

Duquesne University School of Law

From the Selected Works of Jane Campbell Moriarty

2010

Will History be Servitude?: The NAS Report on Forensic Science and the Rule of the Judiciary

Jane Moriarty, *Duquesne University School of Law*



Available at: https://works.bepress.com/jane_moriarty/16/

WILL HISTORY BE SERVITUDE?: THE NAS REPORT ON FORENSIC SCIENCE AND THE ROLE OF THE JUDICIARY

Jane Campbell Moriarty*

1993

*[The trial judge should undertake] a preliminary assessment of whether the reasoning or methodology underlying the testimony is scientifically valid and of whether that reasoning or methodology properly can be applied to the facts in issue. We are confident that federal judges possess the capacity to undertake this review.*¹

2009

*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.*²

*“[T]he undeniable reality is that the community of forensic science professionals has not done nearly as much as it reasonably could have done to establish either the validity of its approach or the accuracy of its practitioners’ conclusions,”³ and the courts have been “utterly ineffective” in addressing this problem.*⁴

I. INTRODUCTION

For several decades, the prosecution and its witnesses have maintained that despite little research and virtually no standards, they can match a fingerprint,

* © 2010 Jane Campbell Moriarty, Professor of Law & Director of Research and Development, University of Akron School of Law. The author wishes to thank Dan Medwed, Joelle Moreno, Jay Koehler, Frank Yates, Elisa Krackow, David Harris, Jeff Rachlinski, and Adina Schwartz, all of whom provided helpful insights into either law or psychology; Matt Powell for help with tracking down out-of-the-ordinary research; Akron Law students Erick Rigby and Marisa Main for terrific research assistance; The University of Akron School of Law for support for this Article; and the editors of the *Utah Law Review* for their patience and fine work. All errors remain the author’s.

¹ *Daubert v. Merrell Dow Pharm., Inc.*, 509 U.S. 579, 592–93 (1993).

² NATIONAL RESEARCH COUNCIL OF THE NATIONAL ACADEMIES, STRENGTHENING FORENSIC SCIENCE IN THE UNITED STATES: A PATH FORWARD 7 (2009) [hereinafter NAS REPORT].

³ *Id.* at 108–09 (quoting Jennifer L. Mnookin, *The Validity of Latent Fingerprint Identification: Confessions of a Fingerprinting Moderate*, 7 LAW, PROBABILITY & RISK 127, 134 (2008)).

⁴ *Id.* at 109 (quoting Peter J. Neufeld, *The (Near) Irrelevance of Daubert to Criminal Justice and Some Suggestions for Reform*, 95 AM. J. PUB. HEALTH 107, 109 (2005)).

handwriting, bullet and bullet cartridge, hair, dental imprint, footprint, tire track, or even a lip print to its unique source (collectively, “individualization evidence”). Not only can they match it, they claim, they can do so often without any error rate.

In the last few decades, with the help of lawyers and academics, litigants have challenged the underlying reliability of individualization evidence. Scholars in various disciplines have written about the startling state of individualization evidence, including its lack of standards, research, and established error rates, and its failure to rely upon statistical probabilities to estimate the likelihood of a match. Since its inception, the Innocence Project has exonerated more than 250 people, a majority of whose convictions have involved inaccurate or even fraudulent forensic science testimony, including individualization evidence.

Despite the lack of proof that such evidence is scientifically reliable (and continued exculpations), courts have rejected most challenges to individualization evidence and continue to admit such testimony. With every exoneration, proof mounts that forensic science cannot do what it claims to be able to do with the precision alleged. By not requiring minimal standards for the reliability of individualization evidence, courts have allowed the forensic science system to operate without any checks and balances and to convict innocent people in numbers we can only estimate.

In February 2009, the National Academy of Sciences issued its long-awaited and groundbreaking report on the status of forensic science, *Strengthening Forensic Science in the United States: A Path Forward* (“the NAS Report”).⁵ The NAS Report is a scathing indictment of both the state of the forensic science system and judicial rulings on such individualization evidence.

This Article discusses the findings of the NAS Report, relevant cases that predate the report, and some cases decided since the report. It posits that the judiciary, which has created a standard of reliability, has failed to hold prosecutorial expert evidence to that standard. Using examples from history and modern cognitive science explanations, the Article tries to explain why the judiciary has been so unwilling to rigorously examine forensic science evidence and urges the judiciary to rethink its perspective going forward.

While the NAS Report suggests an overhaul of the current system, that overhaul is a contentious idea that may well not occur in the near (or even longer) future. Thus, a current crisis exists that the judiciary must address in its day-to-day decision making. The Article suggests how the judiciary can become a more effective crucible for testing the strength and limitations of forensic science.

II. THE NAS REPORT

In February 2009, the National Academy of Sciences issued its report on the status of forensic science, *Strengthening Forensic Science in the United States: A Path Forward*.⁶ It is a detailed discussion, with each chapter providing a

⁵ *Id.*

⁶ *Id.*

compendium of a separate subject relevant to forensic science. The NAS Report, highlighting the myriad shortcomings and failures of what we call forensic science, seems quite shocked at the current situation, remarking that “[t]he Law’s greatest dilemma in its heavy reliance on forensic evidence, however, concerns the question of whether—and to what extent—there is any *science* in any given ‘forensic science’ discipline.”⁷

In addition to detailing the shortcomings of the scientific evidence, the NAS Report explains how the judicial system has utterly failed in its regulation of forensic science in criminal cases.⁸ Due to the system failure on every level to improve the quality of forensic science, the report calls for a virtually complete overhaul of the forensic science system, suggesting the creation of the National Institute of Forensic Science (NIFS), an independent agency to oversee forensic science in the United States.⁹

By any interpretation, the report is a critique of both the current state of forensics as practiced in the United States and the judiciary’s unwillingness or inability to require minimum standards for forensic evidence. While the NAS Report compliments the forensic science community on the “valuable evidence that has contributed to the successful prosecution and conviction of criminals as well as to the exoneration of innocent people,”¹⁰ it simultaneously cautions that, although forensic science has advanced, it is now clear that “in some cases, substantive information and testimony based on faulty forensic science analyses may have contributed to wrongful convictions.”¹¹ These wrongful convictions, the NAS Report notes, reflect the “potential danger of giving undue weight to evidence and testimony derived from imperfect testing and analysis.”¹² Additionally, “imprecise or exaggerated expert testimony has sometimes contributed to the admission of erroneous or misleading evidence.”¹³

The reasons for the problems inherent in forensic science in the United States, the report explains, are myriad, including the following:

- Disparities in the forensic science system between federal resources and states and in the various standards for the medical examination system;¹⁴
- Lack of mandatory standardization, certification, and accreditation of laboratories, as well as no uniformity in the certification of forensic practitioners;¹⁵

⁷ *Id.* at 87.

⁸ *Id.* at 109 (stating that the courts have been ineffective in addressing the issue of the validity of forensic science).

⁹ *See id.* at 1–33 (explaining the shortcomings of forensic science and providing a summary of sought-for changes in the system).

¹⁰ *Id.* at 4.

¹¹ *Id.*

¹² *Id.*

¹³ *Id.*

¹⁴ *Id.* at 5–6.

¹⁵ *Id.* at 6.

- A wide variability across disciplines with respect to techniques, methodologies, reliability, types and numbers of potential errors, research, general acceptability, and published material;¹⁶ and
- Problems relating to interpretation of forensic science—the questionable conclusion about individualization where a specimen is “matched” to a particular individual or source.¹⁷

The NAS Report emphasizes that “law enforcement officials and the members of society they serve need to be assured that forensic techniques are *reliable*.”¹⁸ To that end, the NAS Report calls for a body of research to establish the limits and measures of performance of the various forensic disciplines, as well as research to address the impact of sources of variability and potential bias.¹⁹ Without this research, the NAS Report insists, “[w]e must limit the risk of having the reliability of certain forensic science methodologies judicially certified before the techniques have been properly studied and their accuracy verified by the forensic science community.”²⁰ Nonetheless, the NAS Report concedes that “some courts appear to be loath to insist on such research as a condition of admitting forensic science evidence in criminal cases,”²¹ apparently believing the forensic science disciplines are currently incapable of offering such validation.²² The NAS Report finds that the judicial approach to forensic disciplines has been “if you can’t meet the standard, we’ll eliminate the standard”—a frightening approach, given the clear concordance between forensic science and actual guilt or innocence.²³

In addition to arson,²⁴ biological evidence, chemical analysis, and the medical examination system in the United States, the NAS Report reviews non-DNA individualization evidence (fingerprints, hair, handwriting, toolmarks, shoeprints and tire tracks, and forensic odontology, among others).²⁵ The report describes what a conclusion of individualization actually requires:

¹⁶ *Id.* at 6–7.

¹⁷ *Id.* at 7–8.

¹⁸ *Id.* at 12.

¹⁹ *See id.*

²⁰ *Id.*

²¹ *Id.*

²² *Id.*

²³ *See id.* at 87 (“The degree of science in a forensic science method may have an important bearing on the reliability of forensic evidence in criminal cases.”). For a searching empirical analysis explaining the relationship between forensic science errors and wrongful convictions, see Brandon L. Garrett & Peter J. Neufeld, *Invalid Forensic Science Testimony and Wrongful Convictions*, 95 VA. L. REV. 1 (2009).

²⁴ The foundations of arson investigation, erroneously believed by many to rest on firm scientific footing, suffers from multiple shortcomings. The NAS Report discusses the “paucity of research” and states that “many of the rules of thumb that are typically assumed to indicate an accelerant was used . . . have shown not to be true.” NAS REPORT, *supra* note 2, at 173.

