document examiners, on the

⁷³ See, e.g., *United States v. Smallwood*, No. 5:08-CR-38, 2010 WL 4168823, at *3–9 (W.D. Ky. Oct. 12, 2010) (examining the reliability of fire investigators’ testimony in depth).

⁷⁴ For example, because many of the features of fire scenes that were previously thought to indicate arson, such as certain burn patterns or “alligator” glass, have been shown through testing to be fully consistent with nonincendiary fires.

⁷⁵ However, there does still appear to be significant resistance to change. As recounted in the Arson Report commissioned by The Innocence Project,

other hand, if it were to be shown that they simply cannot reliably do many of the tasks that they claim to do, they would be at much greater risk professionally.⁷⁶ And, indeed, it does seem that the field of forensic document examiners is substantially more resistant to testing and peer review than the field of fire origin and cause.⁷⁷

In addition, falsification is more plausible in the case of arson indicator assumptions. It is not uncommon for investigators or other experts to create experimental fires and then study the resulting artifacts.⁷⁸ In this way, it has now become clear, for example, that fires involving accelerants do not necessarily burn hotter than accidental fires, and that “flashover” conditions in a compartment fire⁷⁹ can easily result in the indicators that previously were believed to signal arson. It therefore becomes more difficult (and easier to refute) for an arson expert to testify, as in the case against Cameron Todd Willingham, that

In 1985, the National Fire Protection Association Standards Council recognized the lack of reliability of fire investigations, and formed the Technical committee on Fire Investigations to prepare a standard document. Unfortunately, the first edition of NFPA 921, Guide for Fire and Explosion Investigations, was not published until shortly after the Willingham fire. Even if it had been published, there is little chance that it would have been accepted. The fire investigation community resisted this document and the principles it espoused for most of the 1990s.

Lentini et al., *supra* note 55, at 39.

⁷⁶ Or at least some of them would. Professor Risinger has pointed out that the results of those competency tests that have been done suggest that the skills of document examiners may be bimodal, with some people doing quite well at certain tasks and others doing quite poorly. Risinger, *supra* note 12, at 493 (observing that the results of two studies on accuracy rates of expert documents examiners tended to be “very uneven, with both very good and very bad performers in the expert group” and that “there is some reason to think that both expert and non-expert performance may be widely spread, or even bi-modal”).

⁷⁷ Risinger recounts this resistance, which includes extreme difficulty in persuading handwriting experts to participate in testing, severe limitations upon the parameters of voluntary proficiency tests run by forensic trade organization, and refusal to disclose data upon which reported results are based. *See id.* at 480–94.

⁷⁸ *E.g.*, JAMES H. SHANLEY, JR., FED. EMERGENCY MGMT. ADMIN., U.S. FIRE ADMIN., REPORT OF THE UNITED STATES FIRE ADMINISTRATION PROGRAM FOR THE STUDY OF FIRE PATTERNS (1997).

⁷⁹ A compartment fire is “a fire that is confined within an enclosure such as in a room or building.” Lentini et al., *supra* note 55, at 8 n.7.

“aluminum melts at 1,200° normal. Wood fire does not exceed 800°. So to me, when aluminum melts, it shows me that it has a lot of intense heat. It reacts to it. That means its temperature is hot. The temperature cannot react. Therefore the only thing that can cause that to react is an accelerant. You know it makes the fire hotter. It’s not normal fire.”⁸⁰

This is not to say that such erroneous and misleading testimony will never be admitted, but it stands to reason that it may become less commonplace.

With respect to handwriting identification, though, such falsification is much more difficult—if not impossible—because it amounts to trying to prove a negative. It is possible to demonstrate that accidental, nonaccelerated fires can burn at the same temperatures as incendiary, accelerated fires or that crazed patterns in glass indicate not that a “fire burned fast and hot” but rather that the glass was subjected to rapid cooling or mechanical breakage. It is not possible to show that an expert’s opinion that a particular person wrote a particular piece of text is inaccurate. The most one might do is to show that its accuracy is unproven.⁸¹

One might wonder whether the 2009 publication of the NAS Report on forensics—which highlights, among other issues, the serious lack of data underlying many of the claims of forensic document examiners⁸²—has had any effect on admissibility of such testimony. The short answer is that so far it has not. Indeed, a recent decision by the District of Columbia Court of Appeals, the District’s equivalent of a state supreme court, suggests that prosecutors and courts may continue to be reluctant to engage in any serious examination of the reliability of the various types of evidence discussed in the NAS Report. In *Gee v. United States*,⁸³ the D.C. Court of Appeals considered the trial court’s rejection of

⁸⁰ *Id.* at 10–11 (quoting testimony of primary prosecution expert witness Manuel Vasquez from the transcript of Willingham’s trial).

