PREFERRING THE “WISE MAN” TO SCIENCE: THE FAILURE OF COURTS AND NON-LITIGATION MECHANISMS TO DEMAND VALIDITY IN FORENSIC MATCHING TESTIMONY

PROFESSOR JULIUS EPSTEIN

A single conversation with a wise man is better than ten years of study. Chinese Proverb

The saddest aspect of life right now is that science gathers knowledge faster than society gathers wisdom. Isaac Asimov, Isaac Asimov's Book of Science and Nature Quotations, 1988.

INTRODUCTION

Forensic evidence is often critical in investigating cases and proving culpability. Studies from the 1980s “found [that] clearance rates of offenses with evidence scientifically analyzed were about three times greater than in cases where such evidence was not used[.]” and this trend has continued. Notwithstanding both its importance and its utility, much forensic evidence other than DNA has been criticized for its lack of scientific validation, with the most significant and trenchant assessment coming from the National Research Council of the National Academy of Sciences (“NAS”) or

---

1. Professor of Law, Widener University School of Law. The author thanks the School of Law for a 2013 summer research grant that funded the writing of this article, and colleague Professor John Nivala for his insights. The author also thanks science and law scholars Professors Edwin Imwinkelried, David Kaye, and Jennifer Mnookin for their patient responses to inquiries regarding the Daubert rule and its potential laxity.


4. Testimony offering conclusions unsupported by science are a documented problem. The May 2013 decision of the Mississippi Supreme Court to stay an order of execution was brought about by a United States Department of Justice (DOJ) concession that FBI analysts testified to match conclusions without scientific support. Manning v. State, 112 So. 3d 1082, 1082 (Mins. 2013). The DOJ acknowledged that at Manning’s trial “the microscopic hair comparison analysis testimony or laboratory report presented in this case included additional statements that exceeded the limits of science and was, therefore, invalid.” Letter from John Crabb, Jr., Special Counsel to Deforest R. Allgood, Dist. Attorney’s Office (May 4, 2013), http://www.reprieve.org.uk/static/downloads/139767216-DOJ-Letter-Manning-case_re_hair_analysis2.pdf.
“Academy”) in its 2009 report “Strengthening Forensic Science in the United States: A Path Forward” (the “NAS Report”). There had been earlier criticisms, particularly from academics, questioning whether disciplines such as latent print and handwriting examinations were science and capable of “individualization,” the conclusion that the crime scene print came from one suspect to the exclusion of all other persons in the universe. There was even prompting from within the forensics practitioners communities for scientific research to validate (or confirm the self-perceived validity of) their methodologies and outcomes.

These earlier concerns voiced by academics and forensic practitioners notwithstanding, it was the issuing of the NAS Report that sounded a clarion call—in no small part due to the organization’s prestige. The NAS was established by Abraham Lincoln and has 300 Nobel laureates among its current 2,100 members. The Academy’s prestige is such that “[f]orensic scientists from all disciplines and from every corner of the nation were awaiting the report[…]”

The “clarion call,” sounded by the NAS Report had two facets critical to this article: (1) the report concluded that but for DNA, no forensic discipline had been shown to be scientifically valid in claims of “individualization,” and

---


7. At a conference following the release of the NAS Report, one speaker urged that “lack of validation is not the same thing as admissibility, nor is lack of validation the same thing as invalid.” Kenneth E. Melson et al., Comment to Session #5: The Prosecutors’ Perspective, The Future of Forensic Sciences Symposium, held by Cleveland Marshall Law School (Mar. 19, 2009), http://mediastore.ualh.edu/Mediastore/Play/c64a197e208f4dc649d5b3c47ca9d1821d?catalog=215e4879-3b56-424c-8b37-46be4bec6d8d.

8. Kenneth E. Melson, Embracing the Path Forward: The Journey to Justice Continues, 36 NEW ENG. J. ON CRIM. & CIV. CONFINEMENT 197, 199-203 (2010). The Consortium of Forensic Science Organizations, an umbrella group for the American Academy of Forensic Sciences, American Society of Crime Laboratory Directors, American Society of Crime Lab Directors – Laboratory Accreditation Board, International Association for Identification, National Association of Medical Examiners, and the Society of Forensic Toxicologists – American Board of Forensic Toxicology, was a moving force in convincing Congress to task the NAS to study and report on the state of forensic disciplines in the United States. Id. at 197 & n.2.


10. Melson, supra note 8, at 197.

11. “With the exception of nuclear DNA analysis, however, no forensic method has been rigorously shown to have the capacity to consistently, and with a high degree of certainty,
(2) “the courts have been ‘utterly ineffective’ in addressing this problem.”

This article, accepting the dual premises of the NAS Report, examines (1) whether the judiciary has, post-Report, remedied its diagnosed affliction of being “utterly ineffective,” and (2) whether any other mechanism or entity has compensated for the judiciary’s failing. This article then demonstrates and concludes that, but for one forensic discipline involving impression or pattern evidence—ballistics and firearm identification—the answer for the judiciary is “no.” Even in regard to that discipline, the remediation by courts is modest at best. The failure to re-examine and respond to the question of validation may be attributable to three causes: (1) the relatively lax thresholds of both the Frye and Daubert admissibility tests as applied in criminal
demonstrate a connection between evidence and a specific individual or source.” NAS Report, supra note 5, at 7. A more detailed discussion of this and related findings is set forth in section 1. See infra p. 85.

12. NAS Report, supra note 5, at 109 (citation omitted). See also Joseph Sanders, “Utterly Ineffective”: Do Courts Have a Role in Improving the Quality of Forensic Expert Testimony?, 38 FORDHAM URB. L.J. 547, 549 (2010).

13. These terms, with some overlap, have been defined as follows: “Pattern and impression evidence includes any markings produced when one object comes into contact with another object, such as fingerprints, shoeprints, toolmarks, and tire treads. It also includes pattern analysis, such as is used when evaluating handwriting, typewriting, and writing instruments.” Law Enforcement Standards Office, Pattern and Impression Evidence, NAT’L INST. OF STANDARDS & TECH., http://www.nist.gov/oles/forensics/pattern_impression_evidence.cfm (last visited Dec. 24, 2013).

14. There has been some critical re-examination of non-impression, non-individualization disciplines or claims of expertise, particularly in the subject of shaken baby syndrome. See, e.g., Ex parte Henderson, 384 S.W.3d 833, 833-34 (Tex. Crim. App. 2012) (affirming grant of new trial where advances in biomechanics called into question earlier testimony that death was caused by deliberate shaking of infant). Another re-examined non-impression, non-individualization discipline or claim of expertise is arson causation testimony. See, e.g., Lee v. Glunt, 667 F.3d 397, 407-08 (3d Cir. 2012) (permitting habeas proceeding to go forward based on claim that “new developments in fire science . . . show[] that the fire expert testimony at Lee’s trial was fundamentally unreliable[] . . .”). This article focuses on disciplines that claim the ability to individualize.

15. Frye v. United States, 293 F. 1013 (D.C. Cir. 1923). The key passage defining what came to be the Frye standard is:

Just when a scientific principle or discovery crosses the line between the experimental and demonstrable stages is difficult to define. Somewhere in this twilight zone the evidential force of the principle must be recognized, and while courts will go a long way in admitting expert testimony deduced from a well-recognized scientific principle or discovery, the thing from which the deduction is made must be sufficiently established to have gained general acceptance in the particular field in which it belongs.

Id. at 1014.

16. Daubert v. Merrell Dow Pharm., Inc., 509 U.S. 579 (1993). In its most succinct form, the Daubert test requires that:

[The trial judge must determine at the outset] . . . whether the expert is proposing to testify to (1) scientific knowledge that (2) will assist the trier of fact to understand or determine a fact in issue. This entails a preliminary assessment of whether the reasoning
proceedings; (2) the lack of scientific training and education among the judiciary, corps of prosecutors, and defense counselors who are responsible for understanding and testing the limits of such proof—a problem compounded by often inadequate funding; and (3) a stasis or inertia resulting from decades or more of reliance on these disciplines and their perceived continued utility. To date, there has also been an inadequate response from institutions outside of the judiciary, be they scientific or expert working groups, forensic science commissions, or law enforcement.

Section 1 of this article provides a summary of the NAS Report findings, both in general terms and in regard to specific forensic disciplines. Section 2 examines a pre-NAS Report challenge to forensic evidence, the *Llera Plaza* litigation on the admissibility of latent print testimony. The case is used as a springboard to illustrate the limits of the *Daubert* and *Frye* tests in restricting the scope of forensic testimony, conclusions of individualization, and source attribution. Section 3 summarizes post-NAS Report decisional law in which the Report was cited and considered in challenges to forensic evidence. The third section compellingly demonstrates that the Report has had minimal impact, and that admissibility decisions remain largely consistent with those that preceded the Report’s issuance.

In section 4, this article assesses alternative mechanisms, particularly non-litigation models (but for a brief assessment of what the article will colloquially term “activist judging”), and their utility in improving and refining the presentation of forensic evidence to ensure that it does not exceed in conclusions what current science shows to be valid. The article avers that these mechanisms—technical working groups, state forensic science commissions, and government law enforcement systemic reviews of errors in forensic testimony—have to date brought about minimal change, with one

or methodology underlying the testimony is scientifically valid and of whether that reasoning or methodology properly can be applied to the facts in issue.

*Daubert*, 509 U.S. at 592-93. *Daubert* set out to define the test under the then-extant version of Federal Rule of Evidence 702. *Id.* at 588-89. Subsequent amendments of that Rule sought to synthesize the *Daubert* formulation:

Rule 702. Testimony by Expert Witnesses

A witness who is qualified as an expert by knowledge, skill, experience, training, or education may testify in the form of an opinion or otherwise if:

(a) the expert’s scientific, technical, or other specialized knowledge will help the trier of fact to understand the evidence or to determine a fact in issue;
(b) the testimony is based on sufficient facts or data;
(c) the testimony is the product of reliable principles and methods; and
(d) the expert has reliably applied the principles and methods to the facts of the case.

*Fed. R. Evid.* 702.

critical exception. When a crisis casts doubt on a discipline, to the point that it is in the interest of those in the particular agency or discipline to get ‘ahead of the curve’ in working toward reform, greater change has been effectuated.18

The conclusions of this article are simple. The Frye and Daubert thresholds are inadequate to the task as they tolerate a lack of rigor, and are dependent upon lawyers’ and the judiciary’s understanding of scientific principles. Movement toward a scientifically-validated form of presenting forensic evidence is crisis-responsive and dependent on the pressures such crises generate. What may provide continued momentum for such change is the history of DNA exonerations, the development of innocence projects and their acceptance as credible “players” in the system deserving a voice at the table, and the lurking effect of the NAS Report.

