Who speaks for science? A response to the National Academy of Sciences Report on forensic science

SIMON A. COLE†
Associate Professor of Criminology, Law & Society, University of California, Irvine, CA 92697-7080, USA

[Received on 15 May 2009; revised on 15 September 2009; accepted on 16 September 2009]

This response focuses on the treatment of latent print identification by the recent National Academy of Science (NAS) Report on forensic science. It begins by situating the Report in the historical context of a decade of controversy over the validity of latent print identification. Stark disagreement between the academic and judicial communities over this issue created a situation in which the question of which of these two communities would ‘speak for science’ became contested. The Report’s support of the academic position demonstrated the lack of support among non-practitioners for the claims of extreme discrimination and accuracy advanced on behalf of latent prints. The Report in some sense constitutes the response of institutionalized science to this issue. Nonetheless, it is still unclear whether the Report will function, as some may have hoped, as a ‘court of last resort’ on this issue or whether the courts themselves will again arbitrate it. The response then turns to the issue of how latent print conclusions can be reported in the wake of the NAS Report. The Report expresses clear disapproval of the reporting framework currently mandated by latent print professional organizations, creating a tension around the reporting of analyses. The response concludes that semantic resolutions to this tension are undesirable compared to resolutions based on empirical data.

Keywords: National Academy of Science; forensic science; fingerprint; expert witnesses; general acceptance.

1. Introduction

Of the many topics addressed by the recent National Academy of Science (NAS) Report on forensic science, this response will focus on the forensic discipline I know best, a discipline, that enjoys great prominence in the report: latent print analysis. In Section 2, I will place the NAS Report, as it pertains to latent print identification, in historical context. I will argue that, by the time of the publication of the NAS Report, a robust controversy was underway that had become starkly polarized between scholars and scientists on the one hand and practitioners and judges on the other. In Section III, I argue that the significance of the NAS Report lies in its status as an intervention by a mainstream scientific institution in this ongoing controversy. As such, I suggest that the Report raises the question of ‘Who Speaks for Science?’—i.e. whether it will ultimately be judges or scientists who have the last word on the issues raised by latent print analysis. I suggest that the resolution of this issue implicates the broader issue of to what extent scientific institutions like the NAS have ‘jurisdiction’ over scientific matters that arise in legal settings. In Section IV, I assume that the NAS Report will

† Email: scole@uci.edu

© The Author [2009]. Published by Oxford University Press. All rights reserved.
be treated as authoritative, and I ask a different version of the question ‘Who Speaks for Science?’ How, I ask, will the results of latent print analyses be reported, given that the NAS Report states that the current mode of reporting is unacceptable? I explore some of the possible ways that latent print conclusions may be reported, and I conclude that the reporting of conclusions is a crucial, and underexplored, issue for forensic science.

2. Background

To situate the Report in historical context, insofar as it addressed latent print identification, the Report marked the first intervention by a credentialed mainstream scientific institution into a debate that had been going on for at least a decade. Latent print identification was introduced to the United States during the first decade of the 20th century, and it was first deemed admissible evidence in a reported decision in 1911 (*People v. Jennings*, 1911). Only in the last decade of the 20th century, however, were serious questions raised about the validity and admissibility of latent print analysis. During the 1990s, a number of scholars raised questions about the validity of latent print analysis (e.g. Berger, 1994; Faigman et al., 1997; Woodworth, 1997; Cole, 1998; Saks, 1998; Starrs, 1999; Mnookin, 2001; Stoney, 2001). In 1999, these questions resulted in a legal challenge to the admissibility of latent print evidence (*United States v. Mitchell*, 2004) under the Supreme Court’s *Daubert* and *Kumho Tire* standards for expert evidence (*Daubert v. Merrell Dow Pharmaceuticals*, 1993; *Kumho Tire v. Carmichael*, 1999). *Kumho Tire* mandates that the trial court assess the reliability of proffered expert evidence. In such proceedings, the burden is clearly on the proponent of the evidence to provide evidence of reliability (e.g., Faigman et al., 1997; Mearns, 2009). The defense, conceding the uniqueness of all human friction ridge skin (*United States v. Mitchell*, 1999a), contended that there was no evidence of reliability.