⁸¹ Of course, this is precisely the basis for arguments that such “expertise” does not amount to science or that, if it is science, it has not been validated sufficiently to satisfy the *Daubert* standard.

⁸² NAS REPORT, *supra* note 20, at 166–67 & n.98.

⁸³ 54 A.3d 1249 (D.C. 2012).

the defendant's attempt to cross-examine the government's fingerprint expert using excerpts from the NAS Report. In the course of analyzing whether the Report satisfied the requirements of the learned treatise exception to the rule against hearsay,⁸⁴ the court repeatedly referred to the "relevant scientific community" as consisting of forensic fingerprint examiners,⁸⁵ essentially validating a version of the guild test⁸⁶ in this context. The court quoted the prosecutor as follows:

The relevant scientific fingerprint community does not consider the NAS Report a learned treatise. The people on the scientific working group on fingerprints, SWGFAST, do not consider it a learned treatise. The FBI at Quantico does not consider it a learned treatise. The fingerprint unit at Scotland Yard does not consider it a learned treatise. These are the leaders in the field. These are the people that are brought together to issue protocols and standards for those folks who are practicing in the field. And they don't consider it a learned treatise. *What they consider it to be is a policy statement.*⁸⁷

For purposes of the learned treatise exception, the trial court undoubtedly was well within its discretion in declining to judicially notice the NAS Report as authoritative in the field.⁸⁸ But the prosecutor's comment—that the Report is merely a "policy statement," insofar as it notes that several common assumptions underlying latent print identification have not been quantified or validated—bodes ill for any prospect that either prosecutors or courts will be inclined to approach these issues as an

⁸⁴ FED. R. EVID. 803(18) (providing an exception to hearsay exclusion for "a statement contained in a treatise . . . if . . . the publication is established as a reliable authority").

⁸⁵ *Gee*, 54 A.3d at 1264.

⁸⁶ See D. Michael Risinger, *Defining the "Task at Hand": Non-Science Forensic Science After Kumho Tire Co. v. Carmichael*, 57 WASH. & LEE L. REV. 767, 770 n.14 (2000) (explaining "guild" test relative to other standards).

⁸⁷ *Gee*, 54 A.3d at 1263 (emphasis added).

⁸⁸ *Id.* at 1266.

epistemologist would.⁸⁹ Similarly, the trial court's statement that it "had read the report"⁹⁰ and that there were certain parts "which [it] disagree[d] with"⁹¹ evinces the long shadow of precedent and common practice when it comes to government forensics.⁹²

So it's true: there are two *Dauberts*. Many commentators view this as a problem, and there are many suggested solutions. The solutions differ based on the particular commentator's views about practicalities, forensics, judges, juries, experts, crime labs, statistics, and criminal justice. Suggested reforms include creating independent crime labs,⁹³ raising the *Daubert* bar in criminal cases,⁹⁴ lowering the *Daubert* bar in civil cases,⁹⁵

⁸⁹ See generally Edward J. Imwinkelried, *The Epistemological Trend in the Evolution of the Law of Expert Testimony: A Scrutiny at Once Broader, Narrower, and Deeper*, 47 GA. L. REV. 863 (2013).

⁹⁰ *Gee*, 54 A.3d at 1262.

⁹¹ *Id.* at 1263.

⁹² Cf. *Pettus v. United States*, 37 A.3d 213, 214 (D.C. 2012) ("Although appellant . . . makes a spirited attack on the general acceptance of all such 'pattern matching' analysis in the light of a recent National Research Council Committee Report, we hold that forensic handwriting comparison and expert opinions based thereon satisfy the bedrock admissibility standard . . . and may be put before a jury, where remaining issues of reliability may be argued, after cross-examination and any counter-expert testimony, as affecting the weight of the opinions."); *United States v. Llera Plaza (Llera Plaza I)*, 179 F. Supp. 2d 492, 516 (E.D. Pa. 2001) ("Since the court finds that ACE-V does not meet *Daubert's* testing, peer review, and standards criteria, and that information as to ACE V's rate of error is in limbo, the expected conclusion would be that the government should be precluded from presenting any fingerprint testimony. But that conclusion—apparently putting at naught a century of judicial acquiescence in fingerprint identification processes—would be unwarrantably heavy-handed."); *vacated and superseded*, *United States v. Llera Plaza (Llera Plaza II)*, 188 F. Supp. 2d 549, 552–53 (E.D. Pa. 2002) (holding that expert testimony regarding latent fingerprint identification was admissible, following the government's assertion that "its prosecutorial effectiveness, both in the case at bar and in other cases in which fingerprint identification could be expected to play a significant role, would be seriously compromised by the preclusion of opinion testimony at the 'evaluation' stage 'that a particular latent print is in fact the print of a particular person'").