SECTION 1 – THE NAS REPORT

That the NAS Report emerged from a comprehensive assessment of forensic disciplines cannot be questioned. The Committee, which met on eight occasions, heard testimony from numerous experts both from within the various disciplines and across science and law,19 read reports, and conducted independent research.20 While emphasizing the significance and value of forensic evidence in solving crime and exonerating the wrongfully accused and/or convicted,21 the report catalogued the following concerns pertinent to this article:

• “Often there are no standard protocols governing forensic practice in a given discipline.”22
• “[E]ven when protocols are in place[,] . . . they often are vague and not enforced in any meaningful way.”23
• “With the exception of nuclear DNA analysis, . . . no forensic method has been rigorously shown to have the capacity to consistently, and with a high degree of certainty, demonstrate a connection between evidence and a specific individual or source.”24

These findings were not restricted to a general assessment of forensic discipline evidence of the types commonly found in criminal investigations and trials. The Committee assessed several disciplines, and identified weaknesses or limitations in each.

18. See infra Section 4. Section 4 also discusses, briefly, how “activist” judges and courts can make a modest difference when the understanding of science is clear.
19. The Committee heard from “federal agency officials; academics and research scholars; private consultants; federal, state, and local law enforcement officials; scientists; medical examiners; a coroner; crime laboratory officials from the public and private sectors; independent investigators; defense attorneys; forensic science practitioners; and leadership of professional and standard setting organizations. . . .” NAS Report, infra note 5, at 2.
20. Id.
21. Id. at 4-5.
22. Id. at 6.
23. Id.
24. Id. at 7.
For latent print examination, the Report concluded as follows:

- “In the United States, the threshold for making a source identification is deliberately kept subjective, so that the examiner can take into account both the quantity and quality of comparable details. As a result, the outcome of a friction ridge analysis is not necessarily repeatable from examiner to examiner.”

- “[C]laims that these analyses have zero error rates are not scientifically plausible.”

The NAS Report is not alone in this view of fingerprint analysis. As Scottish authorities concluded in 2011, the current model for latent print analysis, ACE-V, “is not specific enough to qualify as a validated method for this type of analysis.”

Ballistics and firearm identification received less detailed treatment in the NAS Report, being addressed in a section on toolmark evidence, but resulted in a similar discussion regarding limits:

- “[E]ven with more training and experience using newer techniques, the decision of the toolmark examiner remains a subjective decision based on unarticulated standards and no statistical foundation for estimation of error rates.”

The Report went on to adopt language from an earlier National Academies study:

- “The validity of the fundamental assumptions of uniqueness and reproducibility of firearms-related toolmarks has not yet been fully demonstrated. . . . A significant amount of research would be needed to scientifically determine the degree to which firearms-related toolmarks are unique or even to quantitatively characterize the probability of uniqueness.”

Handwriting and “questioned document” analysis was described as needing a scientific foundation:

25. NAS Report, supra note 5, at 139.
26. Id. at 142.
28. FINGERPRINT INQUIRY, supra note 27, at ch. 36 § 96.
29. NAS Report, supra note 5, at 150-55.
30. Id. at 153-54.
31. Id. at 154 (quoting DANIEL L. CORK ET AL., BALLISTIC IMAGING 3 (2008)).
Preferring the "Wise Man" to Science

“...The scientific basis for handwriting comparisons needs to be strengthened.”

“Although there has been only limited research to quantify the reliability and replicability of the practices used by trained document examiners, the committee agrees that there may be some value in handwriting analysis.”

In sum, each of these disciplines was deemed wanting in both objective standards for individual examiners to apply when evaluating, comparing, and drawing conclusions; and in having a scientific basis for claims of individualization that excluded all other possible sources of the “match.”

The Report found fault as much with the ultimate end-use consumers of forensic evidence—judges and lawyers—as with the disciplines and their practitioners. It began with a general condemnation:

[T]he existing legal regime—including the rules governing the admissibility of forensic evidence, the applicable standards governing appellate review of trial court decisions, the limitations of the adversary process, and judges and lawyers who often lack the scientific expertise necessary to comprehend and evaluate forensic evidence—is inadequate to the task of curing the documented ills of the forensic science disciplines.

A specific concern and finding included that “the courts continue to rely on forensic evidence without fully understanding and addressing the limitations of different forensic science disciplines.” Additionally:

The judicial system is encumbered by, among other things, judges and lawyers who generally lack the scientific expertise necessary to comprehend and evaluate forensic evidence in an informed manner, trial judges (sitting alone) who must decide evidentiary issues without the benefit of judicial colleagues and often with little time for extensive research and reflection, and the highly deferential nature of the appellate review afforded trial courts’ Daubert rulings.

32. NAS Report, supra note 5, at 166.
33. Id. at 167.
34. These deficiencies were not limited to these fields, which are highlighted here both because of their prominence in criminal investigations and trials and because they are the subject of most of the post-report litigation. Similar (if not more substantial) concerns were raised regarding other disciplines that produce “match” conclusions: footwear and tire mark impression evidence (“there is no consensus regarding the number of individual characteristics needed to make a positive identification”), id. at 149; hair comparison (“No scientifically accepted statistics exist about the frequency with which particular characteristics of hair are distributed in the population[,] [and] [t]here appear to be no uniform standards on the number of features on which hairs must agree before an examiner may declare a ‘match.’”), id. at 160; and forensic odontology (“bite mark”) evidence (“the scientific basis is insufficient to conclude that bite mark comparisons can result in a conclusive match”), id. at 175.
35. Id. at 85.
36. Id.
37. Id. at 110.
The Report drew an ultimate conclusion that “[t]he adversarial process relating to the admission and exclusion of scientific evidence is not suited to the task of finding ‘scientific truth.’”\textsuperscript{38} This article now turns to an assessment of that conclusion, first by examining the pre-NAS Report litigation on the admissibility of latent print expert evidence in \textit{United States v. Llera Plaza}, and then by examining post-Report rulings in which the Report was raised as a ground to restrict the nature and scope of forensic evidence testimony.

\textbf{SECTION 2 – THE LLERA PLAZA LITIGATION: A TALE OF TWO LATENT PRINT RULINGS AND THE LIMITED GATEKEEPING UTILITY OF DAUBERT AND FRYE}

\textit{United States v. Llera Plaza} was a capital prosecution of three defendants—Carlos Ivan Llera Plaza, Wilfredo Martinez Acosta, and Victor Rodriguez—for a series of drug-related murders in Puerto Rico and Philadelphia, Pennsylvania.\textsuperscript{39} It was brought in the Eastern District of Pennsylvania. As a federal death penalty-eligible case, it was well-resourced. There were two attorneys per defendant, and an ample budget was available for, \textit{inter alia}, expert assistance and expert witnesses.\textsuperscript{40} It was also tried before one of the most eminent and scholarly jurists of his time, Judge Louis Pollak.\textsuperscript{41}

As one of numerous pre-trial motions, the defense filed a challenge to the admissibility of expert testimony matching crime scene latent prints to the defendants. By agreement with the Government, the challenge would be

\begin{itemize}
  \item[A] Nickel's disqualification
  \item[B] Vacated
  \item[C] The author of this article was one of two counsel assigned to represent Mr. Llera-Plaza, and participated in the latent print challenges.
\end{itemize}

\textsuperscript{38} NAS Report, \textit{supra} note 5, at 12.

\textsuperscript{39} United States v. Llera Plaza, 179 F. Supp. 2d 444, 447 (E.D. Pa. 2001). As it pertains to the admission of latent print evidence, the District Court’s original opinion limiting such evidence was dated January 7, 2002. The January 7 opinion was ultimately withdrawn. United States v. Llera Plaza, 179 F. Supp. 2d 492 (E.D. Pa. 2002) (depublished) [hereinafter \textit{Llera Plaza I}], vacated, United States v. Llera Plaza, 188 F. Supp. 2d 549 (E.D. Pa. 2002) [hereinafter \textit{Llera Plaza II}]. The author of this article was one of two counsel assigned to represent Mr. Llera-Plaza, and participated in the latent print challenges.

\textsuperscript{40} 18 U.S.C. § 3005, as amended by § 60026 of the Federal Death Penalty Act, provides as follows:

\begin{quote}
  Whoever is indicted for treason or other capital crime shall be allowed to make his full defense by counsel; and the court before which the defendant is to be tried, or a judge thereof, shall promptly, upon the defendant’s request, assign 2 such counsel, of whom at least 1 shall be learned in the law applicable to capital cases . . . .
\end{quote}

\textit{Id.} (emphasis added).

\textsuperscript{41} Judge Pollak’s distinguished career included serving as a member of the “brain trust” for Thurgood Marshall in the \textit{Brown v. Board} litigation, his tenure as Dean of Yale Law School and of the Law School of the University of Pennsylvania, and arguing civil rights cases before the United States Supreme Court. Dennis Hevesi, \textit{Louis H. Pollak, Civil Rights Advocate and Federal Judge}, Dies at 89, \textit{N. Y. Times}, May 13, 2012, at A24.
based on testimony from an earlier Eastern District of Pennsylvania prosecution, *United States v. Mitchell*.42

In a January 7, 2002, Opinion that sent shock waves through the Government and the criminal law community,43 Judge Pollak concluded that “no expert witness for any party will be permitted to testify that, in the opinion of the witness, a particular latent print is—or is not—the print of a particular person.”44

The Opinion was all the more remarkable because Judge Pollak first made a determination that the two foundational aspects of latent print identification theory—that each person’s fingerprints are unique and that they are permanent in terms of their distinguishing features—are valid, and indeed so settled beyond dispute to the extent they qualify for judicial notice.45 What proved critical was Judge Pollak’s vigorous reading of the *Daubert* mandate of reliability, which he described as requiring that “the proposed expert’s opinion must actually be based on what Rule 702 terms ‘scientific knowledge.’”46

Applying the *Daubert* criterion of whether the technique “can be and has been tested,”47 Judge Pollak held that: “[E]ven 100 years of ‘adversarial’ testing in court cannot substitute for scientific testing when the proposed expert testimony is presented as scientific in nature. . . . [T]he government had little success in identifying scientific testing that tended to establish the reliability of fingerprint identifications.”48

As to the second criterion, peer review, the critique was even more trenchant:

> It is the case that there are numerous writings that discuss the fingerprint identification techniques employed by fingerprint examiners. But it is not apparent that their publication constitutes “submission to the scrutiny of the scientific community” in the *Daubert* sense. Even those who stand at the top of the fingerprint identification field—people like David Ashbaugh and Stephen Meagher—tend to be skilled professionals who have learned their craft on the


45. Id. at 502.

46. Id. at 503. The opinion went on to detail the core: “‘*Daubert* factors’: (1) whether the technique ‘can be (and has been) tested,’ (2) whether the technique has been ‘subjected to peer review and publication,’ (3) ‘the known or potential rate of error . . . and the existence and maintenance of standards controlling the technique’s operation,’ and (4) ‘general acceptance.’” *Id.* (quoting *Daubert v. Merrell Dow Pharm.*, Inc., 509 U.S. 579, 593-94 (1993)).