²⁵ *Id.* at 136–64. *See generally* Jane Campbell Moriarty and Michael Saks, *Forensic Science: Grand Goals, Tragic Flaws, and Judicial Gatekeeping*, 44 JUDGES’ J. 16, 17 (2005) (explaining that “individualization specialties seek to associate an item of evidence

[A] conclusion of individualization implies that the evidence originated from that source, to the exclusion of all possible sources. The determination of uniqueness requires measurement of object attributes, data collected on the population frequency of variation in these attributes, testing of attribute independence, and calculation of the probability that different objects share a common set of observable attributes.²⁶

DNA evidence, mentioned favorably in the report, possesses all these attributes, which are hallmarks of a proper conclusion about the probability of a match.²⁷ By contrast, the remaining categories of individualization evidence, including fingerprint comparison, possess none.²⁸ The report concludes that “no forensic method other than nuclear DNA analysis has been rigorously shown to have the capacity to consistently and with a high degree of certainty support conclusions about . . . ‘matching’ of an unknown item of evidence to a specific known source.”²⁹ The report also notes that the forensic science community has had “little opportunity to pursue or become proficient in the research that is needed to support what it does.”³⁰

Despite the need for such research to be done to validate methodology and underlying presumptions, the prosecution continues to argue such evidence is reliable, and, to date, many courts seem to be agreeing.³¹ When describing the lack of research or scientific scrutiny concerning individualization specialties, the report concludes that “although the precise error rates of these forensic tests are still unknown, comparison of their results with DNA testing in the same cases has revealed that some of these analyses, as currently performed, produce erroneous results.”³² These “erroneous results” caused wrongful convictions in some cases. Although people on both sides of the aisle may debate how important this lack of

found at a crime scene with its unique source, to the exclusion of all others”). For further discussion of the concept of “individualization,” see John I. Thornton & Joseph L. Peterson, *The General Assumptions and Rationale of Forensic Identification*, in MODERN SCIENTIFIC EVIDENCE: THE LAW AND SCIENCE OF EXPERT TESTIMONY §§ 31:9-31:11 (David L. Faigman et al. eds., 2005).

²⁶ *Id.*

²⁷ *See id.* at 41 (stating DNA analysis, “with its well-defined precision and accuracy, has set the bar higher for other forensic science methodologies”).

²⁸ *See id.* at 43–44.

²⁹ *Id.* at 87.

³⁰ *Id.* at 44.

³¹ *See* Part VI, *infra*, discussing post-NAS Report cases. Even since the report was issued, many federal prosecutors have continued to argue the evidence is completely reliable and the NAS Report should not be taken into consideration. *See* Harry T. Edwards, *The National Academy of Sciences Report on Forensic Sciences: What it Means for the Bench and Bar*, presentation at the Superior Court of the District of Columbia, May 6, 2010, at 4.

³² NAS REPORT, *supra* note 2, at 42.

research is, the inescapable conclusion is that without scientific proof of the foundations necessary to legitimately declare a match, much expert testimony is simply a hunch, supported by experiential “observations of countless samples.”³³ There are numerous reasons that judges admit such testimony so readily: they trust the FBI and other forensic scientists; they assume they would have heard on a more global level if forensic science were unreliable; they share an intuitive belief that forensic comparison is valid; they are comforted by other courts’ decisions admitting the evidence (what I term the “generally accepted by other courts” standard);³⁴ they find the analysts’ experiential knowledge convincing; and they believe that the long history of use confers some field reliability to the work.³⁵ Although many of these justifications might seem compelling on an intuitive, “hunch-like” level, further scrutiny reveals some serious cracks in the foundation of such opinions.³⁶

³³ See Pierre Schlag, *Commentary Law and Phrenology*, 110 HARV. L. REV. 877, 880 (1997) (discussing this methodology as used in phrenology). In *Misconvictions, Science and the Ministers of Justice*, I discuss the methodological parallels between phrenology and individualization specialties. 86 NEB. L. REV. 1, 17–19 (2007).

³⁴ It is important to distinguish legal precedent from this “generally accepted by other courts” principle. The former refers to the prior court rules of law; the latter refers to findings of scientific reliability of evidence.

³⁵ Jennifer L. Mnookin, *The Validity of Latent Fingerprint Identification: Confessions of a Fingerprint Moderate*, 7 LAW, PROBABILITY & RISK 127, 134 (2008) (“We do have a roughly 100-year long, extremely informal ‘natural experiment’ as a result of [fingerprinting’s] quite substantial investigative use. Whenever a fingerprint examiner matches a latent print to a suspect, and then, subsequently, independent evidence emerges to tie the suspect to the location where the print was left, this new evidence corroborates the correctness of the fingerprint identification.”). Although this may be true, we are still left with the nagging lack of certainty about the meaning and importance of a match; whether in *this* case the match was correct; and equally significant, how often is the claim of a match not correct? As Professor Mnookin notes, in a 1995 proficiency test, nearly 20 percent of the test takers found a match where none existed. *Id.* at 135. For more on error rates among fingerprint examiners, see Simon A. Cole, *Grandfathering Evidence: Fingerprint Admissibility Rules from Jennings to Llera Plaza and Back Again*, 41 AM. CRIM. L. REV. 1189 (2004) and Simon A. Cole, *More than Zero: Accounting for Error in Latent Print Identifications*, 95 J. CRIM. L. & CRIMINOLOGY 985 (2005).

³⁶ See ROBERT BURTON, ON BEING CERTAIN: BELIEVING YOU ARE RIGHT EVEN WHEN YOU’RE NOT (2008) (explaining that the certainty of a hunch may well be a misplaced concept). Burton explains, “[i]nternal bias and a misplaced *feeling of knowing* routinely overpower and outsmart the intellect.” *Id.* at 149. However, we have “no mechanism for establishing the accuracy of a line of reasoning until it has produced a testable idea.” *Id.* at 151. Judges might well think about this concept when they continue to “feel” as though fingerprints, handwriting, and toolmarks can be accurately “matched.”

III. THE JUDICIAL ROLE IN THE GATEKEEPING OF INDIVIDUALIZATION EVIDENCE

During his confirmation hearing in 2005, then-Judge John Roberts told Congress, “Judges are like umpires. Umpires don’t make the rules; they apply them. The role of an umpire and a judge is critical. They make sure everybody plays by the rules, but it is a limited role.”³⁷ There is some nugget of truth in this partial description of what judges do, although there is long-standing, serious disagreement that judges do not “make law.”³⁸ The comment, however, seems disingenuous in its overly facile, folksy explanation of the role of the judiciary. And not surprisingly, Justice Roberts’s comment sparked much controversy, not the least of which was from Judge Richard Posner, who sniffed dismissively that “[n]o serious person thinks that the rules that judges in our system apply, particularly appellate judges and most particularly the Justices of the U.S. Supreme Court, are given to them the way the rules of baseball are given to umpires. The rules are created by the judges themselves.”³⁹

Proving the accuracy of Judge Posner’s comments, the Supreme Court has created, from virtual whole cloth, the “rules” governing admissibility of expert testimony, in the *Daubert/Joiner/Kumho Tire* trilogy (“trilogy”).⁴⁰ To be admissible, the party seeking to introduce forensic evidence must be able to establish, to the satisfaction of the trial court, that the proposed expert evidence meets the Supreme Court-created standard of reliability, incorporated into the Federal Rules of Evidence in 2000.⁴¹ A substantial minority of courts use versions

³⁷ *Confirmation Hearing on the Nomination of John G. Roberts, Jr. to be Chief Justice of the United States: Hearing Before the S. Comm. on the Judiciary*, 109th Cong. 55 (2005) (statement of John G. Roberts, Jr., nominee to be Chief Justice of the United States); see also Jeffrey Toobin, *No More Mr. Nice Guy: The Supreme Court’s Stealth Hard-liner*, THE NEW YORKER, May 25, 2009, at 42. Professor Toobin, however, disagrees with Justice Roberts’s self-assessment, stating that even more than Justice Scalia, “Roberts has served the interests, and reflected the values, of the contemporary Republican Party.” *Id.* at 44.

³⁸ Frederick Schauer, *Do Cases Make Bad Law?*, 73 U. CHI. L. REV. 883, 887–88 (2006). It is far beyond the scope of this Article to review the long history of jurisprudence concerning the judiciary as lawmakers. But as Professor Schauer quite accurately comments, it is “far too late in the day to deny that judges are often (some would say ‘always’) engaged in the process of making law.” *Id.* at 888 (citing Duncan Kennedy, *Legal Formality*, 2 J. LEGAL STUD. 351, 378 (1973)).

³⁹ Richard A. Posner, *The Role of the Judge in the Twenty-First Century*, 86 B.U. L. REV. 1049, 1051 (2006). Many other scholars have roundly disagreed with the now-Chief Justice’s view of the judiciary, for various and often competing reasons. See Michael P. Allen, *A Limited Defense of (at Least Some of) the Umpire Analogy*, 32 SEATTLE U. L. REV. 525, 526 n.4 (2009).

⁴⁰ See *Kumho Tire Co. v. Carmichael*, 526 U.S. 137 (1999); *Gen. Elec., Co. v. Joiner*, 522 U.S. 136 (1997); *Daubert v. Merrell Dow Pharms., Inc.*, 509 U.S. 579 (1993).

⁴¹ See, e.g., FED. R. EVID. 702 (providing, in pertinent part, that an expert may testify if “(1) the testimony is based upon sufficient facts or data, (2) the testimony is the product

of the so-called “*Frye* general acceptance test,” while the federal courts and a majority of state jurisdictions use some variant of the *Daubert* reliability standard.⁴²

Collectively, these cases (and the amended Federal Rule of Evidence 702) reflect a concern about whether evidence is reliable as used in a given case. “[R]eliability cannot be judged globally, ‘as drafted,’ but only specifically, ‘as applied.’ The emphasis [is] on the judgment of reliability as it applies to the individual case, to the ‘task at hand.’”⁴³

The trilogy governing the admission of expert testimony claims to envision a flexible standard, in which the trial court, as gatekeeper of the evidence, determines whether expert evidence meets a minimal standard of evidentiary reliability.⁴⁴ Thus, in the case of expert evidence in the federal courts, the Supreme Court has created the rules by which the courts and litigants must abide⁴⁵ and requires the inferior courts to serve as the arbiters of whether the evidence

of reliable principles and methods, and (3) the witness has applied the principles and methods reliably to the facts of the case”); accord *Kumho Tire Co.*, 526 U.S. at 141; *Gen. Elec. Co.*, 522 U.S. at 146–47; *Daubert*, 509 U.S. at 589–90 (requiring sufficient proof that the proposed expert evidence is sufficiently reliable to be admitted). Although Federal Rule of Evidence 702 was amended to add reliability requirements, there is much agreement that the rule reflects the development in the common law. Federal Rule of Evidence 702 was “amended in response to *Daubert* . . . and to the many cases applying *Daubert*, including *Kumho Tire*.” FED. R. EVID. 702 advisory committee’s note on 2000 Amendments.