⁹³ E.g., Paul C. Giannelli, *The Abuse of Scientific Evidence in Criminal Cases: The Need for Independent Crime Laboratories*, 4 VA. J. SOC. POL'Y & L. 439, 464–78 (1997) (arguing that independent crime labs would be a "substantial step in the right direction"); see also NAS REPORT, *supra* note 20, at 24 (recommending congressional allocation of funds to state and local jurisdictions for the purpose of "removing all public forensic laboratories and facilities from the administrative control of law enforcement agencies or prosecutors' offices").

⁹⁴ See, e.g., Milich, *supra* note 7, at 421 (noting that "Georgia currently applies *Daubert* and its provisions for judicial review of the reliability of expert testimony in civil cases but not in criminal cases," and that "this distinction is puzzling" since "*Daubert* was designed to improve the overall reliability of scientific and other expert testimony in the courtroom" and

subjecting forensic analysts and labs to error rate and proficiency testing,⁹⁶ trying to incentivize the forensic communities to adopt scientific norms and practices,⁹⁷ and funding defense experts in a more equitable manner.⁹⁸ All of these suggestions are commendable and certainly would improve upon the current and rather depressing state of affairs in our criminal justice system.

On the other hand, some commentators support differing reliability standards for criminal and civil cases. Professor Risinger, for example, has written several times over the years that he “favor[s] context-inflected standards of required reliability under Federal Rule of Evidence 702.”⁹⁹ However—and quite reasonably—Professor Risinger would like to see “the highest standards being imposed on the prosecution in criminal cases.”¹⁰⁰ I do not disagree that this would be the ideal; my position is simply that this is *never going to happen*.¹⁰¹ Therefore, my reasons for favoring distinct standards are different.

“[i]t is unclear why a state would want that improved reliability only in civil, not criminal, cases”).

⁹⁵ See Allan Kanner & Ryan Casey, *Daubert and the Disappearing Jury Trial*, 69 U. PITT. L. REV. 281, 329 (2007) (describing *Daubert* as a “monster” in civil litigation); Jan Beyea & Daniel Berger, *Scientific Misconceptions Among Daubert Gatekeepers: The Need for Reform of Expert Review Procedures*, 64 LAW & CONTEMP. PROBS. 327, 357–60 (2001) (arguing that *Daubert* holds scientists to “unrealistic expectations” and leads to the “exclusion of an extraordinary number of experts in civil cases”); see also Richard L. Cupp, Jr., *Believing in Products Liability: Reflection on Daubert, Doctrinal Evolution, and David Owen’s Products Liability Law*, 40 U.C. DAVIS L. REV. 511, 516 (2006).

⁹⁶ See Beyea & Berger, *supra* note 95, at 360–67 (discussing expert review as a proposal for change); Paul C. Giannelli, *Wrongful Convictions and Forensic Science: The Need to Regulate Crime Labs*, 86 N.C. L. REV. 163, 208–34 (2007) (making specific recommendations for how to better regulate crime labs).

⁹⁷ See Jennifer L. Mnookin et al., *The Need for a Research Culture in the Forensic Sciences*, 58 UCLA L. REV. 725, 679 (2011) (suggesting that a commitment to “research norms” by forensic science journals would improve the “research culture of the pattern identification services”).

⁹⁸ See Giannelli, *supra* note 93, at 473–74 (arguing that the “always-underfunded criminal defense services are in a state of crisis”).

⁹⁹ Risinger, *supra* note 12, at 558–59.

¹⁰⁰ *Id.* at 559; see also D. Michael Risinger, *The Irrelevance, and Central Relevance, of the Boundary Between Science and Non-Science in the Evaluation of Expert Witness Reliability*, 52 VILL. L. REV. 679, 719–22 (2007); D. Michael Risinger, *A Functional Taxonomy of Expertise*, in 1 MODERN SCIENTIFIC EVIDENCE ch. 2, § 2:16 (David L. Faigman et al. eds., 2012–2013 ed.) (suggesting “the higher the standard of proof applicable to the issue upon which the expertise is offered, the higher the required threshold dependability should be”).

