Tulsa Law Review

Volume 45
Issue 1 *Exhibiting Culture: Museums and Indians*

Fall 2009

Don't Shoot the Messenger by One of the Messengers: A Response to Merlino et al.

Simon A. Cole

DAUBERT REVISITED

DON'T SHOOT THE MESSENGER BY ONE OF THE MESSENGERS: A RESPONSE TO MERLINO ET AL.*

Simon A. Cole**

Oh, I see! It's my fault, is it? Oh, of course, there I was thinking it was your fault because you've been left in charge, or Manuel's fault for not waking you, when all the time it was my fault! Oh, it's so obvious now, I've seen the light!

—John Cleese, *Fawlty Towers*[1]

I. INTRODUCTION

In its Winter 2007 issue, the *Tulsa Law Review* published a Symposium Issue on Daubert, *Innocence, and the Future of Forensic Science*.[2] The Issue might be described as a retrospective on the seminal 1993 United States Supreme Court decision *Daubert v. Merrell Dow Pharms.*,[3] now widely described as the most important expert evidence decision ever written by the Supreme Court, and its application to forensic science.

---

* My title is a reference to the second part of the title of Professor Risinger's contribution to the Symposium Issue of the *Tulsa Law Review* to which this article is a reply. *See* 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). On a completely irrelevant note: the first part of Professor Risinger's title refers to Robert Graves's memoir *Goodbye to All That*, whose discussion of the experience of being gassed in the First World War indirectly inspired my undergraduate thesis on German preparations for chemical warfare between the two world wars.

** Associate Professor of Criminology, Law & Society, University of California, Irvine; Ph.D. (Science & Technology Studies), Cornell University; A.B., Princeton University. This project was funded in part by the National Science Foundation (Award # SES-0347305). Any opinions, findings, conclusions or recommendations expressed in this material are those of the author and do not necessarily reflect the views of the National Science Foundation. I am grateful to Gary Edmond, Craig Cooley, William C. Thompson, Jane C. Moriarty, Tamara Piety, and especially Michael Risinger for reviewing and commenting on a draft of this reply. I am grateful to the student editors for allowing me to write and publish this reply.

1. *Fawlty Towers*, "The Builders" (BBC Sept. 26, 1975) (TV series). This is considered by many among the funniest lines from the BBC television comedy series *Fawlty Towers*, a mini-series about a cantankerous innkeeper, his formidable wife, and the Barcelonés bellhop, Manuel, which starred John Cleese shortly after completing his stint in *Monty Python's Flying Circus* and is considered by many to be among the funniest British television shows ever made. In 2000, it placed first on the British Film Institute's list of the 100 "all-time top British television programmes." *See* British Film Institute, *The BFI TV 100*, http://www.bfi.org.uk/features/tv/100/index.html (last updated Sept. 4, 2006).

2. 43 Tulsa L. Rev. 229 (2007).

3. *Daubert v. Merrell Dow Pharms., Inc.*, 509 U.S. 579 (1993).

Broadly speaking, *Daubert*, and the symposium essays centered around *Daubert,* might be said to concern "the problem of expertise."[4] Given that courts have long allowed expert witnesses to testify[5]—and given the increasing use of such experts—how are courts to evaluate the testimony of proffered "experts"? Ought anyone who claims the mantle of expertise be permitted to testify in that guise? Or, should courts police claims to the title of "expert" by permitting only those experts deemed legitimate to testify? American courts have long tended toward the latter view; *Daubert* made this commitment (in the federal courts and in the many jurisdictions that subsequently adopted *Daubert* or *Daubert*-like rules) explicit. But, this preference only generates another philosophical dilemma: how are courts supposed to adjudicate claims to expertise when many, if not all, of those claims by their very nature are so technical that legally-trained judges cannot reasonably be expected to be competent to sit in judgment upon them?[6] In other words, the law faces a specific instance of the question asked by the philosophical field known as "epistemology": how does one certify knowledge as legitimate?

To make the problem more concrete: if an individual—or, more likely, a group of individuals—claims that they have a forensic assay that performs in a certain way (e.g. identifies the source of trace, with some stated degree of discrimination and some stated degree of accuracy), how are judges to evaluate this claim, given that they are not themselves competent to perform the assay, or even in some cases to understand it? One common approach relies on what is often—somewhat misleadingly—called "the scientific method."[7] Can we muster evidence that supports the claims made concerning this assay? For example, one might design experiments to measure the performance of the assay on a set of controlled test materials.

Of course, we would not realistically expect a judge to perform such an experiment herself. A judge would hardly be the wisest person to select to conduct such an experiment; the task ought to be delegated to someone competent to do so, probably someone generally called a "scientist." However, even more than this, a now familiar insight of the sociology of scientific knowledge holds that *even scientists themselves* regularly evaluate knowledge claims without performing actual experiments themselves. Were each scientist required to directly and independently experimentally validate each knowledge claim that they accept in their work, the enterprise of science would grind to a halt. Instead, scientists accept most knowledge claims through a social process of trust, in which certain markers—prestige of journals, university affiliations, accolades of peers, and so on—are taken as strong evidence of the veracity of the claims made.[8] That this trust-based system is sometimes abused—as when premature stem cell findings survive the peer review process of the prestigious scientific journal *Science* as well as numerous

4. *See generally Expertise in Regulation and Law* 1–2 (Gary Edmond ed., Ashgate Publg. Ltd. 2004).

5. *See* Tal Golan, *Laws of Men and Laws of Nature: The History of Scientific Expert Testimony in England and America* (Harv. U. Press 2004).

6. *See generally* Scott Brewer, *Scientific Expert Testimony and Intellectual Due Process*, 107 Yale L.J. 1535 (1998).

7. *See* Susan Haack, *Defending Science–Within Reason: Between Scientism and Cynicism* (Prometheus Books 2003).

8. Steven Shapin, *A Social History of Truth: Civility and Science in the Seventeenth-Century England* 23 (U. Chi. Press 1994).

other instances[9]—only serves to highlight the continuing assumption that the system works well most of the time.[10] In other words, we have, in part, what has been called a "deference model" for trusting scientific claims: scientists believe claims because they defer to the judgment of trusted peers whose expertise on the specific issue they deem greater than their own (i.e. the peer reviewers of the article). Judicial evaluation of expert evidence, it has been pointed out, also requires such a deference model and perhaps even a stronger form than that found in science itself, and this notion may be found in both the *Daubert* and the *Frye*[11] regimes.[12]

It has not been lost on the philosophically minded that such a process necessarily entails an infinite regress: the judge is asked to evaluate a self-professed claim to expert knowledge, in large part, by deferring to another self-professed group of experts. But how is the judge to know whether the experts to whom she defers are themselves truly as expert as they say? To determine that, the judge would need to consult yet another group of "meta-experts"[13] and so on *ad infinitum.*[14] To this, the sociologically minded have replied that, while the regress is true in principle, in practice the institution we know as "science" continues to function by breaking the regress all the time. The regress is broken simply by treating conventional markers of expertise—prestigious journals, university affiliations, advanced degrees, and so on—as *de facto* indicators of genuine expertise. While such a system is hardly failsafe, it is argued, it works well enough to allow science to generally make progress.

All of this is to say that knowing when to trust scientific knowledge claims is a difficult problem that, in principle, can, one hopes,[15] be addressed by reference to experiment and data, but that in practice, even within science itself, is often based in larger part on a social process of trust. This problem, it might be said, underlies the contributions to the Symposium Issue, which wrestle with the problem of how judges should evaluate *forensic* scientific knowledge and—in the case of most of the contributions—why judges have thus far done such a poor job of it.

It will not be lost on the reflexively minded that the Symposium Issue itself is subject to the same epistemological problem that forms its subject of inquiry.[16] That is to say: the Symposium contributors themselves claim expertise on the topic of the use of expert evidence in the American legal system. And yet, a consumer of their expertise (i.e. a reader of the Symposium Issue) might reasonably ask why she should trust *their* knowledge claims. It turns out, of course, that the Symposium Issue's claims to legitimate knowledge are founded precisely on the same kind of social system of trust

9. *See* Jennifer Couzin, *Stem Cells . . . And How the Problems Eluded Peer Reviewers and Editors*, 311 Sci. 23 (Jan. 6, 2006).

10. And, indeed, the reason sanctions against abuses of the trust of the scientific communication process are so severe is doubtless the awareness that the system is almost entirely self-policing.

11. *Frye v. U.S.*, 293 F. 1013 (D.C. Cir. 1923).

12. Brewer, *supra* n. 6, at 1627.

13. *Id.*

14. *Id.*

15. It should be noted that not all, perhaps not even most, scientific knowledge claims are so simple that they can be straightforwardly evaluated by experimental data.

16. "Reflexivity" is a term used by sociologists of science to denote the process of applying the same sociological scrutiny to one's own knowledge claims that one applies to the knowledge claims of others. *See* Malcolm Ashmore, *The Reflexive Thesis: Wrighting Sociology of Scientific Knowledge* (U. Chi. Press 1989).

that underlies the scientific knowledge claims that form the Issue's subject: a reader would tend to view the claims made in the Symposium Issue as legitimate simply because they appear in a credentialed source—a scholarly journal—rather than a less credentialed source, such as a blog or a vanity press.