⁴² The “*Frye* general acceptance test” requires proof that the novel scientific evidence is generally accepted in the field to which it belongs. The general acceptance test does not analyze the reliability of the proposed evidence; it asks whether novel scientific evidence has reached the tipping point at which it has become generally accepted by scientists in the field. See *Frye v. United States*, 293 F. 1013, 1014 (D.C. Cir. 1923). Many federal cases substantially incorporate some variant of the *Frye* test in their reliability determination, because it is one of the factors listed by the *Daubert* court. See 509 U.S. at 588–89, 598. For a state-by-state breakdown of which states use the respective tests, see 2 JANE CAMPBELL MORIARTY, PSYCHOLOGICAL AND SCIENTIFIC EVIDENCE IN CRIMINAL TRIALS app. 1A (2008) (providing a state-by-state analysis of admissibility standards) [hereinafter MORIARTY, CRIMINAL TRIALS].

⁴³ D. Michael Risinger, *Defining the “Task At Hand”: Non-Science Forensic Science After Kumho Tire Co. v. Carmichael*, 57 WASH. & LEE L. REV. 767, 773 (2000) (citing *Kumho Tire Co.*, 526 U.S. at 141). This point is important, again, for distinguishing between precedent and decisions about reliability of evidence.

⁴⁴ See *Kumho Tire Co.*, 526 U.S. at 150 (for the principle that the inquiry is “a flexible one”) (citing *Daubert*, 509 U.S. at 594).

⁴⁵ See *Daubert v. Merrell Dow Pharm., Inc.*, 43 F.3d 1311, 1316 (9th Cir. 1995) (On remand, Judge Kozinski, writing for the court, applied the standard created by the Supreme Court with the statement, “[m]indful of our position in the hierarchy of the federal judiciary, we take a deep breath and proceed with this heady task.”). Many states, through their supreme courts’ decisions, also have either adopted or rejected *Daubert*, in whole or in part. See MORIARTY, CRIMINAL TRIALS, *supra* note 42, at app. 1A. Thus, very clearly, courts both federal and state are engaged in rule-making in the area of expert testimony.

complies with those rules in this judicially-created gatekeeping role.⁴⁶ Thus, the judiciary has created an entire universe where none previously existed and has appointed itself the master of such universe.

It seems beyond cavil to hope judges would make sure “everybody plays by the rules” the Supreme Court and other federal courts have taken great pains (and several years) to create, refine, and develop. At a minimum, we expect our judiciary to attempt to level the playing field for all participants and to apply equal standards to competing litigants.

Nonetheless, after reading a multitude of cases involving forensic science evidence and empirically driven studies about what courts have done post-trilogy, it is not at all clear that judges “make sure everybody plays by the rules.” Indeed, research suggests that many judges do *not* require the prosecution to play by the same rules that other litigants in both civil and criminal cases do when the subject is forensic science evidence. These findings are not based solely upon my own impressions,⁴⁷ but rather upon empirical data collected and analyzed by others⁴⁸ and the conclusions of the NAS Report: “[T]he vast majority of the *reported* opinions in criminal cases indicate that trial judges rarely exclude or restrict expert testimony offered by prosecutors; most *reported* opinions also indicate that appellate courts routinely deny appeals contesting trial court decisions admitting forensic evidence against criminal defendants.”⁴⁹

It seems abundantly clear that many courts merely give lip service to pressing the prosecution to meet its burden of proof in *Daubert* hearings; much the way some bartenders mix an extra dry martini, as if waving the bottle of vermouth next to the glass will suffice. In other words, the court will find in favor of the

⁴⁶ *Daubert v. Merrell Dow Pharm., Inc.*, 509 U.S. 579, 589 n.7 (1993) (mentioning that FED. R. EVID. 702 is the source of the judge’s gatekeeping duty).

⁴⁷ During the ten years I worked on my treatise, *MORIARTY, CRIMINAL TRIALS*, *supra* note 42, my impression was that courts were not requiring the prosecution to comply with the trilogy requirements. Of course, that was just my impression. Better evidence comes from the studies cited *infra* at note 48.

⁴⁸ For empirical data establishing that judges permit much prosecutorial expert testimony and simultaneously disallow much of defense expert testimony, see D. Michael Risinger, *Navigating Expert Reliability: Are Criminal Standards of Certainty Being Left on the Dock?*, 64 ALB. L. REV. 99 (2000). Recently, Professor Risinger analyzed all reported decisions on defense challenges to fingerprint evidence and concluded that, overwhelmingly, the courts rejected these challenges—not on the basis of an accurate reliability analysis but for a variety of other, less data-driven reasons, noting that “there is some reason to believe that judges as a group are resistant to rejecting prosecution proffers of expert testimony.” D. Michael Risinger, *Goodbye to All That, or, a Fool’s Errand, by One of the Fools: How I Stopped Worrying About Court Responses to Handwriting Identification (And “Forensic Science” in General) and Learned to Love Misinterpretations of Kumho Tire v. Carmichael*, 43 TULSA L. REV. 447, 473 (2007).

⁴⁹ NAS REPORT, *supra* note 2, at 11 (emphasis in original). Of course, the universe of reported decisions is far smaller than that of all decisions courts make, a point the NAS Report concedes. *Id.* Nonetheless, if the reported cases are at all representative, the discrepancy is noteworthy.

government's evidence as long as the government makes the appropriate motions.⁵⁰

Deirdre Dwyer writes compellingly about the asymmetries in the application of *Daubert* between civil and criminal cases.⁵¹ She considers the normative expectations that either all evidence would be treated equally, or that evidence rules “in a criminal action [would be] geared more towards reducing the risk of an erroneous outcome than they are in a civil action.”⁵² After looking at data-driven studies, however, she concludes, “[i]t would seem that the expert evidence of civil plaintiffs, particularly in toxic tort cases, is subject to greater scrutiny than that of civil defendants, while the expert evidence of criminal prosecutors is subject to less scrutiny than that of criminal defendants, or than that of civil parties.”⁵³ This seems to be an odd outcome, indeed, when civil cases involve only monetary damages, while criminal cases deal with stakes of much greater value.⁵⁴

Courts, including the U.S. Supreme Court, have been willing to engage in detailed, scientific analysis of proposed expert evidence in civil cases.⁵⁵ In many of

⁵⁰ This point has been stated elsewhere: “There is almost no [prosecutorial] expert testimony so threadbare that it will not be admitted if it comes to a criminal proceeding under the banner of forensic science.” Moriarty & Saks, *supra* note 25, at 29.

⁵¹ See Deirdre Dwyer, *(Why) are Civil and Criminal Expert Evidence Different?*, 43 TULSA L. REV. 381, 384–87 (2007).

⁵² *Id.* at 384–85.

⁵³ *Id.* at 383. Dwyer does acknowledge the limitations of the studies, pointing out that “[t]here are significant methodological difficulties with inferring general trial conduct from reported decisions.” *Id.*

⁵⁴ See *id.* at 385 (noting that “[t]here are generally greater consequences, socially and physically, of criminal conviction compared with adverse civil judgment, and so we might expect that the rules of evidence in a criminal action are geared more towards reducing the risk of an erroneous outcome than they are in a civil action”).

⁵⁵ The following cases provide a small sample of courts' in-depth opinions on toxic tort expert evidence: *Gen. Elec. Co. v. Joiner*, 522 U.S. 136, 143–47 (1997) (analyzing admissibility of plaintiffs' proposed expert testimony that exposure to PCBs promoted decedent's lung cancer); *McLain v. Metabolife Int'l, Inc.*, 401 F.3d 1233, 1239–52 (11th Cir. 2005) (reversing jury verdict for plaintiffs because plaintiffs' expert evidence that ingestion of defendant product for weight loss was a likely cause of their respective strokes and heart attack was not reliable); *Norris v. Baxter Healthcare Corp.*, 397 F.3d 878, 883–87 (10th Cir. 2005) (explaining why plaintiffs' experts' opinions that silicone breast implants caused autoimmune diseases was not reliably grounded on existing data); *Moore v. Ashland Chem. Inc.*, 151 F.3d 269, 274–79 (5th Cir. 1998) (upholding trial court's decision to exclude expert evidence on allegations that plaintiff's exposure to toluene caused his respiratory disease); *Daubert v. Merrell Dow Pharm., Inc.*, 43 F.3d 1311, 1316–22 (9th Cir. 1995) (on remand from the U.S. Supreme Court, analyzing whether plaintiffs' claim that Bendectin exposure to fetuses in utero caused their limb malformation); *Henricksen v. Conoco Phillips Co.*, 605 F. Supp. 2d 1142, 1152–79 (E.D. Wash. 2009) (analyzing expert testimony that plaintiff's exposure to benzene caused the development of his myelogenous leukemia); *Amorgianos v. Nat'l R.R. Passenger Corp.*, 137 F. Supp. 2d 147, 160–191, (E.D.N.Y. 2001), *aff'd*, 303 F.3d 256 (2d Cir. 2002) (explaining why experts' opinions that

these cases, the district courts, affirmed by courts of appeal on an abuse of discretion standard, have found such testimony insufficiently reliable.⁵⁶ Indeed, the NAS Report notes the different standards in civil and criminal cases, remarking that “courts have not . . . imposed standards ensuring the application of scientifically valid reasoning and reliable methodology in criminal cases involving *Daubert* questions.”⁵⁷ Furthermore, the report concludes that upon reviewing the reported decisions, “at least in criminal cases, forensic science evidence is not routinely scrutinized pursuant to the standard of reliability enunciated in *Daubert*.”⁵⁸

Since *Daubert v. Merrell Dow Pharmaceuticals*, there have been numerous, substantial challenges to the reliability of many types of forensic science, particularly in the area of individualization evidence.⁵⁹ Federal courts have rejected most defense challenges made to the reliability of fingerprint individualization evidence.⁶⁰ With a handful of exceptions in which trials courts have limited conclusions about handwriting and ballistics,⁶¹ courts have rejected challenges to

exposure to xylene vapor glue fumes caused central nervous system and peripheral nervous system injury was not sufficiently reliable to be admissible).