¹⁰¹ As for why it won’t happen—or at least why it thus far has not happened—see Joseph

III. WE KNOW EVERYTHING, WE KNOW NOTHING: A PLEA FOR EPISTEMIC HUMILITY

It has been suggested by one of my eminent co-panelists that “we are moving toward a fundamentally epistemological approach to determining the admissibility of expert testimony.”¹⁰² Certainly this is the import of *Daubert* and its progeny, taken at face value. And it may indeed be the direction in which courts have moved and are moving on the civil side of the docket. As discussed above, however, courts in criminal cases have mostly come to this fork in the road and then stopped dead in their tracks.¹⁰³

What Professor Imwinkelried and others advocate is a rigorous application of the *Daubert* standard to all “knowledge claims by any expert.”¹⁰⁴ As he notes, many *Frye* jurisdictions long distinguished between “science” and “nonscience” expertise, or between hard and soft science, exempting the latter from serious scrutiny.¹⁰⁵ Even following the Supreme Court’s decision in *Daubert*, some federal courts continued to draw this line between science and nonscience and to subject the former to the new, more rigorous¹⁰⁶ *Daubert* test while giving the latter a relative pass.¹⁰⁷

Sanders, “Utterly Ineffective”: Do Courts Have a Role in Improving the Quality of Forensic Expert Testimony?, 38 *FORDHAM URB. L.J.* 547, 555–62 (2010). Professor Sanders summarizes various explanations for judicial laxity in policing the reliability of forensic science evidence in criminal cases, including that judges are simply “grandfathering in the traditional forensic sciences,” that “they believe most criminal defendants to be guilty,” and that they believe the evidence to be reliable. *Id.* at 555–57. Professor Sanders argues, instead, that courts’ reluctance to apply *Daubert* more vigorously to most forensic science evidence is primarily the result of “the judiciary’s contextual approach to knowledge” combined with “the limited ability of science to provide causal information about a particular case.” *See id.* at 557–58. He surmises that judges “understand the evidence is weak but admit it anyway,” *id.* at 561, because it is the best that is currently available and they believe it is better than nothing.

¹⁰² Imwinkelried, *supra* note 89, at 870.

¹⁰³ Or, one might say, they have mainly paused at the fork and then turned back around. As the court said in *United States v. Cline*, with regard to fingerprint identification evidence, “[w]hile certainly exhaustive and informed, the analysis of the *Daubert* factors used in *Llera Plaza* [I] does not persuade this court to depart from the well-traveled path.” 188 F. Supp. 2d 1287, 1294 (D. Kan. 2002) (citing *Llera Plaza I*, 179 F. Supp. 2d 492, 500–01 (E.D. Pa. 2001)).

¹⁰⁴ Imwinkelried, *supra* note 89, at 870.

¹⁰⁵ *Id.* at 871–73.

¹⁰⁶ It should be noted that there was significant debate immediately following the Supreme Court’s *Daubert* decision as to whether, in fact, the new standard would be more

This was certainly a fair reading of *Daubert*.¹⁰⁸ Much of the reasoning in Justice Blackmun's majority opinion focused on the word "science" in Federal Rule of Evidence 702,¹⁰⁹ and the *Daubert*

rigorous than *Frye*. See Seaman, *Triangulating Expert Testimony*, *supra* note 29, at 861 & n.179. As Professor Carlson notes, however, at least "by the time of its 2000 *Weisgram* decision, the Court had adopted a very different tone, alluding to 'the exacting standards of reliability' mandated by Rule 702." Ronald L. Carlson, *The Curious Case of Differing Literary Emphases: The Contrast Between the Use of Scientific Publications at Pretrial Daubert Hearings and at Trial*, 47 GA. L. REV. 837, 843 (2013) (quoting *Weisgram v. Marley Co.*, 528 U.S. 440, 455 (2000)).