Although, in principle, it would perhaps be desirable to independently evaluate all of the knowledge claims made in the Symposium Issue, in practice to do so would bring the knowledge production process to a halt. And so, readers invest a certain amount of trust in the mere fact of being published in a journal. As noted above, there is nothing especially wrong with this system (except to an excessively rigorous epistemologist), and it is the same system used by those scholarly disciplines deemed most trustworthy and prestigious in modern society, the natural sciences. However, it was also noted above that the system has obvious potential deficiencies.

Recently, legal scholars have become increasingly candid in noting that, even beyond the inescapable deficiencies that dog any deferential knowledge evaluation system, *legal* scholarship suffers from additional deficiencies not shared by other disciplines.[17] The now familiar litany of criticisms include the fact that most law journals, including the most prestigious ones, are not peer reviewed;[18] that they are edited by law students lower in the legal hierarchy than the authors of the articles; that these students are unqualified to evaluate much of the scholarship they receive and therefore are suspected of resorting to "crude credentialism" (which might be distinguished from the sort of "reasonable credentialism" that we described above);[19] that the system of multiple submissions allows virtually all scholarship to be published somewhere; and so on.[20]

It is not my intention to engage in that complicated debate here. I raise these prefatory issues merely to make the point that a reader would ordinarily approach the Symposium Issue from a perspective of credulity. Because of the imprimatur of being in a scholarly journal, the reader would expect that the articles contained in the issue include trustworthy expert knowledge about the problem of legal evaluation of expert evidence. To be sure, the reader might find herself disagreeing with, or even doubting, some of the claims made, but, nonetheless, these claims would—and probably should—enjoy some *a priori* advantage over, say for example, claims made in an anonymously posted Internet "blog."

And, yet upon inspection, this assumption becomes undermined because the Symposium Issue turns out to be internally contradictory. Specifically, the issue is internally inconsistent with regard to one hot-button issue discussed by many of the contributors: the validity of latent print (fingerprint) identification. To illustrate, one contribution stated, "as far as courts are concerned, there is *no* accuracy data for latent print source attributions. In other words, there is *no* evidence, of the sort that practitioners of [Evidence-Based Medicine] would consider 'evidence,' as to the

17. Elizabeth Chambliss, *When Do Facts Persuade? Some Thoughts on the Market for "Empirical Legal Studies"*, 71 L. & Contemp. Probs. 17 (2008).

18. Lee Epstein & Gary King, *The Rules of Inference*, 69 U. Chi. L. Rev. 1, 125 (2002).

19. *See* James Lindgren, *An Author's Manifesto*, 61 U. Chi. L. Rev. 527, 533 (1994).

20. *Id.* at 535.

accuracy of this form of evidence."[21] Other articles in the Symposium Issue would appear to echo this judgment. One approvingly quotes the following statement: fingerprinting's "reliability is unverified either by statistical models or fingerprint variation or by consistent data on error rates."[22] Another says "the science on which many identification techniques, such as latent fingerprint[s] . . . are based, is at best of questionable validity."[23] A third bluntly calls fingerprint evidence "lousy science."[24] A fourth, almost as bluntly, calls it "the reigning champion of the non-science forensic sciences . . . ."[25] One article, however, takes a different view. Merlino et al. appear to contradict all of the above assertions when they speak of "the methodology of latent print examination (i.e. ACE-V), its reliability, [and] the methods used to test its reliability . . . ."[26]

In short, the contributors to the Symposium Issue appear to disagree about the validity of latent print identification with a clear majority holding that latent print identification has not been validated. Thus far, there is nothing extraordinary going on; it would appear that we have what is known as a "scholarly disagreement." There is certainly nothing unusual about a symposium issue on a particular topic containing different opinions, and even vigorous debate, on that topic. Consumers of scholarship are certainly accustomed to such circumstances and well equipped to form their own opinions, perhaps based on their own preconceived preferences, but preferably based on the quality of argumentation and evidence put forward by the various debaters.

But the internal contradiction in the Symposium Issue is more than a routine scholarly disagreement. In addition to critiquing the *opinions* of the other contributors, Merlino et al. attack the qualifications to comment of one of the other contributors. They do not merely critique what that contributor says, they critique that contributor's authority to say it. Specifically, they state that one of the authors who argues that

---

21. Simon A. Cole, *Toward Evidence-Based Evidence: Supporting Forensic Knowledge Claims in the Post-*Daubert *Era*, 43 Tulsa L. Rev. 263, 277 (2007).

22. Craig M. Cooley & Gabriel S. Oberfield, *Increasing Forensic Evidence's Reliability and Minimizing Wrongful Convictions: Applying* Daubert *Isn't the Only Problem*, 43 Tulsa L. Rev. 285, 289 n. 37 (2007) (quoting Donald Kennedy, *Forensic Science: Oxymoron?* 302 Sci. 1625, 1625 (2003)).

23. Déirdre Dwyer, *(Why) Are Civil and Criminal Expert Evidence Different?* 43 Tulsa L. Rev. 381, 391 (2007).

24. Susan D. Rozelle, Daubert, *Schmaubert: Criminal Defendants and the Short End of the Science Stick*, 43 Tulsa L. Rev. 597, 597 (2007).

25. Michael J. Saks, *Protecting Factfinders from Being Overly Misled, While Still Admitting Weakly Supported Forensic Science into Evidence*, 43 Tulsa L. Rev. 609, 611 (2007).

26. Mara L. Merlino, Victoria Springer, Jan Seaman Kelly, Derek Hammond, Eric Sahota & Lori Haines, *Meeting the Challenges of the* Daubert *Trilogy: Refining and Redefining the Reliability of Forensic Evidence*, 43 Tulsa L. Rev. 417, 432 (2007). I must admit that I find Merlino et al.'s interpretation of *Daubert* as "refining and redefining . . . reliability" puzzling. "Reliability"–or, as the *Daubert* court surely meant, "accuracy" or "validity," are scientific concepts that cannot necessarily be "redefined" by a court. *See infra* n. 51. As I understand it, *Daubert* changed the admissibility threshold to require evidence of reliability, rather than mere relevance or "general acceptance" in a field. *See* Fed. R. Evid. 702; *Frye.*, 293 F. 1013. *Daubert* did not "redefine" the notion of reliability; it simply required that reliability–defined in the same way it always has been defined–be shown for evidence to be admissible. *Daubert*, 509 U.S. at 590. The notion that *Daubert* "redefined" reliability suggests that latent print identification may have satisfied some prior notion of "reliability" only to have the Court "redefine" reliability. It suggests that the absence of evidence of reliability was caused by shifting legal definitions rather than by the simple omission of empirical data gathering. In fact, establishing the reliability of latent print identification has always simply required empirical evidence concerning the extent to which latent print examiners produce correct results, regardless of what the courts say.

fingerprinting has not been validated gave testimony that failed the *Daubert* requirements, "assert[ed] that the identification sciences are unreliable, biased, or invalid without any properly conducted research as a foundation for such assertions[,]" was appropriately characterized by Judge Michael Brennan, a New York Supreme Court judge, as a faux expert, and was precluded from testifying at trial "to [the] Judge['s] . . . credit."[27]

This is a far more serious matter than a mere difference of opinion precisely because, as discussed above, the reader's trust in what is said in a journal is almost entirely based upon the implicit assumption that what is published in a journal is at least in some degree trustworthy (or at least more trustworthy than something published with what convention holds to be a lesser imprimatur of authority, like, say, a "blog"), in that the authors were at least presumptively qualified to make the statements they made in their contributions. Thus, unlike a routine scholarly disagreement, this attack by one contributor on another contributor's qualifications might leave the consumer of the information contained in a journal somewhat perplexed. If, as Merlino et al. claim, one of the contributors to the Symposium is in the habit of making assertions "without any properly conducted research as a foundation for such assertions,"[28] why was that contribution permitted to survive the editorial review process that supposedly polices the quality of scholarship that appears in scholarly journals? If, on the other hand, Merlino et al.'s claim is false, why was their false claim permitted to survive the editorial review process that supposedly polices the quality of scholarship that appears in scholarly journals? In short, because one group of contributors calls another contributor unqualified to speak, the reader could only conclude that at least one of the contributions must be untrustworthy (either the accuser or the accused), potentially and needlessly, tainting an excellent collection of articles on an important subject.[29]

By now, it will probably come as no surprise if I disclose that I am the contributor whose authority was questioned by Merlino et al. Since the Symposium authors did not have access to one another's manuscripts prior to publication, I was not aware that my article would appear in a Symposium Issue that also contained claims that might lead a reader to conclude that my own article should not be trusted. I write this reply, not because I object to the airing of the fact that I was once (more than once actually) precluded from testifying at a criminal trial involving latent print evidence (as I was also more than once permitted to testify at such trials)—indeed, I have aired that fact in several of my own publications[30]—but, rather, because I am concerned that this fact has

27. Merlino et al., *supra* n. 26, at 444. Note that in New York, the Supreme Court is the felony trial court. The highest court is the New York Court of Appeals.