⁵⁶ See cases cited *supra* note 55. In *Uncertainty and Informed Choice: Unmasking Daubert*, 104 MICH. L. REV. 257, 258–67 (2005), Professors Margaret A. Berger and Aaron D. Twersky argue that toxic tort plaintiffs have a great deal of difficulty, post-*Daubert*, in meeting the courts’ requirements concerning expert testimony.

⁵⁷ NAS REPORT, *supra* note 2, at 96.

⁵⁸ *Id.* at 106.

⁵⁹ See cases mentioned *infra* at notes 60–62.

⁶⁰ See, e.g., *United States v. Mitchell*, 365 F.3d 215, 245–46 (3d Cir. 2004) (allowing presentation of fingerprint comparison but allowing the defense to call “counter-experts” to raise doubt about the analysis); *United States v. Crisp*, 324 F.3d 261, 269–70 (4th Cir. 2003) (basing its decision that fingerprint evidence is reliable, at least in part, on the previous decades of consistent admission of such evidence); *United States v. Havvard*, 260 F.3d 597, 600 (7th Cir. 2001) (holding that fingerprint comparison met *Daubert*’s reliability standard); *United States v. Rogers*, 26 F. App’x 171, 173 (4th Cir. 2001) (allowing fingerprint comparison evidence while seemingly shifting the burden on the defendant to show that the analysis is unreliable); *United States v. Llera Plaza*, 188 F. Supp. 2d 549, 575 (E.D. Pa. 2002) (holding that because fingerprint evidence is “sufficiently reliable for an English court,” it must also be reliable under Rule 702). For an excellent analysis of the shortcomings of fingerprint comparison based upon the twin pillars of uniqueness and individualization, see Simon A. Cole, *Forensics Without Uniqueness, Conclusions Without Individualization: The New Epistemology of Forensic Identification*, 8 LAW, PROBABILITY & RISK 233, 242–46 (2009).

⁶¹ See, e.g., *United States v. Glynn*, 578 F. Supp. 2d 567, 574–75 (S.D.N.Y. 2008) (acknowledging that firearm toolmark comparison “lacks the rigor of science” and “suffers from greater uncertainty than many other kinds of forensic evidence,” yet allows testimony that a firearms match was “more likely than not”); *United States v. Greene*, 405 F. Supp. 2d 104, 124 (D. Mass. 2005) (allowing expert to testify about the points of similarity between a shell casing at the scene and one from defendant’s weapon but disallowing any testimony that the shell casings matched); *United States v. Hines*, 55 F. Supp. 2d 62, 70–71 (D. Mass. 1999) (limiting testimony of handwriting comparison witness to discussing points of

the reliability of those forms of evidence as well.⁶² Microscopic hair comparison, implicated for its role in wrongful convictions in many exonerations,⁶³ has been admitted by both federal and state courts, with seemingly no concern for the potentially substantial rate of error such evidence presents.⁶⁴

Courts do not admit such evidence because it meets the trilogy and FRE 702's standards of reliability. To the contrary, prosecutors and their experts cannot establish the validity of what they claim to be able to do with the precision alleged—a point made quite clear in the NAS Report where it remarks that there is a “notable dearth” of scientific studies to establish the validity of many forensic science methods.⁶⁵ Rather, courts admit such evidence simply because they cannot seem to imagine doing otherwise. “The methods of latent print identification can and have been tested. They have been tested for roughly 100 years. They have

similarity but disallowing conclusion testimony): *United States v. Santillan*, 1999 WL 1201765, 5 (N.D. Cal., Dec. 3, 1999). Post-NAS Report cases are discussed, *infra*, in Part IV.

⁶² See, e.g., *United States v. Adeyi*, 164 F. App'x 944, 946 (2d Cir. 2006) (upholding admission of handwriting analysis relying almost exclusively on precedent); *United States v. Crisp*, 324 F.3d 261, 271–72 (4th Cir. 2003) (affirming decision by the trial court that expert testimony regarding handwriting comparisons and conclusion of authorship were reliable under the *Daubert* standard); *United States v. Battle*, 1999 WL 596966, 4 (10th Cir. Aug. 6, 1999) (upholding the trial court's decision that expert testimony about a conclusion of handwriting comparison met the standard of reliability and relevancy); *United States v. Gonzales*, 90 F.3d 1363, 1371 (8th Cir. 1996) (finding that the district court was within its discretion to admit handwriting comparison evidence). For a full explanation of the history of legal challenges concerning the foundation of handwriting comparison, see Risinger, *Goodbye to All That*, *supra* note 48, at 457–67.

⁶³ INNOCENCE PROJECT, WRONGFUL CONVICTIONS INVOLVING UNVALIDATED OR IMPROPER FORENSIC SCIENCE THAT WERE LATER OVERTURNED THROUGH DNA TESTING (2009), http://www.innocenceproject.org/docs/DNA_Exonerations_Forensic_Science.pdf?phpMyAdmin=52c4ab7ea46t7da4197; Innocence Project, Factors Leading to Wrongful Convictions, http://www.innocenceproject.org/understand/factors_74_chart.php (last visited June 1, 2010).

⁶⁴ See, e.g., *United States v. Santiago*, 156 F. Supp. 2d 145, 151–52 (D.P.R. 2001) (admitting hair comparison analysis after referencing persuasive authority and independent research); *Williamson v. Ward*, 110 F.3d 1508, 1522–23 (10th Cir. 1997) (affirming on other grounds but rejecting the district court's ruling that hair analysis was not admissible); *U.S. v. Matta-Ballesteros*, 71 F.3d 754, 766–67 (9th Cir. 1995) (approving of the admission at trial of expert testimony on hair comparison). For critical commentary on microscopic hair comparison, see the NAS REPORT, *supra* note 2, at 155–61, and Edwards, *supra* note 31, at 6 (stating that “hair comparisons without mitochondrial DNA are highly questionable”).

⁶⁵ NAS REPORT, *supra* note 2, at 8 (“The simple reality is that the interpretation of forensic evidence is not always based on scientific studies to determine its validity. This is a serious problem. Although research has been done in some disciplines, there is a notable dearth of peer-reviewed, published studies establishing the scientific bases and validity of many forensic methods.”).

been tested in adversarial proceedings with the highest possible stakes—liberty and sometimes life.”⁶⁶

In *United States v. Llera Plaza*, Judge Pollack vacated his earlier opinion limiting fingerprint examiners’ testimony to pointing out similarities (rather than testifying as to the conclusion of a match) and decided it was appropriate to permit testimony about conclusions.⁶⁷ “In short,” he begins, “I have changed my mind.”⁶⁸ Relying in part upon testimony admitted in the rehearing he granted the government, he concludes if it is “sufficiently reliable in England,” it should be good enough for U.S. courts.⁶⁹ The *Llera Plaza* court, like many others, was impelled by the long history of fingerprint admission, noting that both English and American courts have admitted fingerprint comparison testimony for nearly a century.⁷⁰

In *United States v. Prime*, the district court stated when allowing conclusions about handwriting, “[w]here a technique has been repeatedly applied and tested by law enforcement and courts for over a century, the court does not believe the absence of scientific data, without more, should sound the death knell for such testimony.”⁷¹ The Court of Appeals for the Ninth Circuit upheld this decision, deciding that as long as the process is generally reliable, any potential error can be brought to the jury’s attention with cross-examination and competing testimony.⁷² While the language the court used is not unusual, what the court found to be “generally reliable” was based upon admitted evidence that the error rate for handwriting comparison was approximately 13 percent!⁷³ Yet the court was not convinced that such evidence might not meet the standard of reliability. One must wonder then at what rate of error *does* the evidence become unreliable?

A most unusual form of individualization evidence was admitted in *State v. Davis*, where the court allowed testimony that the defendant’s lip print matched a

⁶⁶ *United States v. Havvard*, 117 F. Supp. 2d 848, 854 (S.D. Ind. 2000), *aff’d*, 260 F.3d 597 (7th Cir. 2001).

⁶⁷ *United States v. Llera Plaza*, 188 F. Supp. 2d 549, 575–76 (E.D. Pa. 2002), *vacating*, 179 F. Supp. 2d 492 (E.D. Pa. 2002).

⁶⁸ *Id.* at 576.

⁶⁹ *Id.*

⁷⁰ *Id.* at 572.

⁷¹ *United States v. Prime*, 220 F. Supp. 2d 1203, 1210 (W.D. Wash. 2002), *vacated and remanded on other grounds*, 543 U.S. 1101 (2005), *subsequently aff’d on this issue*, 431 F.3d 1147 (9th Cir. 2005).

⁷² *Prime*, 431 F.3d at 1153.

⁷³ *Id.* (noting that in one professor’s studies, experts arrived at the correct conclusions 87 percent of the time). As noted by the *Llera Plaza* court (and echoed by numerous scholars), the proficiency tests are not rigorous: “[T]he proficiency tests are less demanding than they should be. To the extent that this is the case, it would appear that the tests can be of little assistance in providing the test makers with a discriminating measure of the relative competence of the test takers.” 188 F. Supp. 2d at 565. For further information on proficiency tests of fingerprint examiners, see NAS REPORT, *supra* note 2, at 143–44.

lip print at the scene.⁷⁴ Although there was absolutely no foundation for such evidence, the trial court, without the smallest amount of critical inquiry, found such proof met both *Frye* and reliability standards.⁷⁵ The defense lawyer handling this case acted incompetently by not satisfactorily challenging the testimony,⁷⁶ yet both the trial court and the court of appeals in the direct appeal were satisfied with little more than the *ipse dixit* testimony of the government's witnesses.⁷⁷

While this lip print case may be an outlier in terms of the courts' willingness to accept expert evidence in criminal cases without sufficient foundational proof of reliability (or, indeed, even general acceptance), it is in many ways emblematic in its willingness to take prosecutorial experts at their word with no requirement of supporting data.⁷⁸

Bite mark evidence, also known by its more technical name "forensic odontology," has a history of acceptance by courts despite the fact that the science supporting it is dubious. In fact, the theories upon which the field is predicated—that dentition is unique and that marks found upon a victim can be linked unequivocally to the perpetrator—are not supported by current data.⁷⁹ The problems, like the rest of forensic science individualization evidence, are legion. It is unlikely that human dentition is unique; very few teeth are actually used to make a bite mark; and bite marks become distorted on human skin due to the passage of time, thus not maintaining the accuracy of the marks.⁸⁰ Moreover, unlike DNA evidence, which uses databases to generate probabilities that the suspect left his DNA at the crime scene, there is no bite mark database from which probabilities can be generated⁸¹—although such probabilistic testimony has been admitted

⁷⁴ *People v. Davis*, 710 N.E.2d 1251, 1256 (Ill. App. Ct. 1999), *conviction vacated*, No. 94 CF 76, 2006 WL 2641753 (Cir. Ct. Ill. Mar. 7, 2006).