Whatever may be said about the accuracy of latent print identification, the arguments mounted in its defense were surprisingly weak, indeed irrational. At the *Daubert* hearing in *Mitchell*, the government contended, among other things, that the reliability of latent print identification was demonstrated by the uniqueness of friction ridge skin and the longstanding use of the technique in casework and criminal trials (what would later become known derisively as ‘adversarial testing’) (*United States v. Mitchell*, 2004, p. 238), that the error rate of the technique could be meaningfully parsed into ‘methodological’ and ‘human’ categories (p. 240) and that the ‘methodological error rate’ was zero (p. 227). Mitchell replied that one could not infer the accuracy of the technique from the uniqueness of its target of analysis or from casework or trials, that there was no meaningful distinction between the examiner and the method, and that the error rate, though unknown, was certainly not zero (Epstein, 2002).

Over the ensuing decade, this debate was also joined in the scholarly literature. Most scholars agreed with Mitchell that validation studies of latent print identification were lacking (e.g. Berger, 1994, p. 1354; Faigman et al., 1997; Haber and Haber, 2003, p. 358; Kennedy, 2003; La Morte, 2003, p. 187; Lawson, 2003, p. 65; Cole, 2004, p. 1205; Schwinghammer, 2005; Siegel et al., 2006, p. 35; Dwyer, 2007, p. 391; Mnookin, 2008; Saks & Faigman, 2008, p. 152), that one could not infer the accuracy of the technique from the uniqueness of its target of analysis (e.g. Robertson, 1990, p. 255; Stoney, 1997; Starrs, 1999, p. 243; Champod and Evett, 2001, p. 115; Epstein, 2002, p. 613; Kaye, 2003b, p. 1080; Lawson, 2003, p. 9; Benedict, 2004, p. 528; Cole, 2004, p. 1199; Moriarty, 2004, Section 12, p. 24; Zabell, 2005, p. 163; Saks & Faigman, 2008, p. 155) or from casework or trials (e.g. Saks, 2000; Mnookin, 2001, p. 65; Faigman, 2002; Giannelli, 2002; Haber
and Haber, 2003, p. 343; Kaye, 2003b, p. 1081; Benedict, 2004, p. 529; Cole, 2004, p. 1210; Zabell, 2005, p. 168; Siegel et al., 2006, p. 44; Koehler, 2008, p. 1085), that latent print examiners could not support their claims to be able to accurately ‘individualize’ latent prints (e.g. Stoney, 1997, p. 72; Saks, 2000, p. 881; Mears & Day, 2003, p. 731; Cole, 2006; Mnookin, 2008; Saks & Faigman, 2008, p. 152), that there was no meaningful distinction between the examiner and the method for calculating error rates (e.g. Mnookin, 2001, p. 60; Kaye, 2003b, p. 1083; Cole, 2005, p. 1039; Zabell, 2005, p. 177; Koehler, 2008, p. 1088; Saks & Faigman, 2008, p. 159) and that the error rate, though unknown, was certainly not zero (e.g. Starrs, 1999, p. 243; Saks, 2000, p. 885; Mnookin, 2001, p. 59; Epstein, 2002, p. 633; Haber and Haber, 2003; La Morte, 2003, p. 184; Lawson, 2003, p. 43; Mears & Day, 2003, p. 732; Cole, 2005; Zabell, 2005, p. 178; Siegel et al., 2006, p. 40; Cooley, 2007, p. 390; Saks & Faigman, 2008, p. 159). Some scholars and forensic practitioners, however, agreed with the government that the reliability of latent print identification was demonstrated by the uniqueness of friction ridge skin (Moenssens, 2003) and the longstanding use of the technique in casework and criminal trials (Moenssens, 2002), that the error rate of the technique could be meaningfully parsed into methodological and human categories and that the methodological error rate was zero (testimony of Stephen Meagher and Bruce Budowle; United States v. Mitchell, 1999b).

What was the status of this debate prior to the NAS’s intervention? The questions raised about latent print identification provoked divergent responses from courts and from scholars (Haber and Haber, 2008; Mnookin, 2008). Courts deemed latent print evidence admissible with near unanimity (Faigman et al., 2007). In contrast, almost all scholars who addressed the issue concluded that latent print examiners’ claims to be able to individualize—that is to determine the source of a latent print of unknown origin to the exclusion of all other possible sources in the universe (SWGFAST, 2002)—were not supported by evidence and, further, that measurements of the accuracy of latent print analysis were lacking.