28. *Id.*

29. As already noted, the situation becomes even more acute if the reader is familiar with the perceived credibility problems with legal scholarship and is aware of the uncomfortable facts that *none* of these articles were peer reviewed and that they were selected through a process of invitation and evaluation by law students, and so on. These issues are, of course, common to most law reviews, including many of those considered the most prestigious and authoritative. At the same time, I am well aware that there were some unusual circumstances in the case of the production of this particular Symposium Issue that caused this apparent contradiction between its contents to fall through the cracks and that, under better circumstances, the contradiction might have been detected and addressed earlier in the publication process by, for example, allowing the authors to read and comment upon one another's papers.

30. Michael Lynch & Simon A. Cole, *Science and Technology Studies on Trial: Dilemmas of Expertise*, 35

been presented in such a way as to cast doubt on the veracity of the assertions I made in my contribution to the Symposium Issue. In scholarly discourse, the rhetorical tactic of *argumentum ad hominem*, attacking one's opponents rather than their arguments—"shooting the messenger," in colloquial terms—is considered not only unsporting but, more importantly, a sign of a weak position.[31]

In this reply, I will endeavor to clearly articulate—and support with evidence—just what the facts are with regard to the latent print evidence and to the exclusion of my testimony. In Part II, I will discuss use of my testimony by Merlino et al. as empirical evidence supporting the reliability of latent print identification. I will show that the preclusion of my testimony cannot serve as evidence of the reliability of latent print identification. Nor can it serve, as Merlino et al. seem to claim, as an explanation for the admissibility of latent print testimony in most U.S. criminal cases. Furthermore, Merlino et al.'s focus on one judge's characterization of my testimony apparently distracted them from engaging with the arguments of scholars, including myself, who have argued that the reliability of latent print identification has not been established. In Part III, I discuss the evidence that Merlino et al. adduce in support of the reliability of latent print identification. I show that their arguments are unconvincing and that, in most cases, legal scholars have pointed out the flaws in relying on the arguments they muster. I show that Merlino et al. have not engaged with these arguments.

II. IS EVIDENCE RELIABLE IF ITS CRITICS ARE UNRELIABLE?

The Merlino et al. article contains several parts, one of which is "an empirical content analysis of published judicial decisions concerning cases in which forensic document and latent fingerprint expert testimony were challenged following the 1993 *Daubert* decision."[32] This analysis, it is claimed, explains why "the majority of [admissibility] challenges to [fingerprint] evidence have been unsuccessful."[33] This is an important question, one that is taken up by several other contributions to the Symposium Issue. Most of the contributors blame pro-prosecution bias on the part of the judiciary, a hypothesis that had earlier been advanced by Professor Risinger.[34] One article argues "the judiciary has failed to apply *Daubert's* 'exacting standards' to forensic evidence offered by the prosecution,"[35] a view that, since the publication of the Symposium Issue, has now been endorsed by a National Academy of Science report on forensic science.[36]

Soc. Stud. of Sci. 269, 294–295 (2005); Simon A. Cole, *A Cautionary Tale About Cautionary Tales About Intervention*, 16 Org. 121, 126 (2009) [hereinafter Cole, *A Cautionary Tale*]; Simon A. Cole, *Does "Yes" Really Mean Yes?: The Attempt to Close Debate on the Admissibility of Fingerprint Testimony*, 45 Jurimetrics 449, 462 n. 84 (2005) [hereinafter Cole, *Does "Yes" Really Mean Yes?*].

31. *See e.g.* Epstein & King, *supra* n. 18, at 125.

32. Merlino et al., *supra* n. 26, at 418.

33. *Id.* at 443.

34. D. Michael Risinger, *Navigating Expert Reliability: Are Criminal Standards of Certainty Being Left on the Dock?* 64 Alb. L. Rev. 99 (2000).

35. Cooley & Oberfield, *supra* n. 22, at 285 (footnote omitted).

36. Natl. Research Council of the Natl. Acads., *Strengthening Forensic Science in the United States: A Path Forward* 85, 108–109 (Natl. Acads. Press 2009). "[T]here are serious issues regarding the capacity and quality of the current forensic science system; yet, the courts continue to rely on forensic evidence without fully understanding and addressing the limitations of different forensic science disciplines." *Id.* at 85. "[T]he undeniable reality is that the community of forensic science professionals has not done nearly as much as it

Another claims "there is some reason to believe that judges as a group are resistant to rejecting prosecution proffers of expert testimony"[37] because:

> when the issue is expertise that has been admitted without question for generations, and which has played a role in convicting many people (often in cases tried in front of those same judges in the past), and which the many judges who have been prosecutors before ascending to the bench have used in trying cases and convicting defendants, then the resistance becomes intense.[38]

A third suggests "that, almost regardless of the quality of the science involved, judges tend to admit scientific evidence when it favors the prosecution while refusing to admit it when it favors the defense."[39] Another contributor suggests four different answers: poor litigation by the defense bar; concerns about opening "the floodgates to appeal" if fingerprint evidence were excluded; a belief that admissibility, even if erroneous, is harmless because fingerprint evidence is always accompanied by corroborative evidence; and the belief that government experts (which latent print examiners usually are), unlike experts in civil litigation, are unbiased because they are public servants.[40] Finally, it is possible that evidence is routinely admitted simply in the spirit of inclusion, of erring on the side of allowing potentially relevant evidence. But this fails to explain why there is an "exclusionary ethos" in civil cases.[41]

Notice, however, that none of these contributors suggest that the reason latent print identification survives admissibility challenges is that the proponents of the evidence satisfied the *Daubert* reliability requirement. Merlino et al. offer a quite different answer to this question, one that, unlike the others,[42] purports to be based on empirical data. Merlino et al. state, "[c]ritical examination of published decisions and transcripts of these cases offers some insights into why" most admissibility challenges to latent print evidence have been unsuccessful.[43] They then state that "[a] comprehensive discussion of case law is beyond the scope of this paper . . . ."[44] Instead, they "offer as one example,"[45] a single unpublished trial court decision, *People v. Hyatt.*[46]

In *People v. Hyatt,* as Merlino et al. correctly note, the defendant sought to proffer my testimony at trial as rebuttal to the testimony of a latent print examiner who was proffered to testify that the defendant was the source of a particular latent print from the

---

reasonably could have done to establish either the validity of its approach or the accuracy of its practitioners' conclusions,' and the courts have been 'utterly ineffective' in addressing this problem." *Id.* at 108–109 (footnotes omitted).

37. 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, 473 (2007).

38. *Id.* at 473–474 (footnote omitted).

39. Rozelle, *supra* n. 24, at 597 (footnotes omitted).

40. Dwyer, *supra* n. 23, at 391–392.

41. Gary Edmond & David Mercer, Daubert *and the Exclusionary Ethos: The Convergence of Corporate and Judicial Attitudes towards the Admissibility of Expert Evidence in Tort Litigation*, 26 L. & Policy 231 (2004).

42. Except for Professor Risinger's earlier work, *supra* n. 34.

43. Merlino et al., *supra* n. 26, at 443.

44. *Id.*

45. *Id.*

46. No. 50115U (N.Y. Sup. Ct. Oct. 10, 2001).

crime scene to the exclusion of all other possible sources in the universe.[47] The state successfully challenged the admissibility of my testimony under New York's version of the *Frye* general acceptance standard.[48]

A. Two Quibbles with Merlino et al.'s Characterization of My Testimony

Before proceeding further, it is necessary to correct two assertions by Merlino et al. about my testimony in *Hyatt,* one of minor, and one of major, significance. First, Merlino et al. note that the prosecutor established that I "never consulted with the examiner who actually conducted the investigation to learn what methods and techniques were used."[49] While it is true that I did not consult with the latent print examiner, this classic line of cross-examination questioning was irrelevant to the testimony I was prepared to give. As I made clear in my testimony, I was prepared to testify that I had been able to find no study validating *any* method of latent print analysis.[50] Therefore, even giving the examiner the benefit of the doubt by assuming that he used what was then widely viewed, at least in North America, as the premier method available (the same "ACE-V method" described at length in Merlino et al.'s article),[51] I could still testify that I had been able to find no study validating whatever method the examiner used. In other words, I gave the benefit of the doubt to the examiner by assuming that he had used "ACE-V," the most credible method available. The only alternative to that assumption was that he used a less regarded method, not a better one. Under such circumstances, it was unnecessary to interview the examiner to determine which method he used.