⁷⁵ *See id.* at 1258–59.

⁷⁶ *People v. Davis*, No. 94 CF 76, 2006 WL 2641753 (Cir. Ct. Ill. Mar. 7, 2006) (holding that defendant was denied his constitutional right to effective counsel and overturning the defendant's conviction).

⁷⁷ *See Davis*, 710 N.E.2d at 1258–59.

⁷⁸ On the other hand, this case may not be that much of an outlier because federal courts have permitted much expert testimony in criminal cases that appears to have dicey scientific grounding, including profile testimony about specific types of pedophiles. *See, e.g.*, *U.S. v. Hitt*, 473 F.3d 146, 151, 158–59 (5th Cir. 2006) (upholding admission of expert testimony "to explain the behavior of those accused of sexual offenses"). Federal courts also have admitted allegedly scientific comparisons between photographs of items and the items themselves. *See, e.g.*, *U.S. v. McKreith*, 140 F. App'x 112, 116 (11th Cir. 2005) (allowing expert testimony on the comparison of the defendant's shirt that was seized in his apartment with the shirts depicted in bank surveillance images).

⁷⁹ *See Erica Beecher-Monas, Reality Bites: The Illusion of Science in Bite-Mark Evidence*, 30 CARDOZO L. REV. 1369, 1375–88 (2009); NAS REPORT, *supra* note 2, at 173–76.

⁸⁰ *See Beecher-Monas, supra* note 79, at 1378–84.

⁸¹ *See id.* at 1385–87.

(without accurate foundation) in trials.⁸² Bite mark analysts often testify using phrases such as “consistent with” or “positively match.”⁸³ Yet such statements have no grounding in science, because they are purely subjective conclusions unsupported by data.⁸⁴ In addition, as with other forms of individualization forensic science evidence, statements about a match are neither scientifically supportable⁸⁵ nor scientifically meaningful.⁸⁶ Early research, however, appears to indicate that such statements are convincing to juries.⁸⁷

The NAS Report finds plenty to criticize about forensic odontology, concluding that no scientific studies support the claims that odontologists can demonstrate sufficient detail for positive identification. In numerous instances, “experts diverge widely in their evaluations of the same bite mark evidence,” which raises serious questions about the value and scientific objectivity of the discipline.⁸⁸ Of course, like every other form of individualization evidence, judges have admitted it readily into the courtroom.⁸⁹

In *United States v. Crisp*, the court admitted both fingerprint comparison and handwriting comparison evidence despite strong, solid challenges to the evidence.⁹⁰ Regarding fingerprint comparison, the court held that “the principles underlying fingerprint identification . . . bear the imprimatur of a strong general acceptance, not only in the expert community, but in the courts as well. . . . Put simply, . . . [there is] no reason . . . to believe that this general acceptance of the

⁸² See, e.g., *State v. Garrison*, 585 P.2d 563, 566 (Ariz. 1978) (affirming trial court’s decision admitting expert testimony that the probability of two sets of teeth being identical was eight in one million, based on “articles written throughout the literature that do mention the possibility or the numerical values of finding two (sets of teeth) that match.”).

⁸³ Beecher-Monas, *supra* note 79, at 1386.

⁸⁴ See A.P.A. Broeders, *Of Earprints, Fingerprints, Scent Dogs, Cot Deaths, and Cognitive Contamination—A Brief Look at the Present State of Play in the Forensic Arena*, 159 FORENSIC SCI. INT’L 148, 153–54 (2006) (explaining why forensic science individualization does not meet the scientific standard).

⁸⁵ See NAS REPORT, *supra* note 2, at 87 (“[N]o forensic method other than nuclear DNA analysis has been rigorously shown to have the capacity to consistently and with a high degree of certainty support conclusions about ‘individualization’ (more commonly known as ‘matching’ of an unknown item of evidence to a specific unknown source).”).

⁸⁶ See NAS REPORT, *supra* note 2, at 21 (noting the need for standardized testimony).

⁸⁷ See Dawn McQuiston-Surrett & Michael J. Saks, *Communicating Opinion Evidence in the Forensic Identification Sciences: Accuracy and Impact*, 59 HASTINGS L.J. 1159, 1188 (2008) (explaining studies conducted to determine how jurors analyze such terms as “match” and “consistent with” and finding that jurors overestimated the meaning of such terms); accord NAS REPORT, *supra* note 2, at 21 (“[U]se of such terms can and does have a profound effect on how the trier of fact . . . perceives and evaluates scientific evidence.”).

⁸⁸ See NAS REPORT, *supra* note 2, at 176 (“[T]here is considerable dispute about the value and reliability of the collected data for interpretation.”).

⁸⁹ See cases collected in Beecher-Monas, *supra* note 79, at 1372–74, and accompanying notes.

⁹⁰ See *United States v. Crisp*, 324 F.3d 261, 271–72 (4th Cir. 2003).

principles underlying fingerprint identification has, for decades, been misplaced.”⁹¹ Thus, in a sweeping but unsupported statement, the court concluded that “the district court was well within its discretion in accepting at face value the consensus of the expert and judicial communities that the fingerprint identification technique is reliable.”⁹² With regard to handwriting comparison, the court stated, “[t]he fact that handwriting comparison analysis has achieved widespread and lasting acceptance in the expert community gives us the assurance of reliability that *Daubert* requires.”⁹³ By not actually evaluating the testimony using *Daubert*’s suggested factors,⁹⁴ the court evaded the problem that the evidence did not meet reliability requirements, a point noted by the dissent, which explained in detail how the government “utterly failed to meet its burden” of establishing reliability under *Daubert*.⁹⁵

The NAS Report criticizes both handwriting and fingerprint comparison. The report finds that the standards for fingerprint comparison are subjective, as are the declarations of a match.⁹⁶ In fact, not only is the outcome of analysis not necessarily repeatable from one examiner to another, research cited by the report indicates that experienced examiners do not even necessarily agree with their own prior conclusions!⁹⁷ Analysis of the methodology used by the government to declare fingerprint matches is so entirely without foundation or objective standards that even the validity of the method cannot be tested.⁹⁸ Moreover, even if the two foundations of the specialty are true—that fingerprints are unique and persist unchanged throughout life—that does not imply that “anyone can reliably discern whether or not two friction ridge impressions were made by the same person.”⁹⁹ Similarly, handwriting comparison fares even more poorly. Conclusions are entirely subjective, and the NAS Report recommends that its “scientific basis . . . needs to be strengthened.”¹⁰⁰

Regarding microscopic hair comparison, the NAS Report was even more critical, commenting on its high rate of error and stating “the committee found no scientific support for the use of hair comparisons for individualization in the absence of nuclear DNA.”¹⁰¹ In sum, the non-DNA individualization evidence is

⁹¹ *Id.* at 268–69.

⁹² *Id.* at 269.

⁹³ *Id.* at 271.

⁹⁴ See *Kumho Tire Co. v. Carmichael*, 526 U.S. 137, 152 (1999) (“[A] trial court should consider the specific factors identified in *Daubert* where they are reasonable measures of the reliability of expert testimony.”).

⁹⁵ *Crisp*, 324 F.3d at 273 (Michael, J., dissenting).

⁹⁶ NAS REPORT, *supra* note 2, at 139 (noting that the threshold for making a source identification is “deliberately kept subjective”).

⁹⁷ *Id.*

⁹⁸ *Id.* at 143 n.31 and accompanying text.

⁹⁹ *Id.* at 143–44.

¹⁰⁰ *Id.* at 166.

¹⁰¹ NAS REPORT at 161.

simply unsupported by research, and serious questions are raised about its evidentiary reliability.

Yet the report is willing to absolve the courts of any real responsibility in enforcing the reliability mandate, citing various reasons: judicial decisions about reliability are “flexible,” and much discretion is granted to the trial court;¹⁰² judges (and lawyers) “lack the scientific expertise necessary to comprehend and evaluate forensic science in an informed manner”;¹⁰³ judges work alone and “often [have] little time for extensive research and reflection”; and, echoing many others, the adversarial process is “not suited to the task of finding ‘scientific truth.’”¹⁰⁴

There is merit in each of these justifications for courts’ unwillingness to subject most of forensic science to any meaningful analysis. And yet, the justifications seem weak, particularly in light of those same courts’ ability to tackle far more complicated questions of scientific causation in exceptionally technical civil cases. And the justification is sorely inadequate when considering the difference in what is at stake in the respective cases. Regarding challenges to fingerprint comparison, Professor Mnookin writes, “The real embarrassment is the way that the courts have been a willing accomplice in this turf battle, in the process abnegating their gatekeeping responsibilities under *Daubert*.”¹⁰⁵

In civil cases, courts seem quite up to the task of evaluating microbiology, teratology, and toxicology evidence, discussing both science and statistics with plenty of acumen.¹⁰⁶ Yet when it comes to evaluating the shortcomings of lip prints and handwriting, courts are unable to muster the most minimal grasp of why a standardless form of comparison might lack evidentiary reliability or trustworthiness.¹⁰⁷ Rather, they have relied on tired aphorisms and biased heuristics so as to avoid stating the obvious: there are no standards, there is no legitimate basis for the matches testified to, and forensic individualization is currently built on sand.

¹⁰² *Id.* at 96–97.

¹⁰³ *Id.* at 110.

¹⁰⁴ *Id.*

¹⁰⁵ Mnookin, *supra* note 35, at 141.

¹⁰⁶ See cases cited *supra* note 55. Granted, part of the courts’ apparent comfort in the civil cases may arise from the parties’ well-structured challenges and responses to the expert testimony. Many criminal practitioners are not developing the challenges as well as their civil counterparts. Nonetheless, serious, science-based challenges have been raised against fingerprints, toolmarks, and handwriting comparison in many criminal cases, and nearly universally courts have rejected them. See cases cited *supra* notes 60–64.

¹⁰⁷ *Daubert*, 509 U.S. 579, 590 n.9 (describing the standard as “evidentiary reliability—that is, trustworthiness”).