¹⁰⁷ Most notably for purposes of this Essay, this was the approach taken by Judge McKenna in *United States v. Starzeczyzel*, which has been called "the original handwriting expertise reliability case of the modern era." Risinger, *supra* note 12, at 495. More generally, prior to the Supreme Court's clarification of the issue in 1999, courts in the Second, Ninth, and Tenth Circuits continued to apply pre-*Daubert* standards to nonscientific evidence. See, e.g., *Compton v. Subaru of Am., Inc.*, 82 F.3d 1513, 1518–19 (10th Cir. 1996), *overruled by Kumho Tire Co. v. Carmichael*, 526 U.S. 137 (1999) (declining to apply *Daubert* test to automotive engineer's expert testimony); *United States v. Rice*, 52 F.3d 843, 847 (10th Cir. 1995) (applying traditional Rule 702 analysis to tax attorney's expert testimony); *United States v. Muldrow*, 19 F.3d 1332, 1338 (10th Cir. 1994) (applying traditional Rule 702 analysis to police officer's expert testimony about drug trafficking); *Thomas v. Newton Int'l Enters.*, 42 F.3d 1266, 1270 n.3 (9th Cir. 1994) (declining to apply *Daubert* to worker's expert testimony based on experience as a longshore worker); *Iacobelli Constr. v. Cnty. of Monroe*, 32 F.3d 19, 24–25 (2d Cir. 1994) (declining to apply *Daubert* test to affidavits of geotechnical expert); *Tamarin v. Adam Caterers, Inc.*, 13 F.3d 51, 53 (2d Cir. 1993) (declining to apply *Daubert* test to expert accountant's testimony). Critiquing this approach, Professor Michael Saks colorfully described the reasoning thus: "Forensic handwriting examination flunks the *Daubert* test. Because it flunks *Daubert*, it is not science. Because it is not science, it need not pass the *Daubert* test. A much looser 'test' of soundness applies. Under that test, it is admissible." Michael J. Saks, *The Aftermath of Daubert: An Evolving Jurisprudence of Expert Evidence*, 40 JURIMETRICS 229, 237 (2000).

¹⁰⁸ In addition to those lower courts that interpreted *Daubert* this way, evidence scholars also debated whether the *Daubert* test—or at least the factors as outlined in *Daubert* itself—was confined to scientific expertise or applicable to all expertise. The weight of scholarly opinion was in favor of applying some version of *Daubert* gatekeeping broadly to asserted expertise. See, e.g., David L. Faigman, *Making the Law Safe for Science: A Proposed Rule for the Admission of Expert Testimony*, 35 WASHBURN L.J. 401, 422–27 (1996) (discussing the applicability of *Daubert* beyond scientific expert testimony).

¹⁰⁹ At the time *Daubert* was decided, Rule 702 provided: "If scientific, technical, or other specialized knowledge will assist the trier of fact to understand the evidence or to determine a fact in issue, a witness qualified as an expert by knowledge, skill, experience, training, or education, may testify thereto in the form of an opinion or otherwise." Justice Blackmun's majority opinion reasons that "[t]he adjective 'scientific' implies a grounding in the methods and procedures of science." 509 U.S. 579, 590 (1993). Similarly, the opinion states that "in order to qualify as 'scientific knowledge,' an inference or assertion must be derived by the scientific method," and "the requirement [in Rule 702] that an expert's testimony pertain to 'scientific knowledge' establishes a standard of evidentiary reliability." *Id.*

factors¹¹⁰ were explicitly drawn from descriptions of the scientific method by philosopher of science Karl Popper and others.¹¹¹

Under the *Daubert* trilogy—and most explicitly as clarified by *Kumho Tire*—all expert opinion testimony offered in federal courts is subject to *Daubert* gatekeeping and scrutiny.¹¹² The so-called “opinion rule”¹¹³ has been put to rest in the federal courts and in those jurisdictions that follow the federal rules in this respect.

A small number of jurisdictions, however, continue to follow some version of the opinion rule. Georgia happens to be one of these, along with California.¹¹⁴ Even with Georgia’s adoption of the Federal Rules of Evidence, the new Georgia evidence code retains the distinction between scientific and nonscientific evidence,¹¹⁵ which is roughly analogous to California’s distinction between scientific and “experiential” expertise. Even more unusual, Georgia appears to be the only jurisdiction that maintains different standards of admissibility, even in the case of *scientific* evidence, in criminal and civil cases.¹¹⁶ In the balance of

¹¹⁰ See *id.* at 592–95 (outlining an illustrative, but not exhaustive, list of factors including: “whether the reasoning or methodology underlying the testimony is scientifically valid,” “whether the theory or technique has been subjected to peer review and publication,” the “known or potential rate of error,” and “general acceptance” within the scientific community).

¹¹¹ See *id.* at 593 (citing KARL POPPER, CONJECTURES AND REFUTATIONS: THE GROWTH OF SCIENTIFIC KNOWLEDGE 37 (5th ed. 1989)). One of the foremost scholars of law, philosophy, and science described the Court’s approach thus:

Apparently equating the question of whether expert testimony is reliable with the question of whether it is genuinely scientific, taking for granted that there is some scientific ‘methodology’ which, faithfully followed, guarantees reliable results, and casting about for a philosophy of science to fit this demanding bill, the *Daubert* Court settled on an unstable amalgam of Popper’s and Hempel’s very different approaches

Susan Haack, *Disentangling Daubert: An Epistemological Study in Theory and Practice*, 5 J. PHIL., SCI. & L. 1, 2 (2005).