The second issue is more significant. Merlino et al. write that I was proffered to "assert that the identification sciences are unreliable, biased, or invalid . . . ."[52] This is particularly important because Merlino et al. state that I made these assertions "without any properly conducted research as a foundation . . . ," and that is why "it is to Judge Hynes's [sic] credit that he . . ." precluded me from testifying at trial.[53]

I have never written or testified that latent print identification is unreliable or invalid. I have written and testified, in *Hyatt* and other cases, that its reliability[54] is

47. *See* Merlino et al., *supra* n. 26, at 436 (noting that the only incriminating conclusion latent print examiners are permitted to reach is that the defendant *is* the source of the print).

48. *See People v. Wesley*, 633 N.E.2d 451 (N.Y. 1994); *Frye*, 293 F. 1013.

49. Merlino et al., *supra* n. 26, at 444.

50. *See* Tr. Transcr. at 38–41, *Hyatt*, No. 50115U (available at http://www.onin.com/fp/ny_v_hyatt_simon_cole_testimony_4oct01.pdf).

51. Merlino et al., *supra* n. 26, at 432.

52. *Id.* at 444.

53. *Id.* The Judge's name was Michael Brennan. Charles Hynes was the Kings County District Attorney at the time.

54. In this article, I will follow Merlino et al.'s terminology by using the term "reliability," the term used by the Supreme Court in *Daubert,* to mean what scientists would mean by the terms "accuracy" or "validity." For scientists, "reliability" refers to consistency of results, regardless of their truth value. "Accuracy" refers to the degree to which results are, in fact, correct. Paul C. Giannelli, *The Admissibility of Novel Scientific Evidence:* Frye v. United States, *a Half-Century Later*, 80 Colum. L. Rev. 1197, 1201 n. 20 (1980). Legal scholars generally agree that when the *Daubert* Court said "reliability," it meant "accuracy," because it would defy common sense for the *Daubert* Court to have intended to create an admissibility threshold that merely insisted that expert evidence produce consistent results, regardless of whether those results were correct.

unknown and that it has not been validated.[55] I have written and testified that there are good theoretical reasons to believe that latent print identification is subject to the same sorts of normal psychological biases as other observational tasks as well as perhaps a pro-prosecution bias.[56] I have never written or said that the "identification sciences" as a whole are "biased." The distinction between asserting that the reliability of latent print identification is unknown and asserting that latent print identification is unreliable would not seem to be so subtle a distinction that Merlino et al. could not have been expected to grasp it.

Thus, the crux of my proffered testimony was not, as Merlino et al. claim, that latent print identification is unreliable, but rather, that we have no good data on its reliability. If this was the nature of my testimony, then what would count as "properly conducted research as a foundation for such" an assertion? I cannot think of anything other than what I actually did: using the tools of the historian of science, I researched the scientific, professional, and legal literature about latent print identification in search of data concerning the reliability of the technique. I found none, and I so testified. I cannot help noting that, if my assertion were false, it could have been more efficiently refuted by producing the reliability data that I claimed did not exist than by attacking my research methods. That an advocate, on cross-examination, might choose the less efficient route of attacking me, rather than producing the data, is perhaps understandable. For scholars to do so is less understandable.

Although Merlino et al.'s symmetrical notion that my testimony should be held "to Cole's own high standards"[57] has superficial appeal, it actually satisfies neither logic nor efficiency because latent print examiners and I make very different knowledge claims. A latent print examiner testifies that the defendant is the source of a particular latent print to the exclusion of all other possible sources. Since the consumer of the evidence (the court) has no way of determining the actual truth of this claim, *Daubert* directs that it inquire instead about the general reliability of people like this witness making claims of

---

55. *See* Tr. Transcr., *supra* n. 50, at 28–29.

> Q: Is it your opinion, Dr. Cole, your position, based upon your experience and study, that the fingerprint matching process that is used by examiners is not reliable?
>
> A: It's my opinion that its reliability has not been measured. It may or may not be reliable, but we don't know because we haven't tried to find out, we haven't tested it or measured it. It's an unknown reliability.
>
> . . .
>
> Q: But are you saying that it's not reliable evidence?
>
> A: It is of unknown reliability. It may turn out that it's reliable, it may turn out that it's not very reliable, we just don't know how reliable it is, I don't know either.

Id. at 28–29, 61.

56. Simon A. Cole, *More Than Zero: Accounting for Error in Latent Fingerprint Identification*, 95 J. Crim. L. & Criminology 985, 1060–1061 (2005); *see also* Itiel E. Dror & David Charlton, *Why Experts Make Errors*, 56 J. Forensic Identification 600 (2006); Itiel E. Dror et al., *Contextual Information Renders Experts Vulnerable to Making Erroneous Identifications*, 156 Forensic Sci. Intl. 74 (2006); Itiel E. Dror et al., *When Emotions Get the Better of Us: The Effect of Contextual Top-down Processing on Matching Fingerprints*, 19 Applied Cognitive Psychol. 799 (2005); Dan E. Krane et al., *Sequential Unmasking: A Means of Minimizing Observer Effects in Forensic DNA Interpretation*, 53 J. Forensic Sci. 1006 (2008); Larry S. Miller, *Procedural Bias in Forensic Science Examinations of Human Hair*, 11 L. & Hum. Behav. 157 (1987); D. Michael Risinger et al., *The* Daubert/Kumho *Implications of Observer Effects in Forensic Science: Hidden Problems of Expectation and Suggestion*, 90 Cal. L. Rev. 1 (2002).

57. Merlino et al., *supra* n. 26, at 444.

this nature. Hence, there is a demand for reliability data about latent print identification. Of course, reliability data will not definitely inform the court about the truth of the witness's claims in the instant case. Even if general reliability data for latent print identification shows high reliability, the examiner could be wrong in the particular case. And, if reliability data is low, the examiner could be correct. Knowing general reliability is, therefore, a second-best solution for those situations (which are most situations) when the accuracy of specific claims cannot be directly ascertained.

When I testify as a rebuttal witness, I testify that the aforementioned reliability data does not exist. I reached this conclusion by conducting a literature review. One could perhaps inquire about the reliability of the methods I used to reach that conclusion—whether I properly conducted the literature review—but why opt for the second best solution? The claim that reliability data does not exist can be directly and conclusively refuted simply by producing reliability data. Of course, one cannot prove a negative, and the fact that no one has yet produced reliability data after eight years of admissibility battles and more than a score of scholarly articles on the subject[58] does not absolutely prove that reliability data does not exist. But, nonetheless, the more effective and efficient way to attack the claim that I make would be to refute it directly rather than to attack the methods that produced it. That a prosecutor would, on tactical grounds, choose this option is perhaps understandable, but that scholars would endorse it is less so.

B. The Use of Hyatt *in Merlino et al.'s Empirical Project*

But let us turn now to the more important issue: the role of *Hyatt* in Merlino et al.'s argument. Merlino et al., recall, promised an explanation as to why admissibility challenges to latent print identification almost always fail, an explanation derived from an empirical analysis of "case law"[59] (actually cases made available in legal research databases).[60] They, then, in the interest of brevity, chose to discuss only one of those

58. *See infra* n. 96.

59. Merlino et al., *supra* n. 26, at 418.

60. Although it is not crucial to my argument, I feel it is necessary to comment briefly on Merlino et al.'s use of concepts drawn from the discipline in which I was trained, Science & Technology Studies, to explain the preclusion of my testimony. They state that "[b]oth boundary work and experimenters' regress are clearly evident in the transcript [of *Hyatt*]." *Id.* at 443. "Boundary work" is a term devised by Professor Gieryn to denote the process by which scientists define the boundaries of what counts as science or, in some cases, as a particular scientific discipline. Thomas F. Gieryn, *Boundary-Work and the Demarcation of Science from Non-Science: Strains and Interests in Professional Ideologies of Scientists*, 48 Am. Sociological Rev. 781 (1983). Merlino et al. apparently derive their notion of boundary work from Professor Jasanoff. Sheila Jasanoff, *What Judges Should Know About the Sociology of Science*, 32 Jurimetrics J. 345 (1992). The boundary work in *Hyatt* is clear enough: the court reasoned that only latent print examiners are qualified to serve as experts on the validity (or "reliability") of latent print identification (reasoning that, incidentally, would also result in only practicing astrologers being qualified to serve as experts on the validity of astrology). Jasanoff had invoked the notion of "boundary work" as an explanation of *scientists'* behavior in an article intended to educate judges about the behavior of scientists. Moreover, Jasanoff's stance toward boundary work was descriptive, not normative; she noted that it might reflect either efforts to "insulate[ ] scientific work from unexpected and possibly ill-motivated challenges by inadequately credentialed critics" or efforts "to deflect criticism that is not only unwelcome but well founded." *Id.* at 349. Jasanoff's message to judges—"What Judges Should Know"—was that they should be alerted to the fact that some utterances by scientists might be properly understood as exercises of boundary work rather than as factual statements about the world. As Jasanoff notes, such exercises in boundary work might serve good purposes or ill ones. *Id.* Merlino et al., however, invoke boundary work as an explanation for a *judge's* behavior. Moreover, they seem to view the judge's employment of boundary work

cases, *Hyatt,* in depth. It should be clear, however, from the preceding discussion that *Hyatt* cannot explain why admissibility challenges to latent print identification are seldom successful for the simple reason that *it was not an unsuccessful admissibility challenge to latent print identification.* Instead, it was a successful admissibility challenge to the proffered testimony of a fingerprint rebuttal witness. In other words, Merlino et al. purport to explain the admissibility of latent print evidence by reference to a case that was not about the admissibility of latent print evidence.