In 2005, a group of scientists and scholars (including the author) working with the New England Innocence Project filed an amicus curiae brief in an appeal of latent print admissibility to the Supreme Judicial Court of Massachusetts (Siegel et al., 2006). This brief, signed by more than 15 scientists and scholars from a wide variety of disciplines, including members of the NAS, stated that the reliability of latent print individualization had not been demonstrated,. It showed that the arguments that had previously been made by various scientists and scholars were not those of an extreme few, but rather the views of almost every non-practitioner scientist or scholar who had examined the question. However, despite having stated in an earlier case that ‘the relevant scientific community’ should ‘be defined broadly enough to include a sufficiently broad sample of scientists so that the possibility of disagreement exists’, not ‘so narrowly that the expert’s opinion will inevitably be considered generally accepted’ the court ruled that the trial court had acted within its discretion in limiting the relevant scientific community to latent print examiners only (Commonwealth v. Patterson, 2005, p. 25, quoting Canavan’s Case 2000).

One way of characterizing the status of the debate prior to the NAS’s intervention is to ask what we might call ‘the Frye question’. Although the Report, like most evidence scholarship today, focuses on the more recent Daubert/Kumho approach to expert evidence, the ‘general acceptance’ approach to assessing expert evidence espoused in Frye v. United States (1923) is still adhered to in many state courts, espoused by some evidence scholars (Schwartz, 1997), and incorporated into Daubert and Kumho themselves. The Frye approach asks whether a claim made by an expert is ‘generally accepted in the relevant scientific community’. Among the most prominent of the many criticisms of this approach are that this supposed ‘test’ only raises further questions, including who
should constitute the relevant scientific community and what constitutes general acceptance. Frye has also been criticized as crude ‘head counting’ (e.g. People v. Leahy, 1994; Brim v. State, 1997). However, other courts have defended head counting as the appropriate way to measure general acceptance (e.g. Jones v. United States, 1988; Goeb v. Tharaldson, 2000).

The year before the release of the NAS Report, I undertook the sort of crude head count envisioned by at least some interpretations of the Frye approach. In doing so, I should note that I was not necessarily endorsing the position that courts should make decisions about the admissibility of evidence based on crude head counts. Rather, I was using the device of the head count to convey a general sense of the state of opinion within the scholarly community on this issue. Although a head count in a particular direction might not necessarily be dispositive on the issue of admissibility, it might be considered a sort of ‘red flag’. I surveyed three sources to measure the opinions of scientist and scholars on the question of whether latent print examiners’ claims of individualization had been validated: sworn expert testimony, amicus curiae briefs and published literature. The results showed that the vast majority of scholars had found evidence validating latent print examiners’ claims to accurate individualization lacking. A full discussion of this survey may be found elsewhere (Cole, 2008b), but, in sum, a rough ‘head count’ of non-practitioner scientists and scholars yielded 25 who had expressed the view—through sworn testimony, signing an amicus curiae brief, publishing an article or a combination of those means of expression—that latent print individualization lacked validation and three who had, through the same mechanisms, stated the contrary. As noted elsewhere (Cole, 2008b), it possible to quibble with the numbers by eliminating some individuals for a various reasons (e.g. disqualifying those who have testified as expert witnesses or disqualifying legal scholars for not being ‘scientists’), but none of these alternative interpretations supports any conclusion other than that very few non-practitioner scientists or scholars had made public statements holding that latent print individualization had been validated. This analyses showed that, while some of the earliest scholarly arguments had been dismissed as the misinformed opinions of ‘a handful of critics’ (Moenssens, 2005) ‘on the margins of a professional discipline’ (Moenssens, 2002) with a ‘self-serving agenda’ (IAI, 2007b), by 2008 the point that latent print individualization lacked validation had been articulated by a large number of scholars from a wide variety of disciplines who were harder to dismiss as being motivated by financial interests or having a ‘self-serving agenda’.