IV. “HISTORY MAY BE SERVITUDE”:¹⁰⁸ THE ERRONEOUS JUDICIAL FOCUS ON THE LONG HISTORY OF USE OF INDIVIDUALIZATION EVIDENCE

Courts have found a variety of reasons to admit individualization evidence, yet one of the most common justifications courts provide is the long history of use. Yet a long history of use confers no particular proof of validity. From the time of Nicader of Colophain, 200 B.C., through the height of their popularity in the mid-nineteenth century, leeches were used to “bleed” the sick as a form of treatment for virtually all ailments, from pneumonia to hemorrhoids.¹⁰⁹ As late as 1920, the use of lancets and leeches for bloodletting was favored by some physicians to treat pneumonia.¹¹⁰ One would think that reflective consideration would have shown the lack of bloodletting efficacy, yet physicians continued to employ it, year after year, decade after decade, century after century. It was not until the rise of experimental methodology and restraints on empirical methods began to gain ground that the use of bloodletting and leeches in everyday medicine began to drop off and subsequently evaporate.¹¹¹

Likewise, the “long history of use” argument failed to establish underlying validity of expert testimony in the Salem witchcraft trials. Cotton Mather and those overseeing the trials approved of the well-known publications¹¹² providing expert guidance for courts on the proof of witchcraft—expert treatises that had stood the test of time in other prosecutions.¹¹³ The principles of the “experts” were

¹⁰⁸ See T.S. Eliot, *Little Gidding*, in *THE COMPLETE POEMS AND PLAYS* 142 (Harcourt Brace & World, Inc., 1971).

¹⁰⁹ I.S. Whitaker et al., *Hirudo Medicinalis: Ancient Origins of, and Trends in the Use of Medicinal Leeches Throughout History*, 42 *BRITISH J. OF ORAL & MAXILLOFACIAL SURGERY* 133, 134–36 (2004); accord N. Papavramidou & H. Christopoulou-Aletra, *Medicinal Use of Leeches in the Texts of Ancient Greek, Roman and Early Byzantine Writers*, 39 *INTERNAL MEDICINE J.* 624, 624–27 (2009) (discussing the historical origins of the medicinal use of leeches). Others claim bloodletting as a form of treatment has a 3,000-year history. See Gilbert R. Seigworth, *Bloodletting over the Centuries*, 80 *N.Y. ST. J. MED.* 2022–28 (Dec. 1980) (citing the use of bloodletting by the ancient Egyptians). In any event, the use of leech bleeding has a long history indeed. David Faigman also has mentioned the leech-bleeding analogy. David L. Faigman, *Anecdotal Forensics, Phrenology, and Other Abject Lessons from the History of Science*, 59 *HASTINGS L.J.* 979, 985 (2008).

¹¹⁰ Seigworth, *supra* note 109, at 2024.

¹¹¹ Recently, scientists have discovered the important anticoagulant properties of leeches and have successfully used leeches in microsurgical application. Whitaker, *supra* note 109, at 136. However, the use of leeches for simply “bleeding cures” for various disease processes turns out not to have much foundation.

¹¹² See RICHARD BERNARD, *GUIDE TO GRAND JURY MEN* (1627); WILLIAM PERKINS, *DISCOURSE ON THE DAMNED ART OF WITCHCRAFT* (1608); see also FRANCES HILL, *THE SALEM WITCH TRIALS READER* 3 (2000).

¹¹³ Jane Campbell Moriarty, *Wonders of the Invisible World: Prosecutorial Syndrome and Profile Testimony in the Salem Witchcraft Trials*, 26 *VT. L. REV.* 43, 57–63 (2001) (discussing the expert authority on which judges relied to detect witchcraft).

recognized by courts for nearly a century, which was a major reason the courts in Salem were swayed by such opinion.

The Oracle of Delphi provides yet a third example of the fallacy of historical reliance.¹¹⁴ Citizens, both of high and low status, would approach the oracle and ask questions, often about quotidian matters such as whether to marry or matters of a more serious nature, such as whether to resist the Persian invasion.¹¹⁵ Despite advice from the oracle not always proving prescient or wise,¹¹⁶ the oracle was consulted for over a thousand years.¹¹⁷

We now appreciate that oracles cannot predict future events, that witchcraft is not real, and that bleeding is generally not effective to cure disease, but old ideas die hard. Perhaps there is a lesson here useful to the judiciary.

V. JUDICIAL DECISION MAKING AND THE PROBLEMS OF COGNITIVE BIAS

Legal, psychological, and sociological scholars have all examined and opined about judicial decision making to determine how judges decide cases.¹¹⁸ The methods of analysis and theories posed are varied, rich, and complex, suggesting that decision making is a product of reason and intuition. Some find that political agendas or background and experience inform decision making, while others argue that judges are influenced by precedent. One theme, however, that resonates throughout much of the literature is that “judges are human.”¹¹⁹ They are swayed

¹¹⁴ HUGH BOWDEN, CLASSICAL ATHENS AND THE DELPHIC ORACLE, DIVINATION AND DEMOCRACY 19 (2005).

¹¹⁵ *Id.* at 29–30.

¹¹⁶ *Id.* at 26–28. It cannot be lost on the reader that many superstitious beliefs persist despite the absence of accuracy. Today, some remnants of the bloodletting years continue to hold sway—recently the use of “cupping”—a form of “dry bloodletting”—has found favor with some alternative medicine fans. For more on its historical use, see Seigworth, *supra* note at 109. For a fuller explanation of modern-day cupping, see Welcome to the British Cupping Society, <http://www.britishcuppingsociety.org/Portal/index.php> (last visited June 1, 2010).

¹¹⁷ BOWDEN, *supra* note 114, at 19.

¹¹⁸ See RICHARD A. POSNER, HOW JUDGES THINK (2008); CASS SUNSTEIN, ET AL., ARE JUDGES POLITICAL? AN EMPIRICAL ANALYSIS OF THE FEDERAL JUDICIARY (2006); THE PSYCHOLOGY OF JUDICIAL DECISION MAKING (David Klein and Gregory Mitchell eds., 2010); Chris Guthrie, Jeffrey J. Rachlinski & Andrew Wistrich, *Blinking on the Bench: How Judges Decide Cases*, 93 CORNELL L. REV. 1 (2007) [hereinafter *Blinking on the Bench*]; Chris Guthrie, Jeffrey J. Rachlinski & Andrew Wistrich, *Inside the Judicial Mind*, 86 CORNELL L. REV. 777 (2001); Chad M. Oldfather, *Writing, Cognition, and the Nature of the Judicial Function*, 96 GEO. L.J. 1283 (2008); Frederick Schauer, *Do Cases Make Bad Law?*, 73 U. CHI. L. REV. 893 (2006); Dan Simon, *A Psychological Model of Judicial Decision Making*, 30 RUTGERS L.J. 1 (1998).

¹¹⁹ For more on this idea, see Chad M. Oldfather, *Judges and Humans: Interdisciplinary Research and the Problems of Institutional Design*, 36 HOFSTRA L. REV. 125, 128 n.11 (2007) (collecting articles referencing the different ways that judges are subject to the same frailties as other humans). Jeff Rachlinski also reminded me that many

by heuristic decision making, friendships, beauty, the strength of a case, public opinion, fear of reversal, and the normal set of cognitive biases to which we all are subject: expectation bias, hindsight bias, confirmation bias, tunnel vision, and so forth.

In an interesting empirical article, *Blinking on the Bench: How Judges Decide Cases*, the authors write about the dual-process models of cognition (intuitive and deliberative).¹²⁰ Roughly explained, intuitive decision making is spontaneous, effortless, relies on pre-existing heuristics, and is fast,¹²¹ while the deliberative process is slower, requiring more “effort, motivation, concentration, and the execution of learned rules.”¹²² The *Blinking on the Bench* authors conclude that judges tend to use intuitive decision making for the everyday problems they see on the bench.¹²³ Thus, “[w]hen ruling on the admissibility of evidence at trial, judges often have little choice but to think intuitively.”¹²⁴

Many might consider intuitive decision making in this role to be beneficial because judges have “typical” evidence questions, think about admissibility questions all the time, and have abundant experiential knowledge that informs the snap judgments that are needed at trial. The use of heuristics to think quickly and decide intuitively has great value. We don’t deliberate when a ball is thrown to us; we either catch it or get bonked on the head. Similarly, in trial, judges need to decide evidentiary questions in real time. They cannot over-deliberate every time an objection is raised.

But there is a decided downside to such quick thinking. There is growing evidence that intuitive and impressionistic decisions about evidence may be more error-prone than a more deliberative process. In an experiment with judges, researchers discovered that when subjects were given a problem they thought they

judges have experience as prosecutors: this background may lead them to think of prosecutorial experts as “tools to build [the] cases” rather than as partisans. (e-mail on file with author).

¹²⁰ *Blinking on the Bench*, *supra* note 118, at 6–9. For a more detailed discussion of this distinction, see D. KAHNEMAN, P. SLOVIC & A. TVERSKY, *JUDGMENT UNDER UNCERTAINTY: HEURISTICS AND BIASES* (1982); T. GILOVICH, D. GRIFFIN & D. KAHNEMAN, *HEURISTICS AND BIASES: THE PSYCHOLOGY OF INTUITIVE JUDGMENT* (2002). For a more precise explanation of the distinction, see Steven A. Sloman, *Two Systems of Reasoning*, in GILOVICH ET AL., *supra* note 120, at 379. Other scholars have written about the dual process model of decision making. In the area of scientific evidence law, Joseph Sanders has written a most compelling and instructive article. See Joseph Sanders, *Kumho and How We Know*, 64 L. & CONTEMP. PROBS. 373, 393 (2001) (describing the concepts of “experiential processing” and “rational processing”).

¹²¹ *Blinking on the Bench*, *supra* note 118, at 7.

¹²² *Id.* (quoting Shane Frederick, *Cognitive Reflection and Decision Making*, 19 J. ECON. PERSP. 25, 26 (2005)).

¹²³ *Id.* at 27.