This lumping together for analysis of cases with different procedural postures is a general weakness of Merlino et al.'s methods. Merlino et al. performed a content analysis on a data set generated from "a subset of the total number of cases published on Lexis in which the admissibility of expert testimony about forensic document examination or latent print examination was challenged."[61] This data set contains both trial and appellate court opinions. But the decisions facing trial and appellate courts are quite different. A trial court decides, within its discretion, whether to admit evidence. An appellate court decides whether the trial court's decision was an abuse of that discretion.[62] These are quite distinct decisions, but Merlino et al. mixed these two sets of observations in a single set of combined results.

Hyatt, meanwhile, differs perhaps even more greatly from the other cases in Merlino et al.'s data set. Whereas the other cases involve challenges to the admissibility of the testimony of a latent print examiner, *Hyatt* involved a challenge to the admissibility of the testimony of a social scientist claiming expertise on the validity, or lack thereof, of latent print identification. These are two quite different things. I am at a loss to understand why Merlino et al. would select a single case in a procedural posture unique to their data set to explain the rest of the data set. Generally, if one were going to select only one case to explain a data set, one would want to select the most representative case.[63] As far as I can determine, Merlino et al. had a data set of thirty-eight challenges to the admissibility of latent print testimony and one challenge to the testimony of a rebuttal witness. In their contribution to the Symposium Issue, they invoke the one case to explain the outcome of the other thirty-eight. In methodological terms, their sample is not merely unrepresentative of their population of interest; it is not

---

as appropriate.

"Experimenters' regress," which Merlino et al. define, somewhat incompletely, as "deconstruction of the facts," is a notion posited by Professor Collins which disputes the widely held Popperian notion that theories can be falsified through "crucial experiments." Merlino et al., *supra* n. 26, at 442. Collins argues that, in fact, scientists can always dismiss unwelcome experimental results by invoking the claim that the experiment was performed incompetently. Strictly speaking, Collins argues, there is no limit to the extent to which such arguments can be deployed, although, in practice, the community of scientists does impose such limits. H.M. Collins, *Changing Order: Replication and Induction in Scientific Practice* (SAGE Publications 1985). Merlino et al. never make clear where in the *Hyatt* transcript they perceive the experimenters' regress. *See* Merlino et al., *supra* n. 26, at 443.

61. Merlino et al., *supra* n. 26, at 419 n. 11.

62. *See* Risinger, *supra* n. 37, at 461–462, 468–469.

63. I dismiss the possibility that Merlino et al. were intending to smear me. If they had been, they surely would have reproduced and emphasized the part of the *Hyatt* opinion that characterizes me as a "junk scientist" for claiming that there are no reliability studies of latent print identification, as has another of my critics. *See e.g.* André Moenssens, *Court Excludes Fingerprint Critic's Testimony as "Junk Science"*, http://www.forensic-evidence.com/site/ID/Cole_junksci.html (last accessed Nov. 10, 2009). How pointing out the absence of evidence supporting an expert's knowledge claims counts as "junk science" remains a mystery to me.

even included in their population of interest. This is certainly an odd methodological approach from scholars who take me to task for supposedly making assertions "without any properly conducted research as a foundation . . . ."[64]

A further question, however, is what sort of explanation for the admissibility of latent print evidence Merlino et al. are trying to posit by discussing the *Hyatt* case in depth. *Hyatt* itself, of course, offers no explanation as to why latent print evidence should be admissible since that question was not at issue in the case. Instead, by citing *Hyatt,* Merlino et al. appear to be taking the position that *it is the exclusion of the evidence's critics from testifying at trial that renders the evidence admissible.* In other words, the legal principle being expressed appears to be: evidence is reliable if its critics are unreliable. As Basil Fawlty would say, "[i]t's *my* fault."[65] Latent print evidence is reliable not because of empirical data supporting its reliability, but because of something that a socio-legal scholar did.

I want to be clear that I am taking no position here on the *Hyatt* holding itself. I will simply say that I would imagine that somewhat reasonable legal arguments could be made both for and against allowing the testimony of a rebuttal witness who states that there is no data concerning the reliability of a technique used by another witness, and, indeed, courts have split on that issue.[66] For example, the issue of whether the testimony of rebuttal social science experts on the reliability of eyewitness identification should be admissible is an area of lively legal controversy.[67]

Merlino et al.'s "[c]ritical examination"[68] of Judge Brennan's decision does not discuss any legal principles that underlies it, other than to say that the decision was "to [his] credit," and they do not explain why they believe my testimony was appropriately precluded. One possibility is that they believe there is some principle of law served by precluding rebuttal witnesses to certain forms of expert evidence under certain circumstance. But Merlino et al. articulate no such principle. Another possibility is that they believe that such rebuttal testimony might be appropriate in some cases, but that in my particular case I relied on inappropriate or unreliable methods to reach the conclusion to which I was prepared to testify. But, if that is the case, everything hinges on their misunderstanding of the nature of my testimony. If I were proposing to testify that latent print identification is "unreliable" (whatever that means), it would, indeed, be inappropriate to give such testimony without having conducted an empirical study of its reliability. But if, as was indeed the case, I were proposing to testify that there is no empirical study of the reliability of latent print identification, I don't know what method, other than the one I used—searching for such a study in the appropriate literature—Merlino et al. would have had me use. The third possibility is that Merlino et al. believe

64. Merlino et al., *supra* n. 26, at 444.

65. *See supra* n. 1 and accompanying epigraph.

66. Cole, *Does "Yes" Really Mean Yes? supra* n. 30, at 459–460; D. Michael Risinger, *Handwriting Identification,* in *Science in the Law: Forensic Science Issues* 113 (David L. Faigman et al. eds., West 2002).

67. *See e.g.* Rozelle, *supra* n. 24, at 600. I would imagine that an argument for my testimony would be rooted in the Sixth Amendment right to rebut the government's testimony. For what it is worth, I would think that the argument in favor of allowing a counter-expert to rebut the testimony of an expert witness, like a latent print examiner, would be even stronger than the argument in favor of allowing a counter-expert to rebut the testimony of a lay eyewitness.

68. Merlino et al., *supra* n. 26, at 443.

that my testimony was false. If this were true, it would cast doubt not merely on my proffered testimony, but on my scholarship as well. Was my testimony false? As stated above, the most efficient and persuasive way to demonstrate that it was would be to produce a reliability study. But, as I will discuss in Part III, when given the opportunity, Merlino et al. produced no such study.

The *Hyatt* court's reasoning is quoted by Merlino et al. so I leave the issue of its legal persuasiveness to the reader to determine for herself.[69] My concern here is not with the preclusion of my testimony but with the astonishing transformation of the preclusion of my testimony into supposed evidence of the reliability of latent print evidence. Again, the argument appears to be that if rebuttal witnesses to a form of proffered evidence are lacking in reliability, then the evidence itself must be reliable. Surely, Merlino et al. do not want to advance this as a principle—legal, rational, or otherwise? Should we admit para-psychological evidence because one of its chief debunkers, James "The Amazing" Randi, is a magician and therefore is "lacking in scientific method?"[70] Even if we, for the sake of argument, take the *Hyatt* court at its word and assume that it was correct in stating that my testimony was "lacking in scientific method," that does not in any way render latent print evidence reliable.

III. SO, IS LATENT PRINT IDENTIFICATION RELIABLE?

In addition to being faulty reasoning, Merlino et al.'s hanging of the reliability of latent print identification on the supposed unreliability of its supposed critics has a second unfortunate consequence: preventing them from fully engaging in the scholarly debate over the reliability of latent print identification. As discussed above, at least four contributors to the Symposium Issue take the position that the reliability of latent print identification has not been demonstrated, whereas Merlino et al., though they do not explicitly say so,[71] appear to take the opposite view. In this section, I discuss the evidence put forward by Merlino et al. in support of the claim that the reliability of latent

69. For the full text of the decision, see *Hyatt*, No. 50115U.

70. In fact, Randi is a proponent of a rigorously defined "scientific method." *See generally* T.J. Pinch & H.M. Collins, *Private Science and Public Knowledge: The Committee for the Scientific Investigation of the Claims of the Paranormal and Its Use of the Literature*, 14 Soc. Stud. of Sci. 521 (Nov. 1984).