¹¹² *Kumho Tire*, 526 U.S. at 147.

¹¹³ This term refers to the practice of distinguishing between scientific and nonscientific expert opinion and subjecting the latter to less rigorous scrutiny. Professor Faigman argues that the federal courts continue to apply this distinction in practice, though it has been rejected by the explicit doctrine. See Faigman, *supra* note 3, at 706–08. I agree with this characterization, but rather than critiquing the practice, I suggest making it explicit.

¹¹⁴ CAL. EVIDENCE CODE § 801 (West 2012).

¹¹⁵ O.C.G.A. § 24-7-702(b) (2013); see also *supra* notes 6–13 and accompanying text.

¹¹⁶ O.C.G.A. §§ 24-7-702, 24-7-707 (2013). In *Mason v. Home Depot U.S.A., Inc.*, the Georgia Supreme Court held, among other things, that this distinction did not violate the Equal Protection Clause of either the Georgia or United States Constitutions because “the

this Article, I shall try to make the (difficult) case that this may be a good thing, or, at the very least, that it may not be a terrible thing.

But first I would like to indulge in a short diversion about science, knowledge, and truth. A few years ago, *The New Yorker* published an article by science writer Jonah Lehrer about the so-called “decline effect.”¹¹⁷ The article is entitled, “The Truth Wears Off” and is subtitled, “Is there something wrong with the scientific method?”¹¹⁸ In it, Lehrer describes a widespread and rather unsettling phenomenon whereby scientific findings based on well-designed, peer-reviewed, theoretically supported studies steadily and systematically erode over time when attempts are made to replicate them. According to the article, this phenomenon has been noticed in biomedical and pharmaceutical research, in cognitive psychology, and in evolutionary biology, among other fields.¹¹⁹ Various explanations have been offered—regression to the mean over time, publication bias, selective reporting of data—but none of these seems fully to account for the magnitude of the effect.

Lehrer concludes the article with these thoughts:

Although many scientific ideas generate conflicting results and suffer from falling effect sizes, they continue to get cited in the textbooks and drive standard medical practice. Why? Because these ideas seem true. Because they make sense. Because we can't bear to let them go. And this is why the decline

parties to civil cases are not similarly situated to those engaged in criminal prosecutions.” 658 S.E.2d 603, 607 (Ga. 2008). In a vigorous dissent, Justice Hunstein argued that “[t]here is no rational reason to subject evidence affecting an individual’s life and liberty to less rigorous standards of admissibility than that applied to evidence affecting mere property.” *Id.* at 613 (Hunstein, P.J., dissenting).

¹¹⁷ See Jonah Lehrer, *The Truth Wears Off: Is There Something Wrong with the Scientific Method?*, NEW YORKER, Dec. 13, 2010, at 52, available at http://www.newyorker.com/reporting/2010/12/13/101213fa_fact_lehrer?printable=true¤tPage=all.

¹¹⁸ In what has to be the height of irony, Jonah Lehrer has since been accused of plagiarism, fabricating quotations, and other journalistic misconduct and has been fired from his position as a staff writer at *The New Yorker* and as a blogger at Slate.com. Nonetheless, his observations in this particular article—and the underlying research that he references—have not been implicated in the scandal surrounding some of his other work.

¹¹⁹ Lehrer, *supra* note 117, at 54–56.

effect is so troubling. Not because it reveals the human fallibility of science, in which data are tweaked and beliefs shape perceptions. (Such shortcomings aren't surprising, at least for scientists.) And not because it reveals that many of our most exciting theories are fleeting fads and will soon be rejected. (That idea has been around since Thomas Kuhn.) The decline effect is troubling because it reminds us how difficult it is to prove anything. We like to pretend that our experiments define the truth for us. But that's often not the case. Just because an idea is true doesn't mean it can be proved. And just because an idea can be proved doesn't mean it's true.¹²⁰

This description of the stickiness of earlier findings that “seem true” and “make sense” is also a very apt description of how courts behave when faced with challenges to long-accepted forensic evidence such as latent fingerprint and handwriting identification testimony. The *Daubert* framework encourages trial judges to continue to cite settled precedent just as the scientific establishment continues to cite major studies even after they have been undermined by subsequent research and failures to replicate the initial results. Because of the high level of discretion and deference granted to trial courts under Rule 702,¹²¹ there is little incentive for judges to exclude forensic evidence offered against a criminal defendant and compelling reasons for admitting it. Prosecutors often claim that they will be seriously prejudiced if they cannot use a certain type of evidence, and judges are understandably reluctant to hamstring prosecutions by excluding crucial evidence.¹²² Because it is exceedingly unlikely that an appellate court will overturn the trial court's decision, trial judges are quite unlikely to behave as epistemologists under these circumstances. Rather, they are likely to behave like judges and

¹²⁰ *Id.* at 57.