47. Id. at 504.

48. Id. at 506. The District Court cited to substantial expert testimony from the *Mitchell* proceedings confirming the dearth of scientific testing. *Id.* at 507-08 (citing to and quoting testimony by forensic scientists David Stoney and James Starrs).
job and without any concomitant advanced academic training. It would thus be a misnomer to call fingerprint examiners a “scientific community” in the Daubert sense.49

For the third Daubert criterion of “error rate,” Judge Pollak accepted as an assumption that the error rate of the methodology of latent print examination and comparison was “zero,” and that the point of focus should be practitioner error rate. Based upon the Mitchell record, he could conclude only that the Government evidence was “(modestly) suggestive of a discernible level of practitioner error.”50 He also found that “no one standard” controlled the “technique’s operation,” nor were there controlling standards which governed either “the matching of a latent print to a known fingerprint”51 or the qualification standards for print examiners.52

Recognizing that “general acceptance” was part of the discarded Frye analysis but still a potential consideration under Daubert and Rule 702, Judge Pollak gave little weight to this fourth criterion because, although generally accepted among latent print examiners, “fingerprint examiners[ ] . . . do not constitute a ‘scientific community’ . . . .”52 He further determined that reliability was a precondition of giving weight to general acceptance. “The failure of fingerprint identifications fully to satisfy the first three Daubert factors militates against heavy reliance on the general acceptance factor.”53

The outcome of this analysis was a partial exclusion of latent print testimony. Identifying common features between crime scene or evidentiary latent prints and those of one or more defendants would be permitted, as “objective” and “descriptive” testimony, proof of indisputable relevance and strong reliability.54 The only restriction was that regarding individualization: there would be no testimony “that a particular latent print is in fact the print of a particular person.”55 That decision would be left to the triers of fact, the jurors.56

There was (and remains) strong academic support for this important but short-lived holding, with it being described as “well-reasoned . . . [and] a correct application of the Daubert and Kumho Tire principles . . . .”57 The Opinion was hailed as one that “fully analyze[s] the scientific validity of fingerprints under the true tenets of the Daubert-Kumho standard . . . .”58 and as

49. Llera Plaza I, 179 F. Supp. 2d at 509 (footnote omitted).
50. Id. at 513.
51. Id. at 513-14.
52. Id. at 515.
53. Id.
54. Id. at 516.
55. Llera Plaza I, 179 F. Supp. 2d. at 516.
56. Id.
an opinion that “clearly reveals[] this does not satisfy Daubert.”\textsuperscript{59} However, it was not so perceived by the Government, which promptly sought and was granted rehearing and the opportunity to present live testimony (with the corresponding right being granted to the defense).\textsuperscript{60}

The new evidentiary hearing took place in February 2002, with an international and highly-credentialed array of witnesses:

[For the Government,] Stephen Meagher, Unit Chief of Latent Print Unit 3 of the Forensic Analysis Section of the FBI Laboratory; and Kenneth O. Smith, Senior Forensic Latent Print Analyst of the U.S. Postal Inspection Service. [For the defense,] Allan Bayle, a London-based consultant on fingerprint identification, with lengthy prior service as a fingerprint examiner at New Scotland Yard; Janine Arvizu, a laboratory quality auditor serving as Senior Technical Consultant at Consolidated Technical Services, Inc., a New Mexico firm; and Dr. Ralph Norman Haber, a psychometrician at Human Factors Consultants, a California firm.\textsuperscript{61}

Included in the testimony was substantial criticism of the proficiency tests\textsuperscript{62} used by the FBI, described by one witness as “not testing their ability. It doesn’t test their expertise. I mean I’ve set these tests to trainees and advanced technicians. And if I gave my experts these tests, they’d fall about laughing.”\textsuperscript{63}

Notwithstanding this and similar testimony, Judge Pollak reversed course with a new guiding principle. “In adjusting the focus of inquiry from ACE-V's status as a ‘scientific’ discipline to its status as a ‘technical’ discipline, one modifies the angle of doctrinal vision.”\textsuperscript{64} Judge Pollak added that “the test of reliability is ‘flexible[ ]’ . . . .”\textsuperscript{65} Proceeding from these premises, he elevated the “general acceptance” by practitioners within the latent print community to a validation of the technique’s reliability.\textsuperscript{66} As to error rate, Judge Pollak transformed the inquiry into one that looked at the absence of proof of erroneous identifications by latent print examiners, and concluded from this negative finding that “there is no evidence that the error rate of certified FBI

---


\textsuperscript{60} United States v. Llera Plaza, 188 F. Supp. 2d 549, 553 (E.D. Pa. 2002) [hereinafter Llera Plaza II].

\textsuperscript{61} Id. at 554.

\textsuperscript{62} A “proficiency test” is one designed to measure, with casework-like samples, a latent print examiner's capabilities. See SWGFAST: Guidelines for Latent Print Proficiency Testing Programs, 54 J. FORENSIC IDENTIFICATION 353, 353 (2004).

\textsuperscript{63} Llera Plaza II, 188 F. Supp. 2d at 558 (testimony of Allan Bayle).

\textsuperscript{64} Id. at 562.

\textsuperscript{65} Id. (quoting Daubert v. Merrell Dow Pharm., Inc., 509 U.S. 579, 594 (1993)).

\textsuperscript{66} Id. at 563-64.
fingerprint examiners is unacceptably high.”67 As to standards, because it was shown that the FBI method of fingerprint analysis and evaluation corresponded with that in Great Britain, “the standards which control the opining of a competent fingerprint examiner are sufficiently widely agreed upon to satisfy Daubert’s requirements.”68

The transformation was completed in the Judge’s concluding statement: “[b]ased on the foregoing considerations, I have concluded that arrangements which, subject to careful trial court oversight, are felt to be sufficiently reliable in England, ought likewise to be found sufficiently reliable in the federal courts of the United States, subject to similar measures of trial court oversight.”69 That “careful trial court oversight” would assess two factors: (1) the expertise of the testifying examiner; and (2) whether the latent print at issue had sufficient clarity to permit an evaluation and comparison.70

What can be said of this sea change in a several week period? Judge Pollak properly concluded that latent print analysis was a technique and not a science, but arguably ignored the dilemma that the opinion testimony he was permitting was essentially scientific—a statement of individualization that can only be validated scientifically. That type of source attribution proof is akin to the statistical evidence in a DNA case. However, it is one made without the scientific population studies and stated at a higher degree of likelihood (“this print came from this individual to the exclusion of all others on the planet”) than traditional “random match probability” testimony (“the likelihood that a person picked at random from the population would have this same profile is 1 in xxxxx”) used when there is biological evidence that permits the capture of a DNA profile.71 As David Kaye noted shortly after the second Llera Plaza decision, what is lacking is “research into the validity of the claims of the examiners who testify that a defendant is the only person on earth whose hand could have been the source of the partial fingerprint found at the crime scene.”72

Yet Judge Pollak may not have erred in his application of Daubert. The Daubert “gate-keeping” function is one of questionable stringency, ultimately described by the Court as a determination of “evidentiary reliability.”73 More particularly, the Court defined “the requirement that an expert’s testimony pertain to ‘scientific knowledge’ [as] establishing a standard of evidentiary

67. Llera Plaza II, 188 F. Supp. 2d at 566.
68. Id. at 571.
69. Id. at 576.
70. Id.
72. Kaye, supra note 59, at 1087. Kaye also emphasized that the FBI uses “scientific sounding names for the process of comparing latent and rolled prints—terms like ‘ridgeology’ and ACE-V . . . .” Id. at 1080.
reliability.” While some scholars have argued that this is intended to set a high threshold, there is substantial support for a contrary assessment—one permitting loose hands on the admissibility reins. As described by Professor Robinson:

Law is most effective in guiding judicial behavior when the law has a relatively clear rule, a relatively clear substantive meaning, or where judges face meaningful appellate oversight. 

Daubert decisions fit none of these criteria. As a decision rule, 

Daubert does not dictate how an admissibility determination should be carried out. The opinion easily supports contrasting views on the appropriate stringency of the courts towards admitting borderline expert testimony; it also supports contrasting views on the epistemology of science that judges should adopt.

The potential laxity of this “evidentiary reliability” standard comes, first, from the use of “pertains” as the linkage between reliability and scientific knowledge. The Court’s elaborative footnote makes this more clear:

We note that scientists typically distinguish between "validity" (does the principle support what it purports to show?) and "reliability" (does application of the principle produce consistent results?). . . . [O]ur reference here is to evidentiary reliability—that is, trustworthiness. Cf., e.g., Advisory Committee’s Notes on Fed. Rule Evid. 602, 28 U.S.C. App., p. 755 (the rule requiring

74. Daubert, 509 U.S. at 590.

Daubert as “serv[ing] to affirm . . . more vigorously[ ]” an “exclusionary spirit” and saw the decision as “following the twentieth-century trend towards ever-greater judicial scrutiny of scientific evidence.”


Daubert rulings showed a wide range of variability, “finding some evidence that admissibility standards have tightened in civil cases while remaining unchanged in criminal ones.” Id. at 43 (footnote omitted). This variability was documented by Professor Risinger, who showed that:

As to proffers of asserted expert testimony, civil defendants win their 

Daubert reliability challenges to plaintiffs’ proffers most of the time, and that criminal defendants virtually always lose their reliability challenges to government proffers. And, when civil defendants’ proffers are challenged by plaintiffs, those defendants usually win, but when criminal defendants’ proffers are challenged by the prosecution, the criminal defendants usually lose.