71. The text of Merlino et al.'s article is somewhat ambiguous as to whether they believe that the reliability of latent print identification has been established. As will be discussed *infra* Part III.A, at one point they state, "[t]he reliability of latent print examination methods is supported," though they do not say "established," "by the use of Automated Fingerprint Identification Systems (AFIS)." Merlino et al., *supra* n. 26, at 437. At another point, they write of "terms" that "collectively refer to the methodology of latent print examination (i.e. ACE-V), its reliability, [and] the methods used to test its reliability . . . ." *Id.* at 432. At another point, they state that "forensic practitioners have responded to the questions about the reliability of their testimony by seeking ways to both improve their disciplines and demonstrate to judges, attorneys, academicians, and fellow experts that their underlying assumptions, methods, and conclusions meet the requirements of the *Daubert* trilogy." *Id.* at 418. This would seem to indicate that they believe that latent print identification already satisfies *Daubert's* reliability requirement. However, elsewhere they state that latent print examiners "are working to define and establish valid and reliable measures of proficiency and error," which would seem to suggest that the establishment of reliability remains an ongoing project. *Id.* at 444. It should be noted that it is not sufficient to simply say that the establishment of reliability will always be a work-in-progress, as methods improve and science progresses. While the reliability of latent print identification (i.e. its accuracy rate) might be expected to improve over time, the contributors to the Symposium Issue have argued that no adequate reliability measurements have yet been put forward on behalf of latent print identification. This is a claim that is either true or false.

print identification has been demonstrated. On each point, I show that the evidence is unconvincing. Moreover, on most points, I show that published scholarship exists that states that the arguments they muster do not logically support claims of the reliability of latent print identification.[72] This is, of course, by no means to say that this published scholarship is correct on each of these points. But, if Merlino et al. reject these arguments, it is incumbent upon them to explain why. Instead, as I will show below, none of these arguments are discussed or even cited by Merlino et al.

A. Automated Fingerprint Identification Systems

Merlino et al. make no reference to any empirical study of the reliability of latent print identification. They do, however, refer to various evidentiary items that, I assume, they believe support the reliability of latent print identification. Only on one occasion, however, do they affirmatively state that a particular piece of evidence supports the reliability of latent print identification: "[t]he reliability of latent print examination methods is supported by the use of the Automated Fingerprint Identification Systems (AFIS)."[73] AFIS search prints against databases of prints and produce lists of candidate matches. AFIS manufacturers make no claim that the top candidate is the true source of the searched print. Indeed, AFIS will produce candidate lists even if the true source is not in the database. Therefore, AFIS do not generate conclusions of source; human latent print examiners do. What, then, can we infer from the existence of AFIS searches? Merlino et al.'s argument is that AFIS searches generate candidate prints that are very similar to the searched prints. "Similarity, however, is not identity."[74] Presumably, the inference is that no two candidate prints are identical, or all are unique. But, if this is meant to support the reliability of latent print identification, it invokes what Saks and Koehler have called the "individualization fallacy" and what I have called the "fingerprint examiner's fallacy," the claim that the reliability of a forensic technique can be inferred from the uniqueness of the target object.[75] This reasoning has been widely criticized in the forensic literature.[76] The relevant question under *Daubert* is the

72. Since publication of the Symposium Issue, a National Research Council Report has concluded that neither latent print identification nor forensic document examination has demonstrated their ability to reach correct conclusions about the sources of forensic traces. Natl. Research Council of the Natl. Acads., *supra* n. 36, at 7.

73. Merlino et al. *supra* n. 26, at 437. AFIS are computer systems designed to search unknown prints against databases of known prints. Although AFIS are sometimes used to match complete sets of prints, when searching latent prints they do not produce "matches" but rather candidate lists against which human latent print examiner can compare the unknown print. *See id.*

74. *Id.*

75. Michael J. Saks & Jonathan J. Koehler, *The Individualization Fallacy in Forensic Science Evidence*, 61 Vand. L. Rev. 199 (2008); Simon A. Cole, *Grandfathering Evidence: Fingerprint Admissibility Rulings from* Jennings *to* Llera Plaza *and Back Again*, 41 Am. Crim. L. Rev. 1189, 1197–1203 (2004).

76. Keith Inman & Norah Rudin, *Principles and Practice of Criminalistics: The Profession of Forensic Science* 54 (CRC Press 2001); John I. Thornton & Joseph L. Peterson, *The General Assumptions and Rationale of Forensic Identification,* in *Science in the Law: Forensic Science Issues* 1, 25 (David L. Faigman et al. eds., West 2002); David A. Stoney, *Measurement of Fingerprint Individuality*, in *Advances in Fingerprint Technology* 327 (Henry C. Lee & R. E. Gaensslen eds., 2d ed., CRC Press 2001); *see also* Christophe Champod & Ian W. Evett, *A Probabilistic Approach to Fingerprint Evidence*, 51 J. Forensic Identification 101, 115 (2001); Didier Meuwly, *Forensic Individualisation from Biometric Data*, 46(4) Sci. & Just. 205, 207 (2006); David J. Balding, *Weight-of-Evidence for Forensic DNA Profiles* 54 (Wiley & Sons 2005); B.W.N. Robertson, *Fingerprints, Relevance and Admissibility*, 2 N.Z. Recent L. Rev. 252, 255 (1990).

reliability of latent print identification, not the uniqueness of human friction ridge skin, which, even if true, is only necessary but not sufficient for the testimonial claims latent print examiners make. But Merlino et al. neither rebut, nor even cite, any of this literature.

Merlino et al. go on to say, "[u]tilizing the full scope of quantitative-qualitative analysis, forensic scientists are able to discriminate between the chance correspondence of limited data sets and proper conclusions of identification based on a complete analysis."[77] In lay terms, I take this to mean that, given an AFIS candidate that appears quite similar to a searched print, professional latent print examiners are able to distinguish between: (1) cases in which the cause of the similarity is that the two prints derive from different sources and yet, by chance, appear quite similar; and (2) cases in which the cause of the similarity is that the two prints derive from the same source. This discriminating ability is, of course, precisely what we are asking about when we ask about the reliability of latent print identification. Merlino et al.'s statement, quoted above, asserts that latent print examiners "are able" to discriminate in this way, but they offer no evidence in support of this claim—no measurements of latent print examiners' discriminating ability at this task. The statement has a footnote, and there one might hope to find reference to some sort of study or empirical data of this kind. Instead, in the footnote, we find a discussion of the supposed *source* of latent print examiners' "discriminating power,"[78] but no reference to any measurement *of* this supposed "power."[79] At the end of the footnote, there are two citations, but these are references to two studies that used *computers,* not latent print examiners, to try to discriminate between prints from the same source and prints from different sources.[80] Merlino et al. are not the first to try to infer the reliability of latent print identification from the existence of AFIS searches. Both latent print practitioners and courts have previously mounted such arguments. I have critiqued these efforts and argued that the reliability of latent print identification cannot be inferred from the existence of AFIS searches.[81] Merlino et al. do not rebut, or even cite, these arguments.

B. One Hundred Years of Empirical Data

Aside from AFIS searches, Merlino et al. do not point to any other data that they explicitly claim supports the reliability of latent print identification. In the guise of discussing the *Daubert* factors, however, they do discuss a number of items that, one might infer, they believe support the reliability of latent print identification. In the interest of completeness I will discuss these items, even though Merlino et al. do not

77. Merlino et al., *supra* n. 26, at 437 (footnote omitted).

78. *Id.* at 437 n. 69. The footnote states, "[a]t least part of this discriminating power stems from the rarity of features found in the friction ridge skin impression." *Id.* This would indeed be one important component of any measurement of discrimination. The other would be the reliability with which the analytic system is able to perceive those features.

79. *Id.*

80. *See* Christophe Champod et al., *Fingerprints and Other Ridge Skin Impressions* (CRC Press 2004); Sargur N. Srihari et al., *Discriminability of Fingerprints of Twins*, 58 J. Forensic Identification 109 (2008).

81. Simon A. Cole, *Is Fingerprint Identification Valid? Rhetorics of Reliability in Fingerprint Proponents' Discourse*, 28 L. & Policy 109 (2006) [hereinafter Cole, *Is Fingerprint Identification Valid?*]; Cole, *Does "Yes" Really Mean Yes? supra* n. 30.

explicitly state that they support the reliability of latent print identification. I will again show that these items of evidence cannot support the reliability of latent print identification and, in most cases, that arguments on this point exist in the published literature.

Merlino et al. state that "[l]atent print examination draws on . . . over one-hundred [sic] years of empirical data collected by practitioners."[82] It is difficult to infer from the text what they mean by this. It is possible that they are referring to latent print examiners' collective experience observing latent prints in which they have supposedly never observed two exactly identical complete fingertip-sized areas of friction ridge skin. Of course, if one were to treat latent print examiners going about their duties as empirical data collection, one would have to note that the data is in no way being systematically collected, recorded, organized, or classified. But, more importantly, this argument, like the argument discussed in the previous section, invokes the "fingerprint examiner's fallacy" because it speaks only to the supposed "uniqueness" of human friction ridge skin, not to the issue of the reliability of latent print identification.

It is also possible that Merlino et al. are referring to casework performed by latent print examiners. In that case, the argument is presumably that each case functions as a *de facto* experiment testing the accuracy of the latent print examiner's conclusion. However, casework cannot constitute empirical data concerning the reliability of latent print identification because the ground truth is not known in casework. Many other latent print proponents have advanced the casework argument,[83] and the argument has been disputed in the scholarly literature.[84] Indeed, the casework argument is even disputed by a federal judge in one of the cases in Merlino et al.'s data set.[85] They offer no rebuttal, or even awareness, of these arguments.