¹²¹ *Gen. Elec. Co. v. Joiner*, 522 U.S. 136, 141–42 (1997).

¹²² For example, consider the Supreme Court's backtracking in the wake of *Crawford v. Washington*, 541 U.S. 36 (2004), and especially its recent decision in *Williams v. Illinois*, 132 S. Ct. 2221 (2012). *Llera Plaza II*, 188 F. Supp. 2d 549 (E.D. Pa. 2002), is also a good example of this phenomenon.

simply cite longstanding precedent to permit expert testimony, at least where the forensic sciences are concerned.

In prior work, I have focused on the role of the jury not only as factfinder but also—perhaps primarily—as a check on government power, a bulwark between the individual and the state machinery of criminal prosecution.¹²³ I've also argued that *Daubert*, by interposing the judge as a gatekeeper of expert witness testimony, is in some tension with Constitutional principles of confrontation and the right to have a jury determine the facts of one's case.¹²⁴ Under a separation of powers, checking view of the jury, both the court and the prosecution are instrumentalities of the state—the judge is not a neutral umpire but is a state actor. This was the view of the Founders, and this theme is prevalent in many recent Supreme Court opinions on the Sixth Amendment right to confrontation and the right to trial by jury.¹²⁵

It might appear at first glance that this view of the judge, jury, prosecution, and defendant argues in favor of the current practice under *Daubert* of extreme leniency in criminal cases when it comes to forensic evidence. After all, if the judge's role as gatekeeper is in tension with constitutional principles, then the more the judge allows through the gate, the more the jury will hear and the better it can fulfill its constitutional role. If it is correct to say that *Daubert* evidences distrust of both juries and of experts and that *Crawford*, *Apprendi*, and other cases, in contrast, reflect distrust of government (including courts) but a strong trust in juries, then it would seem proper for the courts to be more permissive of evidence in criminal cases when it comes to all witnesses.

But the problem with the way that *Daubert* is currently applied in most criminal cases, at least with respect to some of the more questionable types of forensic evidence, is that it offers a false sense of security. Admitting such evidence under *Daubert* is akin to putting up a flimsy guardrail—it encourages people to take risks they would not take if the danger were clear. While the evidence professoriat and the NAS Report question the scientific

¹²³ See generally Seaman, *Triangulating Expert Testimony*, *supra* note 29; Seaman, *Black Boxes*, *supra* note 29.

¹²⁴ See generally Seaman, *Triangulating Testimonial Hearsay*, *supra* note 29.

¹²⁵ See *id.* at 864–69 (discussing distrust of government as basis for trial by jury).

foundations of many long-settled forensic techniques, most courts continue to admit them and to insist that they satisfy the *Daubert* test. This is dishonest and misleading.

Prior to *Kumho Tire's* holding that courts are required to stand guard against unreliable expertise of all stripes,¹²⁶ several lower federal courts had found that certain forensic techniques were not "science" and therefore were not subject to the *Daubert* standard.¹²⁷ Essentially, these courts were applying a version of the "opinion rule" to hold evidence that they deemed "nonscientific" to a lesser standard of admissibility. At the same time, however, there was a move to inform juries that the court did not consider such evidence to be science because it had yet to be subjected to the rigors of validation, peer review, and the like.¹²⁸ In this way, courts attempted to impart some degree of transparency to the process and to permit juries to consider the evidence for what it was worth. Though this route was ultimately foreclosed by the Supreme Court in federal cases, it remains available to the states. Insofar as courts have largely failed, in many criminal justice contexts, to apply the heightened standard of admissibility ostensibly required by their rules, there is little reason to allow them to pretend otherwise.

If the Georgia courts are to apply this rule in a principled way, they should of course permit defense opinion under the same relatively lenient standard in criminal cases. This kind of uniformity of application would permit juries to hear the limitations of the forensic techniques laid out before them and would potentially spotlight Georgia as a shining star, rather than a whipping boy, in the development of rules regarding expert testimony in criminal cases. The Georgia Supreme Court might even take as its cue the approaches of the New Jersey and Oregon Supreme Courts as they have tackled the equally thorny issues

¹²⁶ *Kumho Tire Co. v. Carmichael*, 526 U.S. 137, 141 (1998).