Daubert to challenge and exclude expert testimony).
that a witness who testifies to a fact which can be perceived by the senses must have had an opportunity to observe, and must have actually observed the fact is a ‘most pervasive manifestation’ of the common law insistence upon ‘the most reliable sources of information’” (citation omitted)); Advisory Committee’s Notes on Art. VIII of Rules of Evidence, 28 U.S.C. App., p. 770 (hearsay exceptions will be recognized only “under circumstances supposed to furnish guarantees of trustworthiness”). In a case involving scientific evidence, evidentiary reliability will be based upon scientific validity.77

That the science need only “support” what it “purports” to show confirms the modest threshold for the reliability finding, one that need be established only by a preponderance of the evidence.78 The Court referring to Rule 602 “personal knowledge” is also illustrative, as the Advisory Committee Notes for that Rule make clear that it is apparent knowledge that suffices. “[P]ersonal knowledge is not an absolute but may consist of what the witness thinks he knows from personal perception.”79

The Daubert Court went further in setting a low threshold for scientific validity at the level of evidentiary reliability. It rejected an argument that having judges perform a gate-keeping role will stifle science and be “inimical to the search for truth[.]” by noting that the proper function of the judge will be to exclude “conjecture.”80 The Court championed a broader admissibility standard than Frye’s “general acceptance” test in the face of concerns that trials would be pervaded by “absurd and irrational pseudoscientific assertions[.]” by emphasizing “the capabilities of the jury and of the adversary system generally.”81 “Vigorous cross-examination, presentation of contrary evidence, and careful instruction on the burden of proof are the traditional and appropriate means of attacking shaky but admissible evidence.”82 If “conjecture” and “absurd and irrational pseudoscientific assertions” are the outliers the gatekeeper is to guard against, and to guard against in light of the adversarial trial protections, then the Daubert standard is decidedly lenient and consistent with Judge Pollak’s approach in Llera Plaza II. As Professor Kaye noted, “Kumho Tire arguably permits this outcome—the opinion is quite malleable . . . .”83

Indeed, the Court’s subsequent Rule 702 jurisprudence shows the malleability and laxity of the Daubert standard. In General Electric v. Joiner, the Court confirmed that “the Federal Rules of Evidence allow district courts to admit a somewhat broader range of scientific testimony than would have been admissible under Frye[.] . . .”84 In doing so, the court emphasized the propriety of excluding evidence where it is only “opinion evidence that is connected to

77. Daubert, 509 U.S. at 590 n.9.
78. Id. at 592 n.10.
79. Fed. R. Evid. 602 advisory committee’s note (emphasis added).
81. Id. at 595-96.
82. Id. at 596.
83. Kaye, supra note 59, at 1087.
existing data only by the *ipse dixit* of the expert."³⁸⁵ No language in *Joiner* approved, or intimated approval of, excluding evidence slightly more satisfactory or “reliable” than that born of *ipse dixit* reasoning. The third case in the Rule 702 trilogy, *Kumho Tire Co. v. Carmichael*, emphasized that with either scientific or technical expert evidence, “the law grants a district court the same broad latitude when it decides how to determine reliability as it enjoys in respect to its ultimate reliability determination.”³⁸⁶ While *Kumho* affirmed the exclusion of evidence, its broad discretion equally embraces the decision to permit expert testimony.

Reading the Court’s admissibility threshold as a lenient one is compatible with its treatment of eyewitness identification testimony and the Due Process reliability standard, which described the admissibility issue as one of “evidentiary interest”³⁸⁷ that requires only a level of reliability from which a factfinder might conclude that it is correct. Pre-dating *Daubert* by two decades, the Court used strikingly similar language in extolling the ability of the adversarial trial process to judge weak evidence. “We are content to rely upon the good sense and judgment of American juries, for evidence with some element of untrustworthiness is customary grist for the jury mill. Juries are not so susceptible that they cannot measure intelligently the weight of identification testimony that has some questionable feature.”³⁸⁸

The conclusion that *Daubert* is inherently tolerant in its admissibility threshold is not uniformly held. The Court itself wrote in 2000 of “the exacting standards of reliability [expert] evidence must meet[,]”³⁸⁹ albeit without adding any new criteria or highlighting how District Court judges were to be “exacting” in applying a liberal admissibility test. Also, many scholars have claimed that *Daubert* has stringent standards and instead decried what they consider lax application, but have often done so without extensive discussion of what “evidentiary reliability” means.³⁹⁰ Others acknowledge *Daubert’s*

---

flexibility and permissiveness. Yet without resolving what the Court intended, it is clear that in application there has been a strong preference for admissibility of forensic evidence, particularly from prosecutors. As was noted nearly a decade ago:

A review of federal criminal court cases reveals that 92% of prosecution experts survive defense challenges, while only 33% of defense experts survive challenges by federal prosecutors. A recent study of federal appellate criminal cases found that more than 95% of prosecutors’ experts are admitted at trial, while fewer than 8% of defense experts are allowed to testify.

In light of both the permissiveness inherent in the “abuse of discretion” standard of review for Daubert/Rule 702 admissibility decisions, and the historic record of liberal admissibility of prosecution-offered forensic evidence, it cannot be said that Daubert will be an efficacious preventative against flawed or insufficiently validated forensic evidence.

---

Giannelli, The Supreme Court’s “Criminal” Daubert Cases, 33 SETON HALL L. REV. 1071, 1082 (2003) (emphasizing movement toward “exacting standards” for admissibility); Paul C. Giannelli & Kevin C. McMunigal, Prosecutors, Ethics, and Expert Witnesses, 76 FORDHAM L. REV. 1493, 1531 (2007) (“Daubert v. Merrell Dow Pharmaceuticals, Inc. has been transformed from a case that most courts and commentators believed lowered the barriers to the admissibility of scientific evidence to one that the Court now describes as imposing an ‘exacting’ standard.”) (footnote omitted); Jane Campbell Moriarty, Symposium Foreword, 43 TULSA L. REV. 229, 229-30 (2007) (listing scholars who contend that judges have been lax in their gatekeeping role); Marc T. Treadwell, Evidence, 54 MERcer L. REV. 1487, 1495 (2003) (describing “[t]he exacting and demanding character of the gatekeeper role assigned to district courts by Daubert . . . .”).


92. Joëlle Anne Moreno, What Happens When Dirty Harry Becomes an (Expert) Witness for the Prosecution, 79 TUL. L. REV. 1, 3 (2004) (footnote omitted). See also Peter J. Neufeld, The (Near) Irrelevance of Daubert to Criminal Justice and Some Suggestions for Reform, 95 AM. J. OF PUB. HEALTH 5107, 5110 (2005) (“In the first 7 years after Daubert, there were a mere 211 reported challenges to prosecution experts in state court, and the prosecution defeated the challenge 161 times. From August 1999 through August 2000, there were only 50 reported challenges to admissibility citing Daubert in state criminal cases . . . .”); Sandra Guerra Thompson, Judicial Gatekeeping of Police-Generated Witness Testimony, 102 J. CRIM. L. & CRIMINOLOGY 329, 333 (2012) (“[T]rial courts generally have either been unwilling or unable to perform competent reliability screening in criminal cases.”); NAS Report, supra note 5, at 97 (“[R]eported opinions also indicate that appellate courts routinely deny appeals contesting trial court decisions admitting forensic evidence against criminal defendants.”).

93. There is some support for the proposition that Daubert and the Amendment to Federal Rule of Evidence 702 led to more stringent control over admissibility decisions, and an increase in the exclusion of expert witnesses, in civil cases. See, e.g., Joe S. Cecil, Ten Years of Judicial Gatekeeping Under Daubert, 95 AM. J. OF PUB. HEALTH 574, 574-579 (2005). However, this pattern is not found in criminal cases.

94. As noted by Professor Raeder, the ultimate harm here is inaccuracy in factfinding. “Can interpretation of the rules contribute to jury inaccuracy? Again, the answer is yes, in part,
Although the Daubert Court indicated that its standard was more inclined toward admissibility than the “austere” test developed in Frye,95 the latter has its own vagaries and elasticity. It has been described as the “more lenient” standard,96 perhaps because “where the scientific foundation is weak and general acceptance is high, Daubert excludes while Frye admits.”97 More fundamentally, Frye lacks overall stringency because of its failure to define what constitutes general acceptance in the particular field, and the absence of “standards defining which field to consult . . . leaves the law at the mercy of the practitioners of the respective fields.”98 The malleability of Frye was disparaged by the Third Circuit:

[The vague terms included in the standard have allowed courts to manipulate the parameters of the relevant "scientific community" and the level of agreement needed for "general acceptance." Thus, some courts, when they wish to admit evidence, are able to limit the impact of Frye by narrowing the relevant scientific community to those experts who customarily employ the technique at issue.]99

Finally, Frye is often deemed applicable only to “novel” scientific theories, leaving a gap that precludes re-examination of expert testimony previously deemed admissible but now subject to scrutiny because of questions concerning validity and reliability.100

The doubtful efficacy of the Daubert and Frye standards to ensure that courtroom evidence is limited to scientifically valid propositions is not merely a function of each test’s language; each standard’s laxity is either contributed to or aggravated by additional factors. First among equals is the knowledge because expansive admission of evidence permits jurors to base their opinions on unreliable information leading to unwarranted inferences.” Myrna Raeder, What does Innocence Have to Do With It?: A Commentary on Wrongful Convictions and Rationality, 2003 Mich. St. L. Rev. 1315, 1320 (2003).

95. Daubert v. Merrell Dow Pharm., Inc., 509 U.S. 579, 588-89. The Court emphasized the Frye test’s stringency, describing it as “a rigid general acceptance requirement [that] would be at odds with the liberal thrust of the Federal Rules and their general approach of relaxing the traditional barriers to opinion testimony.” Id. at 588 (internal quotations and citation omitted).


99. United States v. Downing, 753 F.2d 1224, 1236 (3d Cir. 1985). Given that general acceptance remains a permissible factor in a Daubert assessment of admissibility, these criticisms confirm as well the potential lack of stringency under Rule 702.

100. E.g., Commonwealth v. Pukar, 951 A.2d 267, 275 (Pa. 2008) (limiting Frye to novel science). Some jurisdictions have permitted re-examination of previously accepted expert testimony under Frye. Armstead v. State, 673 A.2d 221, 243 (Md. 1996) (“[D]ue process considerations require [trial] courts to intervene if scientific opinion shifts so dramatically that previously accepted methods are considered unreliable[ . . . ]”); see also Ramirez v. State, 810 So. 2d 836, 844 ( Fla. 2001).
deficit of the practitioner and judicial communities. As expressed succinctly by David Faigman in 2006, “[l]awyers, of which judges are merely a subset, generally lack good training in the methods of science. Most lawyers do not speak the language of science.” Faigman was echoing a perception that has been known for decades, following Professor Jonakait’s documented assertion in 1991 that “lawyers and judges usually lack high levels of scientific training and are thus unable to challenge or evaluate science.” The deficit remains today, and was emphasized by the NAS Report: “lawyers and judges often have insufficient training and background in scientific methodology, and they often fail to fully comprehend the approaches employed by different forensic science disciplines and the reliability of forensic science evidence that is offered in trial.” The outcome is that bemoaned by Michael Saks: “[A]s a result of the tradition and education of lawyers and judges, they are poorly prepared to evaluate the claims of forensic science, or any empirical claims. Lawyers and judges come from the world of literary interpretation, not from the world of empirical testing.”


[In 1993, the Carnegie Commission’s Task Force on Judicial and Regulatory Decision Making . . . recommended judicial education on science and technology issues, creation of institutional partnerships between the scientific and legal communities, and establishment of a council for science and justice consisting of lawyers and scientists working together to monitor the development of high-tech litigation.

Patricia E. Lin, Opening the Gates to Scientific Evidence in Toxic Exposure Cases: Medical Monitoring and Daubert, 17 REV. LITIG. 551, 587 (1998). See also Peter J. Neufeld & Neville Colman, When Science Takes the Witness Stand, SCI. AM., May 1990, at 46, 49 (detailing lawyers’ inadequate scientific backgrounds); ANDRE A. MOENSSENS ET AL., SCIENTIFIC EVIDENCE IN CRIMINAL CASES 7 (3d ed. 1986) (“[L]awyers as a group evidence an appalling degree of scientific illiteracy, which ill equips them to educate and guide the bench in its decisions on admissibility of evidence proffered through expert witnesses.”) (footnote omitted).