¹²⁴ *Id.* at 36.

could solve intuitively, the error rate was substantially higher than with a problem the judges had to actually reason their way through to resolve.¹²⁵

Moreover, the problem of cognitive bias is most apparent, according to researchers, in the use of intuitive decision making, and it is where such problems as stereotyping, prejudice, and discrimination are likely to arise.¹²⁶ Relying on heuristic shortcomings can lead to systematic biases.¹²⁷

Thus, it may be worthwhile to consider both the dual processing systems and bias when thinking about courts' decisions on individualizing evidence in the past and going forward into the future.

Part of the courts' persistence in finding individualization specialties reliable may be due to judges using intuitive decision making both before trial and during trial when resolving admissibility questions about fingerprinting, hair, and handwriting comparison. As noted in *Blinking on the Bench*, when judges are presented with a problem they perceive to be simple, they use heuristics to solve it—potentially making errors and not recognizing them due to the belief that the issue before them is not complex. It may well be that fingerprint, handwriting, and microscopic hair comparisons seem to be “simple” problems for judges, leading them to use intuitive decision making about the outcome (although with some rational processing, of course, in the written decision).

Consider the diametric approaches courts employ when deciding admissibility of scientifically complex expert evidence (such as DNA comparison or toxic tort causation) versus the non-DNA individualization evidence such as fingerprints and handwriting. In the complex scientific evidence cases, courts appear to use a more rational processing system and engage in deliberate, analytic reasoning throughout the opinion to determine whether the evidence is reliable.¹²⁸ A good example of this methodology is found in the Supreme Court's opinion in *General Electric v. Joiner*, in which the court engaged in a long methodological analysis of the quality of epidemiological evidence to determine whether plaintiff's proof that PCB exposure had promoted his lung cancer was sufficiently reliable to be admitted.¹²⁹ None of the writing appears intuitively-based.

By contrast, very few cases involving individualization evidence seem to be of the long, methodological analysis present in *Joiner*. To the contrary, courts seem to rely on such heuristic devices as “long history of use” or “generally

¹²⁵ *Id.* at 10–13 (discussing the Cognitive Reflection Test Model from Shane Frederick, *Cognitive Reflection and Decisionmaking*, *supra* note 122, at 26–28, and the judges' performance on the test).

¹²⁶ Thomas Gilovich and Dale Griffin, *Introduction—Heuristics and Biases: Then and Now*, in *JUDGMENT UNDER UNCERTAINTY*, *supra* note 120, at 7.

¹²⁷ See BENJAMIN R. NEWELL, ET AL., *THE PSYCHOLOGY OF DECISION MAKING*, 71 (2007) (citing the oft-quoted Kahneman, D. & Tversky, A., *The Simulation Heuristic*, in *KAHNEMAN ET AL.*, *supra* note 120; see also BURTON, *supra* note 36, at 149).

¹²⁸ This does not mean that error cannot occur in the deliberative process or that judges always get the right answer. All it means, according to research psychology, is that the probability of error is reduced using this method.

¹²⁹ *Gen. Elec. Co. v. Joiner*, 522 U.S. 136 (1997).

accepted by other courts” to support their decisions, rather than engaging the evidence and subjecting it to a rational, science-based analysis.¹³⁰

In fact, the reliance on “long history of use” seems to reflect a common bias that affects intuitive decision making—namely, the concept of “belief perseverance.”¹³¹ This form of bias is “the tendency to maintain existing beliefs in the face of evidence that ought to weaken or even totally reverse those beliefs.”¹³² According to social psychologists, whatever is learned first seems to have a “primacy effect”—information presented earlier has more influence on judgments than information presented later.¹³³ Individuals exposed to subsequent, possibly contradictory or conflicting information, disregard the later information, assume it is less reliable or valid, or interpret later evidence in ways that is consistent with their initial impression (perhaps explaining the long history of leech craft).¹³⁴ Moreover, the problem of “confirmation bias” causes people to seek out information that supports the original belief and to avoid information that contradicts those beliefs.¹³⁵

The combination of belief perseverance and confirmation bias might explain judges’ reluctance to find the so-called matches unsupported—even in the face of ample, compelling testimony that there is absolutely no legitimate support for such conclusions. Consider the contemporary physician writing about leeches, who

¹³⁰ United States v. Crisp, 324 F.3d 261 (4th Cir. 2003) (majority admitting the evidence uses these intuitive heuristics; the dissent, disallowing the evidence, engages in a long, deliberative analysis).

¹³¹ See RICHARD NESBITT & LEE ROSS, HUMAN INFERENCE: STRATEGIES AND SHORTCOMINGS OF SOCIAL JUDGMENT 18–23, 24–42, 45–53, 462 (1980).

¹³² Philip E. Tetlock, *Accountability and the Perseverance of First Impressions*, 46 SOC. PSYCHOL. Q. 285, 285 (1983).

¹³³ *Id.* at 286.

¹³⁴ *Id.* Another bias that might be at work here is the so-called “sunk cost” fallacy, often referred to as “throwing good money after bad.” See Hal R. Arkes & Catherine Blumer, *The Psychology of Sunk Cost*, 35 ORG. BEHAV. & HUM. DECISION PROCESSES, 124–40 (1985). This theory describes an individual’s unwillingness to withdraw from an endeavor after investing money, time, or effort. This bias explains why investors who lose a great deal of money may not be willing to cut their losses and is one explanation for judges’ unwillingness to recognize that much of their prior decision making was premised on erroneous beliefs—they are simply “too invested” in their prior decisions to back out. *Id.*

Other articles have discussed the role of cognitive bias in the creation of and decisions about forensic science, including expectation bias, confirmation bias, and so forth. See Keith A. Findley & Michael S. Scott, *The Multiple Dimensions of Tunnel Vision in Criminal Cases*, 2006 WISC. L. REV. 291; Jane Campbell Moriarty, “*Misconvictions, Science, and the Ministers of Justice*,” 86 NEB. L. REV. 1, 8, 25 (2007); Michael Risinger et al., *The Daubert/Kumho Tire Implications of Observer Effects in Forensic Science: Hidden Problems of Expectation and Suggestions*, 90 CAL. L. REV. 1 (2002).

¹³⁵ Karl Ask and Pär Anders Granhag, *Motivational Sources of Confirmation Bias in Criminal Investigations: The Need for Cognitive Closure*, 2 J. INVESTIG. PSYCHOL. OFFENDER PROFIL. 43, 45 (2005) (discussing aspects of confirmation bias).

muses, “[i]t seems hard to believe that the many educated observers over the centuries were completely wrong in their assessment of clinical improvement following bloodletting.”¹³⁶ This comment seems to reinforce the powerful effect of belief perseverance, even in the face of abundant data to the contrary.¹³⁷

VI. THE STORY POST-NAS REPORT: ONE YEAR LATER

So now that the report, with its excoriation of the current state of individualization specialties, has been published for more than a year, it is interesting to see how courts have responded. As of the beginning of May 2010, there are about a dozen cases mentioning or discussing the report—one U.S. Supreme Court case discussing the Confrontation Clause,¹³⁸ several federal district court cases, and three state cases—a few of which will be discussed. None of the challenges seeking to exclude forensic science evidence on reliability grounds has succeeded in the court. Two decisions have placed some constraints upon the testimony—allowing the conclusion of a match while limiting or excluding the degree of certainty testimony,¹³⁹ and two other federal cases appeared unmoved by the report in analyzing a question of fingerprint comparison admissibility.¹⁴⁰

In deciding that lab results constituted the “testimonial” statements subject to Confrontation Clause mandates, the Supreme Court in *Melendez-Dias v. Massachusetts* mentioned the report, stating:

“[T]here is wide variability across forensic science disciplines with regard to techniques, methodologies, reliability, types and numbers of potential errors, research, general acceptability, and published material.” National Academy Report . . . (discussing problems of subjectivity, bias, and unreliability of common forensic tests such as latent fingerprint analysis, pattern/impression analysis, and toolmark and firearms analysis). Contrary to respondent’s and the dissent’s suggestion, there is little reason to believe that confrontation will be useless in testing

¹³⁶ Seigworth, *supra* note 109, at 2027.

¹³⁷ In fact, in comparisons of clinical judgments and actuarial judgments (controlled studies), clinical judgments fare more poorly. One scholar notes that “[f]ailure to accept a large and consistent body of scientific evidence over unvalidated personal observation may be described as a normal human failing or, in the case of professionals who identify themselves as scientific, plainly irrational.” Robyn M. Dawes et al., *Clinical Versus Actuarial Judgment*, in KAHNEMAN ET AL., *supra* note 120, at 716, 727.

¹³⁸ See *Melendez-Dias v. Massachusetts*, 129 S. Ct. 2527, 2537 n.6 (2009) (noting the report’s conclusion that the forensic science system is badly flawed and refuting the suggestion that forensic examiners are “uniquely reliable”).

¹³⁹ See *United States v. Willock*, Criminal No. WDQ-08-0086, 2010 U.S. Dist. LEXIS 27473, *25 (D. Md. Mar. 23, 2010); *United States v. Taylor*, 663 F. Supp. 2d 1170, 1177 (D.N.M. 2009).

¹⁴⁰ See *United States v. Montalvo-Rangel*, No. SA-10-CR-64-XR, 2010 WL 1484708 (W.D. Tex. Apr. 5, 2010); *United States v. Rose*, 672 F. Supp. 2d 723 (D. Md. 2009).

analysts' honesty, proficiency, and methodology—the features that are commonly the focus in the cross-examination of experts.¹⁴¹

It seems to be a good signal that the U.S. Supreme Court is noticing that “common forensic tests” may be unreliable. In the inferior federal courts, the picture may not be so clear.

In *United States v. Montalvo-Rangel*, the defendant challenged the latent-print examiner's testimony on the basis of the NAS Report. In a one-paragraph discussion, the court overruled the objection with no analysis, only noting that the opinion of the examiner matching the print was given with “great confidence.”¹⁴²

In *United States v. Rose*, the trial court found persuasive the “generally accepted by other courts” rationale: “Before and after *Crisp*, it appears that every federal circuit . . . has found expert fingerprint identification testimony admissible”¹⁴³ *Rose*, however, does discuss the NAS Report but decides that it did not “conclude that fingerprint evidence was unreliable such as to render it inadmissible”¹⁴⁴ In fact, the judge in *Rose* goes so far as to say that “Judge Harry Edwards, who co-chaired the project, made it clear that nothing in the report was intended to answer the ‘question whether forensic science evidence in a particular case is admissible under applicable law.’”¹⁴⁵

As a result of this type of analysis (and several briefs citing Judge Edwards), Judge Edwards responded pointedly:

If courts blindly follow precedent that rests on unfounded scientific premises, this will lead to unjust results. Nothing in established law compels this course. So when the report was released and I said that judges must continue to follow the law, I did not mean to suggest that judges would apply existing law without taking into account the findings in the report that raise serious doubts about the validity and reliability of certain forensic disciplines and practices.