C. Research on Latent Print Examiners' Cognition

Merlino et al. state that Busey and Vanderkolk have "undertaken research" on latent print examiners' cognitive processes and that "preliminary results suggest that expert latent print examiners exhibit discrete adaptive changes in cognitive processing related to the pattern recognition of finger prints."[86] In layman's terms, I take this to mean that latent print examiners' brains react differently when looking at fingerprint patterns than do laypersons' brains. Needless to say, this research, even if it were not "preliminary" would not establish the reliability of latent print identification. And, to their credit, Merlino et al. do not explicitly say that it does.

---

82. Merlino et al., *supra* n. 26, at 434.

83. Cole, *Is Fingerprint Identification Valid? supra* n. 81.

84. *See* Simon A. Cole, *"Implicit Testing": Can Casework Validate Forensic Techniques?* 46 Jurimetrics 117 (2006); Lyn Haber & Ralph Haber, *Scientific Validation of Fingerprint Evidence under* Daubert, 7 L., Probability and Risk 87, 96 (2008); Tamara F. Lawson, *Can Fingerprints Lie? Re-weighing Fingerprint Evidence in Criminal Jury Trials*, 31 Am. J. Crim. L. 1, 37 (2003); Jennifer L. Mnookin, *Fingerprint Evidence in An Age of DNA Profiling*, 67 Brook. L. Rev. 13, 65 (2001).

85. *U.S. v. Llera Plaza*, 179 F. Supp. 2d 492 (E.D. Pa. 2002).

86. Merlino et al., *supra* n. 26, at 435.

D. Descriptive and Experimental Studies: Anatomical Research

In the section titled "Falsifiability,"[87] Merlino et al. state that "[m]any studies, both descriptive and experimental, exist concerning the theories and methods of latent print examination."[88] They then go on to discuss some anatomical research on the formation of friction ridge skin, the anatomical structure of which "fingerprints" are representations, which they characterize as "descriptive studies."[89] But anatomical researchers say nothing about the reliability of latent print identification. Anatomical researchers are concerned with the formation of friction ridge skin, and they have nothing to say about the reliability of latent print identification as practiced by latent print examiners. Merlino et al. state that "the contemporary work of investigators such as William Babler and others . . . provide the scientific mechanism for the uniqueness and permanence of friction ridge skin."[90]

Again, Merlino et al. are hardly the first to make this argument. Proponents of latent print identification have long argued that anatomical research supports claims about its reliability.[91] Merlino et al. do not explicitly state how it is they think the reliability of latent print identification may be inferred from anatomical research, and, indeed, I have argued that anatomical knowledge about the formation of friction ridge skin tells us very little about the reliability of latent print identification.[92] Anatomical knowledge does support the claim that friction ridge skin is highly variable, and accurate latent print identification would be extremely difficult if friction ridge skin were not so variable. But variability does not tell us how accurate latent print identification actually is. I have also argued that Dr. Babler has said nothing about the reliability of latent print identification and that he has never claimed to have proven the uniqueness of friction ridge skin but merely asserted it.[93] Of course, I could be wrong. But Merlino et al. neither rebut nor even cite these arguments.

It is not clear what precisely Merlino et al. had in mind when they stated that there were "experimental" studies on the theory and method of latent print analysis. But this section does not refer to any formal studies beyond the anatomical ones, and Merlino et al. do not use the term "experimental studies" anywhere else in this section.

E. Error Rate

Merlino et al. engage in an extensive discussion of the error rate of latent print identification. Although Merlino et al. do not explicitly connect the notion of error rate to that of reliability, it could reasonably be said that an empirically based estimated error

87. In my view, it is incorrect to characterize the first *Daubert* prong as "falsifiability," rather than as, say, "testing." It seems clear to me that *Daubert* asks not merely whether proffered experts make falsifiable claims, but whether their falsifiable claims have survived vigorous efforts at falsification, i.e. "testing."

88. Merlino et al., *supra* n. 26, at 435.

89. *Id.*

90. *Id.*

91. André Moenssens, *Fingerprint Identification: A Valid Reliable "Forensic Science"?* 18 Crim. Just. 31 (2003); André Moenssens, *The Reliability of Fingerprint Identification: A Case Report*, http://www.forensic-evidence.com/site/ID/pollak2002.html (last updated Jan. 18, 2002).

92. Cole, *Is Fingerprint Identification Valid? supra* n. 81.

93. *Id.*; *see also* Simon A. Cole, *Out of the Daubert Fire and into the Fryeing Pan? The Admissibility of Latent Print Evidence in* Frye *Jurisdictions*, 9 Minn. J. L. Sci. & Tech. 453 (2008).

rate would constitute the estimated reliability (or, more precisely, the estimated accuracy) of latent print identification. Therefore, Merlino et al.'s comments on error rate seem quite relevant.

Merlino et al. argue that estimating the error rate of latent print identification through an experimental validation study, as proposed by Drs. Haber and Haber, "seems quite problematic."[94] Their reason is that the variability, or substructuring, among laboratories that perform latent print analysis prevents the calculation of a meaningful industry-wide error rate. In fact, variability among laboratories does not pose an insurmountable obstacle to the calculation of an error rate. Statisticians, quality assurance auditors, failure analysis engineers, and many others regularly calculate industry-wide error rates, despite comparable variability. The variability of forensic laboratories is no greater or more daunting than the variability of other techno-social endeavors for which "error" or "failure" rates are regularly calculated, like airplane crashes. Variability poses methodological difficulties but not insurmountable ones. Once again, Merlino et al.'s argument has been mounted before and rebutted before, most notably by Professor Koehler.[95] Of course, Professor Koehler could be wrong on this point, but his argument is, again, neither rebutted nor cited.

F. Certification

It is perhaps worth making one clarification regarding the issue of certification, which is also discussed by Merlino et al. The authors note that the International Association for Identification (IAI), a professional organization, administers a certification program. They note that applicants for certification with "two years experience in the field . . . must first document their education, training, and professional experience, and must then pass a detailed written examination comprised of latent print comparison exercises, pattern interpretation of inked prints, and a series of questions" pertinent to fingerprint analysis.[96] Their discussion, however, does not make it entirely clear that certification apparently has little to do with the admissibility or reliability of latent print testimony as practiced in the United States for the simple reason that not all latent print analysts in the United States are certified. The IAI has taken the position that certification is *not* a requirement for competent latent print analysis, and there are many uncertified latent print analysts practicing and testifying in United States courtrooms.[97] Merlino et al. do not report that certification makes a difference in judges' determination of the admissibility of latent print testimony, and indeed, I am aware of no case in which a judge has made IAI certification a requirement for admissibility. Thus, even if certification made latent print identification reliable—and Merlino et al. do not claim that

94. Merlino et al., *supra* n. 26, at 438; Haber & Haber, *supra* n. 84.

95. Jonathan J. Koehler, *Proving the Case: The Science of DNA: On Conveying the Probative Value of DNA Evidence: Frequencies, Likelihood Ratios, and Error Rates*, 67 U. Colo. L. Rev. 859, 873 (1996). Professor Koehler expands on this argument, specifically with regard to latent print identification. Jonathan J. Koehler, *Fingerprint Error Rates and Proficiency Tests: What They Are and Why They Matter*, 59 Hastings L.J. 1077 (2008). However, it should be noted that this latter article was published after the Symposium. *See also* Cole, *supra* n. 56, at 1037.

96. Merlino et al., *supra* n. 26, at 434.

97. James R. McConnell, *Certification (To Be or Not to Be)*, 42 J. Forensic Identification 205 (1992); Pat A. Wertheim, *re: Certification (To Be or Not to Be)*, 42 J. Forensic Identification 280 (1992).

it does—it appears to exert little or no influence on judges who routinely admit latent print evidence whether certified or not.

G. General Acceptance

Merlino et al. argue that latent print identification meets the "general acceptance" requirement that was stated in *Frye v. U.S.* and then integrated into *Daubert* and its progeny case *Kumho Tire v. Carmichael.*[98] This is because "[t]he relevant scientific community for latent print examiners includes members of forensic science organizations such as the International Association of [sic] Identification and the American Academy of Forensic Sciences [AAFS], which encompass many forensic disciplines."[99] Again, this is a familiar argument. I have, however, argued, first, that there is little explicit evidence that practitioners of disciplines other than latent print analysis "accept" that the reliability of latent print analysis has been established—indeed, it is not clear to me that the membership of the AAFS actually does accept the notion that latent print individualization has been validated; and, second, that treating members of the IAI (i.e. professional latent print analysts) as the relevant community for the evaluation of the claim that the reliability of latent print analysis has been established amounts to "self-validat[ion.]"[100] Courts have looked dimly on this notion of self-validation when applied to other forms of evidence. Again, this work is neither rebutted nor cited by Merlino et al., although, in fairness, its publication postdated the writing of their article.