103. Jessica D. Gabel, Forensiphilia: Is Public Fascination with Forensic Science a Love Affair or Fatal Attraction?, 36 NEW ENG. J. ON CRIM. & CIV. CONFINEMENT 233, 250 (2010) (“[L]awyers offering and opposing such evidence must be equipped to address the realities and the shortcomings of the evidence. Unfortunately, most are either ill-prepared to do so or unaware of the burden placed on them.”).

104. NAS Report, supra note 5, at 27.
105. Michael J. Saks, Banishing Ipse Dixit: The Impact of Kumho Tire on Forensic Identification Science, 57 WASH. & LEE L. REV. 879, 888 (2000). Even when a rigorous inquiry is conducted, this lack of scientific education on the part of the judge and lawyers may lead to a poor or inadequate framing of what the admissibility inquiry should be focused on. See Denbeaux & Risinger, supra note 59, at 66-74 (explaining how the inquiry in Llera-Plaza concentrated too much on the global question of admitting latent print testimony and missed the more critical inquiry of where are the scientific limits on accuracy when the particular print at issue is of limited size and quality). See also Erica Beecher-Monas, The Heuristics of Intellectual Due Process: A Primer for Triers of Science, 75 N.Y.U. L. REV. 1563, 1568, 1570 (2000) (questioning courts’ ability to comprehend genuine scientific inquiry).
Education, alone, however, does not explain the NAS Report finding that: “Unlike the extremely well-litigated civil challenges, the criminal defendant's challenge is usually perfunctory. Even when the most vulnerable forensic sciences—hair microscopy, bite marks, and handwriting—are attacked, the courts routinely affirm admissibility citing earlier decisions rather than facts established at a hearing.”

The contrast between civil and criminal challenges must be attributed, in large part, to the limited funding available for the latter. The overwhelming majority of criminal cases involve indigent defendants; a “study of the 100 most populous counties in the United States found that 82 percent of indigent clients were handled by public defenders, 15 percent by assigned counsel attorneys, and 3 percent by contract attorneys.” That services and funding do not match this need cannot be disputed. “The vast majority of criminal defendants in the United States are too poor to afford a lawyer, yet adequate funding and resources for defense counsel remains an elusive goal.”

The underfunding problem is magnified by excessive caseload. A 2007 study determined that “15 of 19 state public defender systems reported caseloads above national standards. . . . Between 1999 and 2000, public defender caseloads increased by 20%, while staffing increased by only four percent. And 73% of county public defenders exceeded national caseload standards.”

The underfunding and excessive caseloads are attributes of institutional providers of indigent defense, which at least are staffed with lawyers with a greater proficiency in criminal law. A compounding factor for criminal defense challenges to forensic evidence arises from the use of individual

106. NAS Report, supra note 5, at 107 (quoting Neufeld, supra note 92, at S109-10).
107. As Professor Giannelli wrote:

The notion that expert testimony in criminal and civil cases should be treated differently does not seem, at least to me, to be a remarkable proposition. The issues are very different. Instead of worrying about the "hired gun" phenomenon as in civil litigation, the criminal defense lawyer often lacks money for any "gun."

Giannelli, supra note 90, at 1072 (footnote omitted).
110. Id. at 7. The caseload varied further depending on the type of public defender office:

State-based public defender programs reported receiving a median of 82 felony (non-capital) cases, 217 misdemeanor cases, and two appeals case[s] per full-time equivalent litigating attorney. County-based public defender offices reported receiving an average of 100 felony (non-capital) cases and 146 misdemeanor cases per full-time equivalent litigating attorney.

“assigned counsel” to represent a substantial percentage of the indigent defendant population, as “defendants with assigned counsel receiving outcomes that, on the whole, are less favorable compared to defendants with public defenders or private attorneys.” This poorer performance evinces an even lesser skill set and, again, a reduced access to the funding necessary to effectively prepare and litigate a challenge to forensic evidence.

The “perfunctory” challenges to forensic evidence described by the NAS Report—arising from this combination of inadequate scientific knowledge and training, and insufficient funding—presumably have a reinforcing effect on judicial acceptance of long-relied-upon forensic evidence of longstanding. The weaker or less articulate the challenge, the less reason to reconsider a long history of acceptance. And that long history has its own dynamic of tolerance, creating a stasis or inertia. The perception is that expert testimony, which has been used in court in a seemingly-accurate way for decades, or longer, should not be lightly tossed aside. As one court rationalized:

Used successfully in criminal trials for over 100 years, fingerprint identification analysis has withstood the scrutiny and testing of the adversarial process. Those of a "scientific" bent certainly can take issue with whether the judges and lawyers have the education or training to engage in "scientific" testing and with whether the courtrooms provide the sterile, laboratory-like and objective atmosphere associated with and probably conducive to scientific analysis. Even so, it seems an unreasonable stretch simply to discard this experiential testing as wholly unreliable and to relegate the testifying opinions of all these fingerprint examiners to ipse dixit.112

This language is not an isolated view, but instead reflects a recurring comfort with grandfathering the “tried and true” expertise of pattern and impression evidence.113 It makes even the most diligent and well-resourced challenge to forensic evidence a formidable undertaking.

Thus, even though courts occasionally speak of a heightened gatekeeping function for judges,114 structurally, and as applied, neither Frye nor Daubert places substantial obstacles in the way of admitting non-validated forensic evidence. This limited scrutiny was well-described by Judge Friendly, albeit in a civil litigation context, more than forty years ago.

---


113. Brodin, supra note 96, at 892-93 (discussing and documenting “the prevalence of ‘grandfathering in’ evidence previously admitted in the jurisdiction under less stringent standards[.]”).

114. See, e.g., Sargon Enters., Inc. v. Univ. of S. Cal., 288 P.3d 1237, 1250 (Cal. 2012) (“Under California law, trial courts have a substantial ‘gatekeeping’ responsibility.”).
There is no bright line that divides evidence worthy of consideration by a jury, although subject to heavy counter-attack, from evidence that is not. Especially because of the guaranty of the Seventh Amendment, a federal court must be exceedingly careful not to set the threshold to the jury room too high. Yet it is the jury system itself that requires the common law judge, in his efforts to prevent the jury from being satisfied by matters of slight value, capable of being exaggerated by prejudice and hasty reasoning . . . to exclude matter which does not rise to a clearly sufficient degree of value; something more than a minimum of probative value is required.115

A threshold of “more than a minimum of probative value” is not an exacting one. This article shows next that, notwithstanding the prestige of the NAS and the clarion call of the Report, this unexacting threshold remains the norm and unvalidated claims of individualization continue to be admitted almost entirely without exception.  

SECTION 3 – POST-NAS DECISIONAL LAW116

As of April 30, 2013, sixty-five reported decisions cited to, and in many instances discussed, the NAS Report.117 Notwithstanding the Report’s conclusion that judges “continue to rely on forensic evidence without fully understanding and addressing the limitations of different forensic science disciplines[,]”118 courts have overwhelmingly declined to revisit admissibility determinations or circumscribe the proposed testimony in pattern and impression evidence cases where the expert testifies to source attribution. Despite repeated challenges to the admission of latent print expert opinion evidence, courts continue to admit this testimony in an unqualified fashion. The trial in United States v. Watkins,119 predated the issuance of the Report. Nonetheless, the appellate court concluded that the admissibility decision was proper, even in light of the Report’s conclusions and the fact that the expert testified that the error rate for this discipline is “zero.”120 The court explained that, “assuming arguendo that the ACE-V method is not error-free, the fact that the fingerprint examiner testified that it was 100% accurate does not by itself mean that the district court erred in determining that the ACE-V method

117. LEXIS search parameters “strength! w/5 forensic w/5 science w/12 path and date aft 1/1/2009” conducted April 30, 2013, in “federal and state cases, combined.” Occasionally the Report was cited regarding issues unrelated to the subject of this article, such as whether lab reports constitute testimonial hearsay. See, e.g., Commonwealth v. Greineder, 984 N.E.2d 804, 808 (Mass. 2013).
118. NAS Report, supra note 5, at 85.
119. 450 F. App’x 511, 515 (6th Cir. 2011).
120. Watkins, 450 F. App’x at 513, 515.
This unqualified acceptance of latent prints was also affirmed by Federal District Courts in Georgia, Michigan, California, and Virginia. The same response has followed in Frye jurisdictions.

The Virginia decision deserves examination because of the proof offered to support the challenge. The defense presented testimony from Professor Jennifer Mnookin, a law professor and scholar in science and the law. As described by the District Court:

Her research focuses on scientific evidence, particularly forensic science evidence and the validation of the scientific processes underlying forensic science evidence. She is currently a member of a working group of the National Institute of Justice and the National Institute of Standards and Technology that studies the effect of human factors on friction ridge analysis.

Professor Mnookin testified that “ACE-V does not describe a scientific process . . . ACE-V lacks objective criteria to determine whether a print is fit for analysis and comparison and, therefore, relies almost entirely on the experience and intuition of the examiner.” The District Court acknowledged that “Dr. Mnookin's criticisms of ACE-V's scientific validity track the conclusions offered by the NRC report.” Nevertheless, the testimony was ruled admissible: “[E]ven widening the circle to include commentators such as Dr. Mnookin and the members of the committee that drafted the NRC report barely changes the balance of opinion. Adding those commentators to the equation does not outweigh the acceptance friction ridge analysis has gained . . . .”

---

121. Watkins, 450 F. App’x at 515.
123. People v. Luna, 989 N.E.2d 655, 673 (Ill. App. Ct. 2013) (“[U]nder the Frye framework in Illinois, various critiques that defendant highlights from the NRC Report go to the weight of the evidence, not to its admissibility . . . .”); State v. Dixon, 822 N.W.2d 664, 674, 676 (Minn. Ct. App. 2012) (same, and approving expert testifying to a conclusion “to a reasonable scientific certainty[ ]”).
125. Id. at 1010.
126. Id.
127. Id. at 1011. The resistance to admissibility challenges has been extended to the exclusion at trial of expert testimony explaining the limits of the current state of knowledge in the field of latent print analysis. In People v. Gonzalez, E052000, 2012 Cal. App. Unpub. LEXIS 1294 (Cal. Ct. App. Feb. 22, 2012), the defense sought leave to present Professor Simon Cole, whose doctorate is in science and technology studies, to testify that fingerprint evidence is not reliable.

Cole’s belief that fingerprint evidence is not reliable is based upon several factors. The first is that he has searched for but found no studies validating the accuracy of fingerprint identification. In addition, he noted that there are no accepted standards for determining a match. The second is the conclusions of the prepublication version of a
While the challengers in the cited federal cases sought outright exclusion of the latent print testimony, it is unclear whether more limited challenges, such as to examiners reporting an individualization to the exclusion of all others or barring testimony claiming a zero error rate, would have resulted in a different outcome—particularly in courts that view the concerns as more appropriate for cross-examination. Overall, the Report has had virtually no impact on trial court acceptance of latent print evidence.