¹²⁷ *See, e.g., United States v. Jones*, 107 F.3d 1147, 1157 (6th Cir. 1997) (admitting handwriting evidence); *Iacobelli Const., Inc. v. Cnty. of Monroe*, 32 F.3d 19, 25 (2d Cir. 1994) (admitting geotechnical data and reports); *United States v. Starzecpyzel*, 880 F. Supp. 1027, 1029 (S.D.N.Y. 1995) (admitting forensic document evidence).

¹²⁸ *See Starzecpyzel*, 880 F. Supp. at 1049 ("[T]he jury will be instructed, in advance of any forensic document testimony, that [the forensic document examiners] offer practical, rather than scientific expertise.").

surrounding eyewitness identification testimony. Grappling with what all informed participants and observers agree are very difficult evidentiary problems, the ideal of the states as little laboratories has some appeal, particularly in light of the generally disappointing results to date in the federal courts.

I therefore propose that other states follow Georgia's lead and at least reconsider the benefits of the much maligned "opinion rule," which allows courts to treat certain types of expertise as "not science"—or at least "not yet demonstrated to be based on scientific methods and principles"—and therefore not (yet) subject to *Daubert*. Furthermore, the states should look to Georgia's example and apply a different, less stringent reliability standard in criminal cases than in civil cases. Though this suggestion may seem perverse,¹²⁹ if implemented carefully, it has the benefit of being more transparent and honest. In addition, "calling out" forensic evidence as "nonscience" might encourage these fields to adopt a more scientific methodology, as several commentators have urged is necessary.¹³⁰ In other incarnations in other jurisdictions, we might see different results. Indeed, a scientific approach to the problem of the admissibility of scientific evidence itself could ultimately point the way out of the maze and onto a road toward reliability and consistency in the treatment of forensic expertise.

IV. CONCLUSION

The bright side of the opinion rule is that it both trusts jurors to evaluate the evidence in the case and allows that evidence to be presented more transparently. Properly applied—and this is obviously crucial—a non-*Daubert* standard would allow defendants to reveal the shortcomings of the forensic evidence offered against

¹²⁹ The opinion rule remains in effect in some states, notably California, and has been criticized as perverse because it allows the most questionable expert opinions to be admitted with little scrutiny whereas more reliable scientific evidence gets the full *Daubert* treatment. See Faigman, *supra* note 3, at 708 ("[T]he California rule means that expert opinion with little or no scientific basis is readily admitted . . . and [it] erects substantial barriers to the introduction of science-based actuarial techniques.").

¹³⁰ See Mnookin et al., *supra* note 97, at 778–79 (arguing for more transparency and better research methods from forensic sciences).

them. It would allow defendants to tell jurors that the evidence is not scientific and to present evidence in support of that claim. This is the direction in which some lower federal courts seemed to be heading prior to the Supreme Court's decision in *Kumho Tire*.¹³¹

Judges probably do not appreciate trying to shoehorn certain forensic testimony into a *Daubert* mold, but it seems pretty clear that most judges are not about to exclude the evidence. A non-*Daubert* approach would allow courts to be more honest about the reliability of such evidence while still admitting it. Coupled with some needed reforms, such as requiring government labs to share their data fully with defense counsel and scholars, tossing *Daubert* out might end up encouraging these fields to move toward a more rigorous scientific model.

The Lehrer article is disturbing¹³²—it reveals the extent to which we rely on questionable data, for example, in making treatment decisions for serious medical conditions. It shows that even the most seemingly reliable scientific evidence, based on gold-standard randomized double-blind studies, may not be as solid as we think. In the criminal context, the flaws underlying much of what passes for reliable scientific evidence carry even more serious consequences for liberty and even life, as the 300 (and counting) DNA exonerations have revealed.¹³³ It may be that we have to make do with the evidence that is available and do the best we can; given these constraints, it might be better to dispense with the pretense of *Daubert* in criminal cases and to encourage more transparency so that juries are clear what they are getting when they consider all of the evidence.

¹³¹ See *supra* note 127.

¹³² See generally Lehrer, *supra* note 117.

¹³³ See THE INNOCENCE PROJECT, <http://www.innocenceproject.org> (last visited Jan. 19, 2013) (listing DNA exonerations to date); see also BRANDON L. GARRETT, CONVICTING THE INNOCENT: WHERE CRIMINAL PROSECUTIONS GO WRONG 84–117 (2011) (discussing failures of forensic sciences).