In summary, Merlino et al.'s arguments in support of the reliability of latent print identification rest upon many of the same arguments that have been mustered by other proponents of latent print identification: that friction ridge skin is unique; that there is anatomical research on the formation of friction ridge skin; the AFIS searches do not produce duplicate prints; the latent print examiners use a process; that there is education and training; that *some* examiners are certified; that some quality control measures are in place; that the reliability of latent print identification is accepted by professional latent print examiners. In repeating these arguments, Merlino et al. have not taken account of the extensive criticisms of these arguments that have been mounted by legal scholars, forensic scientists, and other scholars.

IV. CONCLUSION

In this Reply, I have objected to Merlino et al.'s characterization of my testimony in *Hyatt* and the inclusion of their characterization, unremarked upon, in their contribution to the Symposium Issue of this journal.[101] I have objected on both empirical grounds and in the interest of promoting civil, reputable scholarly discourse.

On empirical grounds, I have argued that my testimony fails to serve as evidence for the argument for which they use it because it was not representative of the population

98. 526 U.S. 137 (1999).

99. Merlino et al., *supra* n. 26, at 441.

100. Bert Black, *A Unified Theory of Scientific Evidence*, 56 Fordham L. Rev. 595, 633 (1988); Cole, *supra* n. 93.

101. Jane Campbell Moriarty, *Symposium Foreword*, 43 Tulsa L. Rev. 229 (2007).

of cases Merlino et al. purport to explain. Moreover, the explanation they construct around my testimony—that the reliability of expert testimony can be imputed from the supposed untrustworthiness of its rebutters—even if it were empirically true, would constitute an astonishingly circular form of scientific or legal reasoning. I suggested that, rather than simply present such a seemingly bizarre legal argument with a straight face, Merlino et al. might reasonably be expected either to "[c]ritically examin[e]"[102] it, or to defend it.

With regard to the promotion of scholarly discourse, I have argued that Merlino et al.'s use of my testimony is problematic for two reasons. First, it may reasonably be read as an attack on my qualifications to speak, rather than my arguments. Presented with a manuscript that characterized one of its contributors as someone who makes empirical assertions "without any properly conducted research as a foundation for such assertions,"[103] one might have expected a journal that takes seriously the policing of published scholarship to investigate the matter. If it were true that I make empirical assertions "without any properly conducted research as a foundation for such assertions,"[104] then perhaps my contribution ought not to be published, or, at least, those assertions should be removed. If, on the other hand, that accusation was false, then perhaps the contribution that made the accusation ought not to be published, or, at least, the groundless accusation should be removed.

Most disappointingly of all, Merlino et al.'s focus on one judge's characterization of my testimony seems to have distracted them from engaging with scholars who hold opinions different from theirs, myself included. Although I have authored numerous scholarly articles which speak to many of the points made by Merlino et al., not a single one is cited. A reader of their article would come away with the perception that I am a professional witness, rather than a scholar. Nor, with two exceptions,[105] do they cite any of the literature by the many other excellent scholars who have published on many of the issues they address.[106] Of course, Merlino et al. are scholars who are free to cite

102. Merlino et al., *supra* n. 26, at 443.

103. *Id.* at 444.

104. *Id.*

105. The exceptions are: Haber & Haber, *supra* n. 84; Mnookin, *supra* n. 84. *See* Merlino et al., *supra* n. 26, at 437, 441.

106. Some of the literature they might have engaged with would include David Faigman et al., *Modern Scientific Evidence: The Law and Science of Expert Testimony* § 27-2.3.1, at 386 (2nd ed., West 2002) ("Woe to fingerprint practice were such [*Daubert* admissibility] criteria applied!"); Michael Saks, *Merlin and Solomon: Lessons from the Law's Formative Encounters with Forensic Identification Science*, 49 Hastings L.J. 1069, 1106 (1998) ("By conventional scientific standards, any serious search for evidence of the validity of fingerprint identification is going to be disappointing .... A vote to admit fingerprints is a rejection of conventional science as the criterion for admission. A vote for science is a vote to exclude fingerprint expert opinions." (footnote omitted)); James E. Starrs, *Judicial Control Over Scientific Supermen: Fingerprint Experts and Others Who Exceed the Bounds*, 35 Crim. L. Bull. 234, 243 (1999) ("Instead of meaning being incapable of error, fingerprint identifications are declared to be infallible on account of the uniqueness of fingerprints to each person . . ."); David A. Stoney, *supra* n. 76, at 383 ("From a statistical viewpoint, the scientific foundation for fingerprint individuality is incredibly weak."); David L. Faigman, *Is Science Different for Lawyers?* 297 Sci. 339, 340 (2002) (fingerprinting has "not been seriously tested"); Paul Giannelli, *Fingerprints Challenged!* 17 Crim. Just. 33, 35 (Spring 2002) ("In its interpretation of *Daubert, Plaza I* is a well-written opinion. *Havvard* is not."); Robert Epstein, *Fingerprints Meet* Daubert*: The Myth of Fingerprint "Science" is Revealed*, 75 S. Cal. L. Rev. 605, 657 (2002) ("Having considered the various indicators of reliability set forth by the Supreme Court in *Daubert*, it is evident that at the present time, latent fingerprint identifications do not constitute reliable evidence."); Jessica M. Sombat, *Latent Justice:* Daubert*'s Impact on*

whomever they wish. But because of its failure to respond to or cite the scholarly literature that addresses the points it makes, Merlino et al.'s article fails to engage in the kind of scholarly discourse that would move the debate forward. One would hope that, in the pages of law journals at least, the debate over the reliability of latent print identification can be based on evidence and logical arguments. Perhaps, a debate of that kind can now begin.

---

the Evaluation of Fingerprint Identification Testimony, 70 Fordham L. Rev. 2819, 2825 (2002) ("the result Judge Pollak reached when he excluded expert testimony concerning fingerprints [in *Llera Plaza I*] was fair."); *Recent Cases*, 115 Harv. L. Rev. 2349, 2352 (2002) ("Fingerprint expert testimony does not survive application of the *Daubert* factors . . ."); Lyn Haber & Ralph Norman Haber, *Error Rates for Human Latent Fingerprint Examiners*, in *Automatic Fingerprint Recognition Systems* 339 (Nalini Ratha & Ruud Bolle eds., Springer 2004) (pointing out that no data have been collected on how accurately latent print examiners match different images of the same finger); Donald Kennedy, *Forensic Science: Oxymoron?* 302 Sci. 1625, 1625 (2003) (Fingerprinting's "reliability is unverified either by statistical models of fingerprint variation or by consistent data on error rates"); David H. Kaye, *The Nonscience of Fingerprinting:* United States v. Llera Plaza, 21 QLR 1073, 1087 (2003) ("As *Llera-Plaza I* so clearly reveals, this [the evidence advanced in support of the admissibility of latent fingerprint individualization] does not satisfy *Daubert*."); Jennifer L. Mnookin, *Fingerprints: Not a Gold Standard*, 20 Issues Sci. & Tech. 47, 47 (2003) ("Judge Pollak's first opinion [restricting latent fingerprint individualization testimony] was the better one."); Tamara F. Lawson, *Can Fingerprints Lie? Re-weighing Fingerprint Evidence in Criminal Jury Trials*, 31 Am. J. Crim. L. 1, 65 (2003) ("Currently fingerprint analysis is under attack because of the lack of study done on the accuracy of the examiners . . ."); Tara Marie La Morte, *Sleeping Gatekeepers:* United States v. Llera Plaza *and the Unreliability of Forensic Fingerprinting Evidence under* Daubert, 14 Alb. L.J. Sci. & Tech. 171, 173 (2003) (discussing "strong indications that the fingerprinting field should not survive a rigorous *Daubert* analysis" (footnote omitted)); Jane Campbell Moriarty, *Psychological and Scientific Evidence in Criminal Trials* vol. 1, § 12:15 (Clark Boardman Callaghan 1996) ("The assumption of the validity of fingerprinting rests upon law, rather than science."); Nathan Benedict, *Fingerprints and the* Daubert *Standard for Admission of Scientific Evidence: Why Fingerprints Fail and a Proposed Remedy*, 46 Ariz. L. Rev. 519, 538 (2004) ("[J]udges have generally relied on their instincts and the long history of judicial acceptance of fingerprint evidence to admit it without serious consideration of the science behind it."); Sandy L. Zabell, *Fingerprint Evidence*, 13 J.L. & Policy 143, 178 (2005). ("*ACE-V is an acronym, not a methodology.*" (emphasis in original)); Michael Mears & Therese M. Day, *The Challenge of Fingerprint Comparison Opinions in the Defense of a Criminally Charged Client*, 19 Ga. St. U. L. Rev. 705, 745 (2003); Katherine Schwinghammer, *Fingerprint Identification: How "The Gold Standard Of Evidence" Could Be Worth Its Weight*, 32 Am. J. Crim. L. 265 (2005).