Essentially the same is true with questioned document (handwriting) evidence. Prior to the issuance of the Report, some federal judges limited the conclusions a handwriting expert could draw. As one court directed: “I find [the expert’s] testimony meets Fed. R. Evid. 702’s requirements to the extent that she restricts her testimony to similarities or dissimilarities between the known exemplars and the robbery note. However, she may not render an ultimate conclusion on who penned the unknown writing.” This response was not uniform, however, as other courts held a contrary view prior to the NAS Report. This accepting view of handwriting expertise has continued unabated even after the Report’s findings. In 2012, and after consideration of the Report, one court approved of the “match” testimony in handwriting (albeit under a Frye “general acceptance” standard), and explicitly “reject[ed] appellant’s argument that the NRC Report represents a scientific consensus as to handwriting identification materially different from that established at the
To date, no reported decision has relied on the NAS Report to restrict a handwriting analyst’s conclusion.133

This categoric approval of impression and pattern evidence is replicated in toolmark identification cases; one court approved testimony by “a tool mark expert that the two sets of [garbage] bags contained markings revealing that they were produced with the same tools and therefore on the same extrusion line.”134 It is only where the particular “expert” has insufficient training that such matching testimony has been disallowed post-Report.135

The only category of pattern and impression evidence where there has been modest judicial scrutiny and restriction has been in the field of firearms identification—the matching of ballistics evidence (a fired projectile or cartridge case) to a particular firearm as the source. The greater stringency of the admissibility standard in this field is in one respect ironic, as there has been significant, rigorous study attempting to show that each firearm does leave unique, identifiable markings.136 At the same time, the explanation may be found in the amorphous and decidedly subjective threshold used by firearms examiners, which allows testimony as to a match when there is “sufficient” correspondence of features.137

134. State v. McGuire, 16 A.3d 411, 429, 437 (N.J. Super Ct. App. Div. 2011) (“[C]oncluding the trial court did not err in admitting the tool mark testimony of [the expert] and his opinion matching the two sets of garbage bags.”). See also Rodriguez v. State, 30 A.3d 764, 768-70 (Del. 2011) approving the admission of tire track and footprint match evidence where the testifying expert was trained primarily in latent print comparison, accepting the trial judge’s assessment that the expert had a “trained eye”).
136. See, e.g., James E. Hamby et al., The Identification of Bullets Fired from 10 Consecutively Rifled 9mm Ruger Pistol Barrels: A Research Project Involving 507 Participants from 20 Countries, 41 ASS’N FIREARM & TOOL MARK EXAM’RS J. 99 (2009). As one study was described:

Bunch and Murphy reported in 2003 on a study in which 10 consecutively manufactured Glock semiautomatic pistol slides were obtained from the factory in Austria. The manufacturing process of the 10 slides - which contain the breechface - was observed and the slides then used to produce test fired cartridge cases for a comprehensive validity study by examiners in the FBI Laboratory’s Firearms-Toolmark Unit (FTU). Using breechface marks, the examiners were able to correctly identify cartridge cases fired by each of the different slides.

Id. at 103.
Yet even with this non-quantifiable standard of “sufficient,” courts have placed only modest restrictions on firearms identification evidence and rebuffed challenges to its admissibility. As to the latter, one California appellate court upheld a lower court decision which found the NAS Report did not raise sufficient concerns to justify a pre-trial hearing on admissibility.138

In Melcher, however, the trial court did preclude the expert from saying that no other gun in the world could have left the same markings, but permitted testimony that “the chance of another weapon creating the same pattern was so remote as to be ‘practically impossible.’”139

This modest limit on the absoluteness or certainty of the expert’s conclusion is found in some, but not all, of the post-Report firearms evidence cases. One court permitted the expert to opine that the cartridge and bullet fragment found at the crime scene were fired by the suspect’s weapon, but did so in a context where the expert would not be permitted to express his opinions with any degree of certainty.140 A slightly greater restriction was imposed in United States v. Taylor, where ballistics “match” testimony was permitted, but not to a degree of scientific certainty, nor to the absolute or practical exclusion of all other firearms.141 The permissible qualifier was that the opinion was “within a reasonable degree of certainty in the firearms examination field.”142 Some other courts have responded similarly.143

Yet even this type of minimal circumscribing of expert testimony has not been deemed essential. As the New Jersey District Court explained in United States v. Otero:

The Court recognizes, as did the National Research Council in Strengthening Forensic Science in the United States: A Path Forward, that the toolmark identification procedures discussed in this Opinion do indeed involve some degree of subjective analysis and reliance upon the expertise and experience of the examiner. The Court further recognizes, as did the National Research Council’s report, that claims for absolute certainty as to identifications made by practitioners in this area may well be somewhat overblown. The role of this Court, however, is much more limited than determining whether or not the


139. Id. at *30.

140. United States v. Willock, 682 F. Supp. 2d 512, 535-36 (D. Md. 2010). In Willock, however, the limitation regarding degree of certainty was not imposed by the court but agreed to by the parties. Id.

141. United States v. Taylor, 663 F. Supp. 2d 1170, 1180 (D.N.M. 2009) (holding that the proffered expert opinion testimony with respect to firearms identification is admissible, but subject to certain limitations and qualifiers).

142. Id.

143. See, e.g., United States v. Willock, 696 F. Supp. 2d 536, 549 (D. Md. 2010). Cf. Jones v. United States, 27 A.3d 1130, 1139 (D.C. 2011) (“In light of the government’s representation and the growing consensus that firearms examiners should testify only to a reasonable degree of certainty[,] . . . we will assume, without deciding, that such experts should not be permitted to testify that they are 100% certain of a match, to the exclusion of all other firearms.”).
procedures utilized are sufficient to satisfy scientists that the expert opinions are virtually infallible. If that were the requirement, experience-based expert testimony in numerous technical areas would be barred. Such an approach would contravene well-settled precedent on the district court's role in evaluating the admissibility of expert testimony. 144

*Otero* accepts the lax view for gatekeeping under *Daubert* and Rule 702 and shows the low threshold for gatekeeping forensic opinion “match” evidence. It confirms that if some strictures are to be placed on the definitiveness of forensic match evidence or the level of certainty that experts may use when providing opinion evidence, a mechanism other than the criminal trial process will be necessary. This article next surveys some possible alternatives.

**SECTION 4 – ALTERNATIVES TO ‘CASE-AND-CONTROVERSY’ LITIGATION**

Several mechanisms, other than *Frye/Daubert* challenges in the context of a specific prosecution, have at least the potential for ensuring that forensic evidence is not overstated, but instead presented within the limits of its current knowledge and scientific validity. These modalities—technical working groups, state forensic science commissions, government law enforcement systemic reviews of errors in forensic testimony, and judicial forays into policy and procedure—avoid the tension of adversarial litigation. However, their responsiveness to concerns about accuracy in forensic evidence presentations is often crisis driven. These mechanisms are often created or respond reactively when events cast serious doubt on a category of forensic evidence to a point that it is in the interest of the stakeholders and practitioners to oversee the particular forensic discipline and cabin its efforts.

In one sense the most powerful tool, yet one exercised only rarely, is that of the judiciary—the ultimate consumer of forensic evidence—in setting policy and practice in forensic cases. This occurs when a court, either through rule-making power or in the context of a single case or appeal, establishes a protocol for all future cases involving a specific category of evidence. To date, only two prominent examples responding to the NAS Report can be identified.

The more far-reaching, yet simultaneously limited, move was that of United States District Court Judge Nancy Gertner in 2010, establishing a protocol for all lawyers presenting, or confronted with, “trace” evidence testimony in a criminal case. Judge Gertner established a protocol, albeit one limited to cases in which she was presiding judge, that lawyers:

- “identify whether or not they seek to introduce trace evidence;
- state whether or not either party seeks a *Daubert/Kumho* hearing prior to trial; and,

---

• state the witnesses required for the Daubert/Kumho hearing and the exhibits that the parties seek to admit.”

Judge Gertner’s Order also directed court-appointed counsel to notify the court well before trial “whether expert funds are sought to deal with the trace evidence[ . . . ].”

The premise for this order was drawn directly from the NAS Report.

The NAS report suggests a different calculus—that admissibility of such evidence ought not to be presumed; that it has to be carefully examined in each case, and tested in the light of the NAS concerns, the concerns of Daubert/Kumho case law, and Rule 702 of the Federal Rules of Evidence. This order is entered to accomplish that end.

Nonetheless, although astonishing in its forthrightness and creation of what is essentially an invitation for opponents of such evidence to litigate challenges, the protocol has not resulted in any reported challenge to forensic evidence. Judge Gertner retired from the bench in 2011, and she personally acknowledged the limited efficacy of her directive:

But the order is not enough. Counsel have to learn that advocacy in cases involving forensic evidence requires familiarity with the kind of issues the NAS Report raised. And further, courts need to make it clear that such familiarity may be one of the benchmarks in evaluating when assistance of counsel is constitutionally ineffective. To be sure, cases that do so are the exception and not the rule.

More subdued, but potentially more consequential, is the decision of the Supreme Judicial Court of Massachusetts in Commonwealth v. Pytou Heang. While rejecting a challenge to a murder conviction and, more particularly, rejecting a challenge to “expert forensic ballistics testimony identifying a particular firearm as the one used in the shootings[,]” the Massachusetts high court elected to “offer the following guidelines to ensure that expert forensic ballistics testimony appropriately assists the jury in finding the facts but does not mislead by reaching beyond its scientific grasp.” Those “guidelines,” actually in the form of mandates, require that in all future ballistics evidence cases the examiner must:

---

146. Id.
147. Id. at 3.
150. 942 N.E.2d 927 (Mass. 2011).
151. Id. at 932.
152. Id. at 944.
• document the findings or observations underlying the ultimate opinion;
• provide the documentation during discovery;
• “explain to the jury the theories and methodologies underlying the field of forensic ballistics[ ]” before an opinion is tendered; and
• when “matching” ballistics evidence to a particular firearm, limit the conclusion to one based upon a “reasonable degree of ballistic certainty.”153

Although not a part of the protocol established for such cases, the court emphasized that “[t]he lack of a firm scientific basis proving the uniqueness of individual characteristic toolmarks is a proper subject for both direct examination and cross-examination.”154