This head counting exercise illustrated the starkly different results that can be generated depending upon whether a court treats as the ‘relevant scientific community’ for a particular forensic technique the practitioners of the technique itself or the broader scientific community (Faigman et al., 1997). More specifically, it showed that, although latent print practitioners themselves attested to the reliability of their analyses, hardly any external observers, with the important exception of judges, had been convinced of this reliability. Contrary to the representation made by the International Association for Identification (IAI) to the NAS Committee, that ‘scientists and legal authorities . . . generally’ hold that ‘the error rate for fingerprint identification is extremely small, statistically insignificant, and not due to the methodology but instead to the inherent risk of error in any human endeavor’ (IAI, 2007b) (without naming any scientists or legal authorities), it would seem that few, if any, non-practitioners scientists or scholars hold this view.

However, the view in the courts was the opposite; nearly every court found latent print evidence admissible, and most of those, in doing so, stated that it was valid. In their efforts to find latent print identification reliable, courts stretched the limits of rational argument. Consider, e.g. the courts that wrote opinions stating that the reliability of latent print identification was established by having
been used in court for 100 years (United States v. Havvard, 2000), for having withstood ‘the test of time’ (United States v. Crisp, 2003) or for having survived a regimen of ‘implicit testing’ (United States v. Mitchell, 2004). Consider the courts that declared that the error rate of latent print identification was ‘vanishingly small’ (United States v. Havvard, 2000, p. 854), ‘essentially zero’ (United States v. Havvard, 2001, p. 599), ‘negligible’ (United States v. Crisp, 2003, p. 269) or ‘microscopic’ (United States v. Mitchell, 2004, p. 241, n. 220), assertions made without reference to any empirical data whatsoever or to what actual numerical error rates led to these verbal characterizations. Consider the courts that approvingly cited and relied upon the notorious, unpublished ‘50K study’ (later criticized in the NAS Report (p. 144)) (United States v. Mitchell, 2004, p. 237; State v. Kim, 2004, p. 7). And yet, at the same time, neither these courts nor any other have ever cited any of the many published—some peer-reviewed—articles critiquing this study (Wayman, 2000; Champod and Evett, 2001; Stoney, 2001; Pankanti, 2004, p. 7). Instead, in an extraordinary feat of legal reasoning, one court ruled that latent print identification satisfied the ‘peer review and publication’ prong of Daubert and Kumho because the 50K study, though never published itself, had been severely criticized in peer-reviewed publications (State v. Sullivan, 2005)!

Thus, latent print admissibility rulings do not appear to have been the finest hour of either legal reasoning or scientific reasoning by judges. Why? Judge Edwards, Co-Chair of the NAS Committee, when asked this question by Julia Leighton at the symposium that occasioned this journal issue, suggested (speaking for himself, not for the Committee) that judges, in perhaps ‘too crude’ terms,

1 The author of the Havvard opinion recently became the first nominee to a Circuit Court judgeship by the Obama Administration, which has promised to ‘develop a strategy for restoring scientific integrity to government decision-making’ and ‘base our public policies on the soundest science’. In noting this mild irony, I am certainly not disputing Judge Hamilton’s fitness for the position for which he was nominated because it would not be appropriate to evaluate a judge based on a single decision. Moreover, the more relevant lesson here might well be that even good judges have had difficulty constructing scientifically defensible arguments in support of latent print identification.

The Report’s idiosyncratic reading of the Havvard case (pp. 103–104) must also be noted. The Report seems at pains to absolve both Judge Hamilton and the government’s expert for what it appears to view as a poor opinion by the Seventh Circuit. Instead, the Report blames the Seventh Circuit for allegedly ‘overstating[ . . . ] the expert’s equivocal testimony’, misreading Judge Hamilton’s supposedly more accurate representation of that testimony and ‘giv[ing] fuel to the misconception that the forensic discipline of fingerprinting is infallible’.

We do not know what, precisely, the expert witness testified to in Havvard because neither the Report nor any of the scholars who have written about the case (including the author) has obtained the transcript. In the absence of a transcript, however, one can reasonably infer that the testimony given in Havvard was substantially similar to the testimony the same witness gave in United States v. Mitchell, not long prior, for which a transcript has been made publicly available. In Mitchell, the witness testified as follows:

Well, when you’re dealing with a scientific methodology such as we have for ever since I’ve been trained, there are distinctions—there’s two parts of errors that can occur. One is the methodological error, and the other one is a practitioner error. If the scientific method is followed, adhered to in your process, that the error in the analysis and comparative process will be zero. It only becomes the subjective opinion of the examiner involved at the evaluation phase. And that would become the error rate of the practitioner (United States v. Mitchell, 1999b, p. 154).