. . .

The point here is simple: *When scientific methodologies once considered sacrosanct are modified or discredited, the judicial system must accommodate the changed scientific landscape.*¹⁴⁶

Two cases involving firearm toolmark comparison engage in a more thoughtful analysis of the meaning of the report, but neither appears to grasp the implication of why a conclusion of a “match” is not currently supportable.

¹⁴¹ *Melendez-Diaz*, 129 S. Ct. at 2538 (2009) (holding that the Confrontation Clause applies to laboratory reports in criminal cases and discussing the findings of the NAS Report).

¹⁴² *Montalvo-Rangel*, 2010 WL 1484708, at *3.

¹⁴³ *Rose*, 672 F. Supp. 2d at 725.

¹⁴⁴ *Id.*

¹⁴⁵ *Id.*

¹⁴⁶ Edwards, *supra* note 31, at 6–7.

United States v. Taylor, a district court decision from New Mexico, provides a detailed explanation of the shortcomings of firearm toolmark comparison, discussing the lack of standards by which a match is declared, the subjective basis for the conclusion of a match, and the lack of standards for even distinguishing between class, subclass, and individual characteristics.¹⁴⁷ “[T]he . . . theory is circular. An examiner may make an identification when there is sufficient agreement, and sufficient agreement is defined as enough agreement for an identification.”¹⁴⁸ The court also quotes the damning language from the NAS Report that, even with better training and new techniques, “the decision of the toolmark examiner remains a subjective decision based on unarticulated standards and no statistical foundation for estimation of error rates.”¹⁴⁹

Nonetheless, the significance of those failures¹⁵⁰ is swept aside as orthogonal to the underlying validity. The court, citing pre-NAS Report cases, finds the practice of cartridge comparison sufficiently reliable to be admitted and permits the expert to give his opinion “within a reasonable degree of certainty in the firearms examination field.”¹⁵¹

*United States v. Willock*¹⁵² is a deeper, more thoughtful analysis. The court cites the conclusion expressed in the NRC Ballistic Imaging Report: “The validity of the fundamental assumptions of uniqueness and reproducibility of firearms-related toolmarks has not yet been fully demonstrated,” and “additional general research on the uniqueness and reproducibility of firearms-related toolmarks would have to be done if the basic premises of firearms identification are to be put on a more solid scientific footing.”¹⁵³ Yet, in its analysis, the court finds the pre-NAS Report decisions compelling, noting that “the furthest” any court to date has gone is to exclude testimony where examiners’ results were not confirmed or documented; or “to restrict the degree of certainty to which the examiners could express their identifications.”¹⁵⁴ The courts seems to note—without disapproval—

¹⁴⁷ 663 F. Supp. 2d 1170, 1177–78 (D.N.M. 2009).

¹⁴⁸ *Id.* at 1177.

¹⁴⁹ *Id.* at 1178 (citing the NAS REPORT, *supra* note 2, at 5–20).

¹⁵⁰ These are not the only shortcomings the court notes. While proficiency tests have been conducted, none of them was done as a blind test, which raises doubts about the value of the tests. *Id.* at 1176.

¹⁵¹ *Id.* at 1180.

¹⁵² *United States v. Willock*, Criminal No. WDQ-08-0086, 2010 U.S. Dist. LEXIS 27473, *25 (D. Md. Mar. 23, 2010).

¹⁵³ *Id.* at *15 (quoting NATIONAL RESEARCH COUNCIL, BALLISTIC IMAGING 3 (2008)).

¹⁵⁴ *Id.* at *16 (citing *United States v. Monteiro*, 407 F. Supp. 2d 351, 374 (D. Mass. 2006); *United States v. Taylor*, No. CR 07-1244, 2009 WL 3347485, at *9 (D.N.M. Oct. 9, 2009); *United States v. Glynn*, 578 F. Supp. 2d 567, 570 (S.D.N.Y. 2008)). Curiously, the court does not mention here *United States v. Green*, 405 F. Supp. 2d 104, 124 (D. Mass. 2005), in which the court disallowed any conclusions about a match, although it references the opinion elsewhere. See *Willock*, 2009 WL 3617748, at *14 (discussing *Green* in the context of admissibility concerns of firearm toolmark identification testimony).

that one 2007 federal case allowed the expert to state a conclusion with “100% degree of certainty.”¹⁵⁵

Quoting *Kumho Tire* for the principle that the *Daubert* factors are not “holy writ,” the *Willock* court does not address the question of whether toolmark identification is science (implicitly failing to recognize that claims of a match are based upon scientific and statistical principles) and largely follows suit with the pre-NAS Report cases.¹⁵⁶ However, *Willock* goes beyond most courts in limiting the testimony: The trial court requires the prosecution to present testimony only from (a) a qualified examiner; (b) who followed the standard theory (“despite its subjectivity”); and (c) who documented in detail his procedures so as to allow another examiner to follow the original steps.¹⁵⁷ However, the court goes one step further than the pre-NAS Report toolmark cases and restricts the examiner to stating opinions and conclusions “without any characterization as to the degree of certainty with which he holds them.”¹⁵⁸ This case is the first to so limit the testimony and may usher in a generation of new decisions.

While both of the courts recognize that toolmark comparison “does not have sufficient rigor to be evaluated as science,”¹⁵⁹ they categorize this testimony as either technological or other specialized testimony¹⁶⁰ and then proceed to admit the testimony with some limitations. What the courts continue to miss, however, is that the declaration of a match (a conclusion of individualization), is, by its very nature, a scientific finding:

[A] conclusion of individualization implies that the evidence originated from that source, to the exclusion of all other possible sources. The determination of uniqueness requires measurements of object attributes, data collected on the population frequency of variation in these attributes, testing of attribute independence, and calculations of the probability that different objects share a common set of observable attributes.¹⁶¹

The judiciary to date still does not seem to understand fully the nature of the problem; while some courts explain the problem quite well, they are unprepared to foreclose the expert from declaring a match. Other courts, like *United States v. Rose*, have simply proceeded along as if the report was meaningless; perhaps

¹⁵⁵ See *Willock*, 2009 WL 3617748, at *16 (citing *United States v. Natson*, 469 F. Supp. 2d 1253, 1261–62 (M.D. Ga. 2007)).

¹⁵⁶ See *id.* at *19.

¹⁵⁷ *Id.* at *20.

¹⁵⁸ *Id.* at *24.

¹⁵⁹ *United States v. Taylor*, 663 F. Supp. 2d 1170, 1179 (D. N.M. 2009) (quoting *United States v. Glynn*, 578 F. Supp. 2d 567, 570 (S.D.N.Y. 2008)).

¹⁶⁰ See *United States v. Mouzone*, Criminal No. WDQ-08-086, 2009 WL 3617748, at *19 (D. Md. Oct. 29, 2009).

¹⁶¹ NAS REPORT, *supra* note 2, at 43–44.

proving the triumph of belief perseverance in the face of contradictory (and compelling) information.

I suggest that we should encourage judges to work with the language of the NAS Report to write an opinion in which they address the following factors: (1) measurement of object attributes; (2) data on population frequency of variation in the attributes; (3) evidence of attribute independence; and (4) calculation of the probability that different objects share a common set of attributes.

By urging judges to use these factors, which are complex and intricate (and follow the template of DNA comparison), judges will be less inclined to resort to heuristic-based intuitive decision making and will instead rely on deliberative process. In working through the deliberative process, it becomes clear very quickly that these elements cannot be met. Thus, the only response is to exclude the evidence or disallow conclusions of a match. Most likely, the latter is the course courts would take, and in my opinion the correct approach. As is true in both love and science, there is nothing simple about a match.

V. CONCLUSION

With other scholars, I have previously argued that because judges apparently were unwilling to exclude prosecutorial forensic evidence, perhaps judges could follow the “middle way” by limiting, if not excluding, the testimony.¹⁶² My suggestion (and that of others) is primarily to let the expert testify about points of comparison, without giving a conclusion to the jury. This approach does not resolve the problems identified in the report; but it possibly cures the worst problems with individualization evidence.

One might have hoped that the courts would, at least, be willing to take this step after the NAS Report waved red flags about the problems of forensic individualization and the worrisome implications of those shortcomings. However, it seems apparent in the few cases that have been decided to date that courts are operating predictably with belief perseverance and are simply assimilating the implications of the NAS Report by interpreting the report to conform with their prior beliefs.

Again, however, social science may provide some clues as to how to affect this particular form of cognitive bias so that judges really understand the dangers of admitting conclusions about matches: namely, by requiring greater accountability of judges in their decision making. If subjects are told ahead of time that they will be accountable for their judgments, they are much less susceptible to primacy or belief perseverance.¹⁶³ Here the role of the Supreme Court is critical: if the Court continues to recognize the problems in forensic science, as Justices

¹⁶² Moriarty, *supra* note 134, at 39–41; Moriarty & Saks, *supra* note 25, at 29. Additional suggestions include greater appointment of defense experts and more testimony about actual error rates. For a fuller discussion, see Moriarty *supra* note 134, at 40–41.

¹⁶³ Tetlock, *supra* note 132, at 290–91 (suggesting that pre-exposure accountability information may reduce the primacy effect).

appeared to in the language of *Melendez-Diaz*, the inferior federal courts likely will realize that they will be held accountable for their decisions on forensic science and will begin to evaluate the testimony in a more critical, thoughtful fashion.

Moreover, the recent comments of Judge Edwards about the report provide a crucial first step to correcting the course that some courts have chosen. The importance of judges holding the prosecution to legitimate reliability standards cannot be underestimated. To paraphrase Judge Gertner, when life or liberty hangs in the balance, we should expect better forensic science evidence than has been historically demanded. “We should require more,”¹⁶⁴ both from our forensic science experts and from our judiciary.

¹⁶⁴ United States v. Green, 405 F. Supp. 2d 104, 109 (D. Mass. 2005).