The protocol established for Massachusetts is, by its terms, applicable to one forensic discipline, and thus offers no assurance of strictures in other categories of cases. It also barely circumscribes such testimony, limiting it only in terms of the degree of certainty the examiner may utilize. Thus, while the Pytou Heang protocol does much to regularize the mechanics of presenting forensic evidence, and ensure greater amounts of information for both factfinders and lawyers, it leaves largely intact the right of the forensic firearms community to testify as it always has—to an actual match. This resolution is incompatible with the court’s acknowledgment of the NAS Report and the fact that “there is little scientific proof supporting the theory that each firearm imparts ‘unique’ individual characteristic toolmarks onto projectiles and cartridge cases.”155 This caveat—that the match is only to a “reasonable degree of ballistic certainty”—is likely to fall on deaf ears, as a jury of non-scientists will have no basis for quantifying what “ballistics certainty” is, or is not. Judicial protocols, then, have not proved to be an effective or rigorous mechanism for ensuring that forensic evidence be restricted to its scientifically-validated conclusions—even when courts are willing to impose them.156

State forensic science commissions have come into being in the recent past; thirteen states have created permanent forensic science oversight bodies by statute, and two other states have had such agencies established by their attorneys general.157 Yet these commissions are unlikely vehicles for change in the presentation of forensic testimony in court, as their primary focus is on lab performance, with a secondary focus on exposing or responding to fraud. The statutory mandate for New York’s commission is illustrative:

153. Pytou Heang, 942 N.E.2d at 944-945.
154. Id. at 945 n.28.
155. Id. at 941.
156. The Massachusetts high court stands out for its willingness to consider and impose protocols, having also done so in cases involving eyewitness identification. Commonwealth v. Silva-Santiago, 906 N.E.2d 299, 312 (Mass. 2009) (proposing a protocol to be used by police when showing a witness a photographic array).
The commission shall develop minimum standards and a program of accreditation for all forensic laboratories in New York state, including establishing minimum qualifications for forensic laboratory directors and such other personnel as the commission may determine to be necessary and appropriate, and approval of forensic laboratories for the performance of specific forensic methodologies.\footnote{158}

In only one reported instance has a forensic science commission even acknowledged the concerns of the NAS Report and the limits of some disciplines in match evidence. A June 2012 meeting of stakeholders in forensic evidence issues, organized by the Texas Forensic Science Commission, resulted in a report that discussed “[p]seudo/junk-science,” defined as “‘science’ introduced as evidence with a lack of adequate underlying research, poor documentation of testing, no repeatable results, no manner of replicating testing, little or insufficient peer review, and an ‘individualized’ approach to analysis.”\footnote{159} The category also includes cases in which scientific principles are overstated in testimony beyond the bounds of scientific integrity, resulting in communication of materially misleading information to a trier of fact. The Texas commission’s report also applied this term to instances where forensic practitioners “overstated in testimony beyond the bounds of scientific integrity, resulting in communication of materially misleading information to a trier of fact.”\footnote{160}

Having identified concerns, the stakeholders reached no resolution. Instead, they focused on the incapacity of current institutions to address such problems:

There are many limitations in the current adversarial process that make it challenging to identify possible pseudo/junk science. First, defense lawyers are not always competent enough to raise the issues. Second, judges are sometimes reluctant to exclude evidence, and they make poor calls on reliance and reliability. Third, there can be legal precedents in appellate court decisions directly impacting a lower court’s ability to act in pseudo/junk science cases.\footnote{161}


\footnotetext{160}{\textit{Id.}\textit{}}

\footnotetext{161}{\textit{Id.} at 17.\textit{}}
Thus, at most, a science commission might provide a forum for raising issues concerning the limitations of forensic disciplines. The limited authority conferred on such bodies makes further action unlikely, if not off limits entirely.

Scientific and technical or expert working groups (SWGs) are of two types: those with a permanent status that seek to set guidelines and standards within a discipline,162 and those organized for a particular task such as writing a report163 or creating an educational document.164 One limitation of the “permanent” groups is membership—it may be drawn entirely from the practitioner community,165 thereby excluding academic and other critics and/or consumers of the discipline’s output. This insularity brings with it the risk of narrowness in perspective and a diminished propensity to adjust, absent a strong external imperative.166 The permanent groups are self-regulating as well, and thus not compelled to weigh or respond to evolving science or

162. As explained by the Federal Bureau of Investigation:

Since the early 1990s, the FBI Laboratory has led the way in sponsoring Scientific Working Groups (SWG) to improve discipline practices and build consensus with our federal, state, and local forensic community partners. Currently, the FBI Laboratory sponsors eight SWGs:

- SWGDAM—DNA Analysis
- SWGDE—Digital Evidence
- SWGDOC—Questioned Documents
- SWGFST—Latent Fingerprints
- SWGGUN—Firearms and Toolmarks
- SWGGIBRA—Illicit Business Records
- SWGGIT—Imaging Technologies
- SWGMAT—Materials


163. E.g., HUMAN FACTORS, supra note 27, at vii-viii.


165. The members of SWGFST, the working group on latent prints, are almost exclusively practitioners. See SWGFST Members, SWGFST (Mar. 27, 2013), http://www.swgfast.org/Members.htm (listing members, with only two academic researchers). The members of SWGGUN (firearms and toolmarks) are all practitioners. SWGGUN Members, SWGGUN, http://www.swggun.org/swg/index.php?option=com_content&view=article&id=46&Itemid=3 (last visited Jan. 19, 2014). The members of SWGDOC, the forensic document examination group, are all government or private entity examiners. Members, SWGDOC, http://www.swgdoc.org/index.php/about-us/members (last visited Jan. 19, 2014). Albeit scientists, all members of SWGDAM, but one, are employed in local, state or federal crime laboratories or medical examiner offices. See SWGDAM Membership 2013, SWGDAM, http://www.swgdam.org/members.html (last visited Jan. 19, 2014).

166. “[I]nsularity is another effect of immersion in a narrow field. People come to see the decisions they make from the perspective of the field in which they work and give little weight to other perspectives.” Lawrence Baum, Probing the Effects of Judicial Specialization, 58 DUKE L.J. 1667, 1677 (2009) (footnote omitted).
law.167 This last point is not intended to diminish the seriousness with which these groups accept their responsibility. Some SWG members have collaborated with academic critics to urge that forensic disciplines be science based.168 However, neither diversity in membership, nor any legal mandate, propels the SWGs to address scientific limitations in their respective fields.169

This shortcoming is apparent in the SWGs’ varying treatment of individualization testimony. The SWG for firearms and toolmarks (SWGGUN) continues to endorse individualization testimony and source attribution based upon “sufficient agreement” between the tool and the marked object.170 The SWG for questioned documents (SWGDOC) also permits source attribution when a document examiner concludes that there is an “identification.”171 The SWG for latent fingerprints (SWGFAST) has

167. The bylaws of SWGGUN, for example, express one purpose of the organization as “[t]o develop a series of consensus guidelines for the firearm and toolmark discipline.” SWGGUN Bylaws, SWGGUN, http://www.swggun.org/swg/index.php?option=com_content&view=article&id=44&Itemid=2 (last visited Jan. 19, 2014). Its objectives include discussing and sharing “ideas regarding forensic analysis methods, protocols and research[ ]” and “monitor[ing] and disseminat[ing] research and technology related to the discipline.” Id.

168. Jennifer L. Mnookin et al., The Need for a Research Culture in the Forensic Sciences, 58 UCLA L. Rev. 725 (2011). Two of the authors of this collaborative piece, Glenn Langenburg and David Stoney, are members of SWGFAST.

169. One SWG, SWGGUN, has taken issue with an earlier NAS Report which stated that “[t]he validity of the fundamental assumptions of uniqueness and reproducibility of firearms-related toolmarks has not yet been fully demonstrated.” NAS Report, supra note 5, at 154. The SWGGUN response to the 2004 report was to defend its “science.” “The NAS Committee’s viewpoint is one that has been the subject of countless Daubert hearings throughout state and federal courts. These hearings have concluded that the science of firearm and toolmark identification is valid and well-established within the scientific community.” SWGGUN NAS Ballistic Study Report Document, SWGGUN, http://www.swggun.org/swg/index.php?option=com_content&view=article&id=38:nas-ballistic-studydocument&catid=13:other&Itemid=23 (last visited Jan. 19, 2014). This defensive posture, affirming that ballistics is a validated science, was repeated after the 2009 NAS Report. “The reliability of the science has been demonstrated and supported through proficiency tests and validity studies over many decades.” SWGGUN Systemic Report, SWGGUN, http://www.swggun.org/swg/index.php?option=com_content&view=article&id=3&Itemid=10 (last visited Jan. 19, 2014).

170. The Theory of Identification as it pertains to the comparison of toolmarks enables opinions of common origin to be made when unique surface contours of two toolmarks are in “sufficient agreement.” Agreement is significant when the agreement in individual characteristics exceeds the best agreement demonstrated between toolmarks known to have been produced by different tools and is consistent with the agreement demonstrated by toolmarks known to have been produced by the same tool. The Foundations of Firearm and Toolmark Identification, SWGGUN, http://www.swggun.org/swg/index.php?option=com_content&view=article&id=66:the-foundations-of-firearm-and-toolmark-identification&catid=13:other&Itemid=43 (last visited Jan. 19, 2014).

171. “The examiner has no reservations whatever, and although prohibited from using the word ‘fact,’ the examiner is certain, based on evidence contained in the handwriting, that the writer of the known material actually wrote the writing in question.” SWGDOC Terminology for Expressing Conclusions of Forensic Document Examiners, SWGDOC 1, http://www.swgdoc.org/images/documents/standards/SWGDOC%20Terminology%20for%2
tempered its source attribution in latent print cases, concluding that “[t]he ability of a latent print examiner to individualize a single latent impression, with the implication that they have definitely excluded all other humans in the world, is not supported by research and was removed from SWGFAST’s definition of individualization.”172 However, latent print examiners are still permitted to attest that “the likelihood the impression was made by another (different) source is so remote that it is considered as a practical impossibility.”173 SWGFAST has proposed only that certain terminology—to the “[e]xclusion of all others” rendered with “100% certainty (as an absolute fact)” by a discipline with “[z]ero error rate” or “infallible method”—should not be used in reports and testimony.174 Only in DNA analysis is there qualified identification testimony, that of the random match probability—a statement of the probability that another person, randomly selected, would have this combination of genetic markers.175 The forensic discipline with the greatest scientific support for its conclusions uses the most tempered language in presenting “match” evidence, and, to the extent that the NAS Report conclusions as to a lack of scientific validation of match conclusions in pattern and impression evidence are correct, scientific working groups have not acknowledged or accepted the same.

One instance where practitioners within a discipline have engaged with academics, critics, and experts from scientific disciplines has been in the expert working group that produced Latent Print Examination and Human Factors: Improving the Practice through a Systems Approach.176 This working group met over a three year period, with the charge to “examine human factors in latent print analysis and to develop recommendations to reduce the risk of error and improve the practice of latent print analysis.”177 Operating in a consensus process,178 the working group concluded the following in terms of “individualization”:

In sum, the existing practice of latent print experts is to provide a firm opinion regarding whether a named individual is the source of a questioned
print (either a universal individualization or an exclusion) or to give no opinion by stating that the unknown impression is insufficient for a definitive comparison or that the comparison is inconclusive). But other ways to describe the possible association include statements about the strength of the evidence (the likelihoods) or the posterior probability . . . . Given the current state of scientific and professional knowledge, however, it is best to avoid testimony based on the theory of global general uniqueness. Recommendation 3.7, that examiners not testify to an identification to the exclusion of all other individuals in the world, requires other, more conservative methods for conveying the probative value of a match. For that reason, this chapter identifies a broad spectrum of alternatives. The Working Group did not reach a consensus on which of these alternatives to universal-individualization testimony is best.179

What can be seen here are both the power and limitation of the expert working group model. Consensus achieved at least a modest breakthrough: a repudiation of testimony claiming “source attribution to the exclusion of all others in the world.”180 However, this language in no way contradicts the International Association for Identification practice of testimony as to “the likelihood the impression was made by another (different) source is so remote that it is considered as a practical impossibility[,]”182 which is essentially the same in the eyes and ears of factfinders, and remains without a clear scientific foundation. Also, at least one expert group member has endorsed the concept of uniqueness (and implicitly the ability to perform source attribution to the exclusion of all others in the world) in a guide to latent print examination published almost contemporaneously with the Human Factors report.183 In sum, the Human Factors report, as a consensus document, offered only modest limitations on expert testimony and has only exhortatory status.

The final principal mechanism for reforming forensic testimony is found in internal law enforcement policy changes. These have historically been crisis-driven, only occurring after years of resistance to challenges or uncritical acceptance of exaggerated or inaccurate claims. Paradigmatic is the FBI response to comparative bullet lead analysis (CBLA), a forensic technique in which analyses were conducted since the 1960s184 “to determine the

179. HUMAN FACTORS, supra note 27, at 138. Recommendation 3.7 reads: “Because empirical evidence and statistical reasoning do not support a source attribution to the exclusion of all other individuals in the world, latent print examiners should not report or testify, directly or by implication, to a source attribution to the exclusion of all others in the world.” Id. at 72.
180. Id.
181. The International Association for Identification is “the oldest and largest forensic association in the world[,]” and its members “assemble[ ] to educate, share, critique and publish methods, techniques and research in the physical forensic science disciplines.” See The International Association for Identification, http://www.theiai.org/ (last visited Jan. 19, 2014).
182. SWGFAST Guideline, supra note 174, at 5.
concentrations of seven elements—arsenic, antimony, tin, copper, bismuth, silver, and cadmium—in the bullet lead alloy of both the crime-scene and suspect’s bullets.”\textsuperscript{185} While the metallurgic analysis of the components was a valid science, the statistical projections and significance were unjustified. The National Research Council (NRC) identified three deficiencies:

- “The available data do not support any statement that a crime bullet came from a particular box of ammunition. In particular, references to “boxes” of ammunition in any form should be avoided as misleading under Federal Rule of Evidence 403.
- Compositional analysis of bullet lead data alone also does not permit any definitive statement concerning the date of bullet manufacture.
- Detailed patterns of the distribution of ammunition are unknown, and as a result, experts should not testify as to the probability that the crime scene bullet came from the defendant. Geographic distribution data on bullets and ammunition are needed before such testimony can be given.”\textsuperscript{186}

What prompted the NRC review was, essentially, an act of whistleblowing by a former FBI lab scientist who published articles and began offering testimony to challenge the statistical evidence offered by the Government.\textsuperscript{187} Coupled with this was the determination that an FBI analyst had lied—under oath—concerning a CBLA analysis.\textsuperscript{188} As a consequence, the FBI requested the NRC assessment.\textsuperscript{189}

The FBI response was significant, albeit at times inconsistent or slow. In 2005, it announced, “after extensive study and consideration, it [would] no longer conduct the examination of bullet lead.”\textsuperscript{190} The Bureau undertook an internal investigation of the 2,500 cases (approximately) where CBLA evidence was used in the investigation, and in cases where such testimony was presented in court, the Bureau sent letters to notify prosecutors that testimony from its experts “exceeds the limits of the science and cannot be supported by the

\begin{flushleft}


\textsuperscript{188} Ragland v. Commonwealth, 191 S.W.3d 569, 580-81 (Ky. 2006).

\textsuperscript{189} NRC FORENSIC ANALYSIS, \textit{supra} note 184, at 2. The Federal Bureau of Investigation (FBI) asked the National Research Council to conduct an impartial scientific assessment of the soundness of the principles underlying CBL, the optimal manner for conducting an examination with CBL, and the scientifically valid conclusions that can be reached with CBL. \textit{Id}.

FBI[] . . . .¹⁹¹ The pace of notification was criticized, particularly because of the failure to ensure defendants received notification and the time limits imposed by many states for claims based upon newly-discovered evidence.¹⁹² Furthermore, the content of the notification was challenged as misleading and incomplete.¹⁹³ Ultimately, however, the FBI entered into a collaborative effort with the National Association of Criminal Defense Lawyers and the Innocence Project to review such cases and ensure notification to defendants.¹⁹⁴ A similar process is now being deployed in cases where


¹⁹³. One such letter, sent by the FBI in 2010, stated that:

Science does not support the statement or inference that bullets, shot pellets, or bullet fragments can be linked to a particular box of bullets. Further, any testimony stating bullets came from the same source of lead is potentially misleading without additional information regarding approximate numbers of other “analytically indistinguishable” bullets that also originated from that same source . . . .

After reviewing the testimony of the FBI’s examiner, it is the opinion of the FBI Laboratory that the examiner properly testified that the examination revealed that the evidentiary specimen(s) probably came from the same melt of lead. However, the reviewers felt that the examiner did not provide sufficient information to the jury to allow them to understand the number of bullets (or shot pellets) made from the melt. Without having evidence concerning the approximate number of the bullets produced from a single melt, the jury could have misunderstood the probative value of this evidence.

examiners testified to hair comparison evidence, testimony since brought into question by DNA exonerations.

However significant the FBI response, it remains subject to criticism for reasons extending beyond its halting, and less than complete remediation, of the problem. It does not stand as proof that internal mechanisms can, and will, catch flawed or over-stated forensic matching testimony. The FBI was aware, as early as 1991, that there was questionable validity to the type of statements its experts were offering in trials. In that year, an internal study cast doubt on the significance of its experts’ testimony in cases where crime scene evidence and bullets seized from a suspect had the same metallurgic composition. Two flaws were identified: (1) the FBI study found that bullets packaged 15 months apart (and thus likely from different ’batches’) had the same composition, calling into question the theory that each batch was unique; and (2) it found that bullets in a single box often had varying compositions—a finding the report noted could have “significant impact on interpretation of results in forensic cases.” Continuing a practice for a dozen years until an insider speaks out and voices the concerns in public repudiates the idea that internal mechanisms are a viable tool to “police” forensic evidence. What prompted internal corrective measures were crises: a well-credentialled insider attack on the purported science, and a massive accumulation of DNA exonerations confirming the invalidity of individualization testimony and the resulting convictions.

CONCLUSION

The proverb quoted at the beginning of this article, “[a] single conversation with a wise man is better than ten years of study,” captures the consequence of how the judiciary and non-judicial bodies have, for the most part, addressed

195. In 2012, the FBI announced that:

The validity of the science of microscopic hair comparison is not at issue; however, based on recent cases, the FBI and Department of Justice are committed to undertaking a review of historical cases that occurred prior to the regular use of mitochondrial DNA testing to ensure that FBI testimony at trial properly reflects the bounds of the underlying science.


196. A review of 137 DNA exoneration cases with available transcripts showed that “[s]ixty-five of the trials examined involved microscopic hair comparison analysis. Of those, 25—or 38%—had invalid hair comparison testimony. Most (18) of these cases involved invalid individualizing claims.” Brandon L. Garrett & Peter J. Neufeld, Invalid Forensic Science Testimony and Wrongful Convictions, 95 VA. L. REV. 1, 47 (2009).

197. Solomon, supra note 192.
non-validated forensic evidence testimony. Without crises—particularly internal whistleblowers and a critical mass of DNA exonerations—a tolerance of prevailing “wise man” testimony is the norm.

Absent a reinvigoration of the Daubert and Frye standards, as applied in criminal cases, or judicial acceptance of the call for heightened scrutiny for categories of evidence that have been shown to result in erroneous or wrongful convictions,\(^{198}\) it seems unlikely that any institution, as currently constituted, will propel the forensic disciplines to constrain their testimony and conclusions in a manner consistent with the determinations of the NAS Report. That there is some possibility of change cannot be denied, particularly as organizations such as Innocence Projects continue to have a “seat at the table” and gain visibility and credibility.\(^{199}\) To be watched are two entities/endeavors. The first is the Subcommittee on Forensic Science (SoFS) of the White House’s National Science and Technology Council, which was tasked in 2012 to “create a White Paper that summarizes the SoFS’s recommended path forward for achieving the NRC report’s goals.”\(^{200}\) The second is the “National Commission on Forensic Science,” an entity intended to “recommend strategies for enhancing quality assurance in forensic science unit[,] . . . [including] developing proposed guidance concerning the intersection of forensic science and the courtroom . . . .”\(^{201}\) Yet the modest language, to develop “proposed guidance,” shows the lack of legal authority the Commission will have. Without another crisis, we may remain in the straits described by Asimov—that of scientific knowledge surpassing, but not being accepted as, societal wisdom.

\(^{198}\)Keith A. Findley, Judicial Gatekeeping of Suspect Evidence: Due Process and Evidentiary Rules in the Age of Innocence, 47 GA. L. REV. 723, 772 (2013) (calling for the United States Supreme Court to “reconceive its due process framework” and/or for courts nationally to “employ the rules of evidence (either as currently drafted or amended) to demand rigorous, empirically based pretrial review of reliability and to construct remedies that courts might actually implement[ ]”).

\(^{199}\)The importance of the Innocence Projects cannot be doubted. In addition to their collaborations with the FBI on hair microscopy and CBLA cases, see \textit{supra} notes 190-94, a member of the Innocence Project was an invited observer/participant in the expert working group that produced the Human Factors report on latent prints. \textit{See Human Factors, supra} note 27, at xi (identifying “Sarah Chu, Innocence Project” as a contributor).


\(^{201}\)Notice of Establishment of the Nat’l Comm’n on Forensic Sci. & Solicitation of Applications for Comm’n Membership, 78 Fed. Reg. 12355 (Feb. 22, 2013). On January 10, 2014, the United States Department of Justice announced the commissioners, who will include “federal, state and local forensic science service providers; research scientists and academics; law enforcement officials; [and] prosecutors, defense attorneys and judges [. . . ]” \textit{U.S. Departments of Justice and Commerce Name Experts to First-ever National Commission on Forensic Science, U.S. DEP’T OF JUSTICE} (Jan. 10, 2014), http://www.justice.gov/opa/pr/2014/January/14-at-029.html. The author of this Article is one of the thirty Commissioners. The opinions expressed herein are in no way meant to represent the opinion of the Commission.