Based on that expert’s unrebutted testimony, the District Court in Havvard wrote that ‘[The expert] testified that the error rate for fingerprint comparison is essentially zero. Though conceding that a small margin of error exists because of differences in individual examiners, he opined that this risk is minimized because print identifications are typically confirmed through peer review’.

Given that the witness stated that the error rate of latent print identification was ‘zero’ and the District Court credited this by stating that it was ‘essentially zero’, it is puzzling why the Report would describe the Seventh Circuit’s characterization of the error rate as ‘essentially zero’ as an ‘overstate[ment]’, suggest that the appellate court was responsible for ‘the misconception that the forensic discipline of fingerprinting is infallible’ or characterize the witness’s testimony as ‘equivocal’.

2 Julia Leighton is General Counsel at the Public Defender Service for the District of Columbia.
had ‘been snookered’—that ‘we just didn’t get it’ (Edwards, 2009, video at 1:02:10). Some commentators have suggested that judges did not apply the same degree of scrutiny to evidence offered by the government in criminal cases that they applied to evidence offered by plaintiffs in civil cases, out of concern for allowing accused criminals to go free (Risinger, 2000; Neufeld, 2005; Cooley and Oberfield, 2007, p. 285; Rozelle, 2007, p. 597). A subtler psychological mechanism has also been suggested, in which judges recognize forensic evidence as, in a sense, their own creation and thus experience cognitive resistance to the notion that their faith in it was perhaps over-hasty (Cole, 2004, p. 1275; Risinger, 2007, p. 473).

However, one might equally well argue that judges simply have a strong intuition that latent print identification is highly accurate and that they are reluctant to overrule this intuition in deference to either legal or scientific formalism. Judges, of course, may well be correct in this intuition. It does not, of course, justify countenancing the refusal to provide any data to support this intuition. Nor does it justify the institutionalized exaggeration of the probative value of the evidence implied by the requirement of testimonial reports of ‘individualization’, reports that overstate the probative value of the evidence, regardless of the accuracy of the technique. Note, however, that by not generating accuracy data and not modifying the testimonial claim of individualization, the government and practitioners put judges in a quandary, in which they must either admit the evidence despite its exaggerated testimonial claim or exclude it despite their intuition of its accuracy—unless, i.e. they opt for the solution of limiting, not admissibility, but the testimonial claim (e.g. State v. Pope, 2008).

If judges have been correct in their intuitions about the high accuracy of latent print identification, then the material damage wrought by their admissibility rulings would be limited to those relatively few individuals who were erroneously implicated by evidence that was presented to jurors—with the courts’ approval—as highly accurate. However, it might also be argued that larger damages have been wrought by the past decade of controversy over latent print identification. Even if one sets aside the ultimate accuracy of latent print identification, one is still left with the nagging facts that courts do not appear to have applied their own admissibility standards in an evenhanded manner; that, when making such rulings, they did not take the trouble to support them with either legal or scientific reasoning that would be recognizable as high quality; and that the courts disregarded or dismissed the views of scientists and scholars on a scientific issue. These are damages wrought to the dynamic relationship between the institutions of science and law (Jasanoff, 2005) and even, potentially, to legitimacy of the courts.

Judicial opinions such as these had an often overlooked but crucial consequence: the judicial opinions themselves came to proxy for the very validation studies whose absence was being debated (Cole, 2004, p. 1275; Risinger, 2007, p. 467). When asked for evidence of the reliability of latent print identification, prosecutors, and even courts themselves, tended to point to court opinions rather than empirical data or studies. In one case discussed in the Report (p. 104), the court referred to ‘the consensus of the expert and judicial communities that the fingerprint identification technique is reliable’ (United States v. Crisp, 2003, p. 269, emphasis added), thus bypassing the external scientific community altogether in instilling legitimacy on the technique. Courts’ notions that they themselves can confer validity on evidence is well illustrated by the following exchange from an admissibility hearing (the author is ‘The Witness’ in this exchange):

THE COURT: Except we do have at least a hundred years of case law where the courts have found that over the passage of time there is some validity to this process called fingerprint identification.
THE WITNESS: I think that scientists and scholars would not view case law as scientific data.

THE COURT: But there is a certain idea—Every system of analysis appreciates a continued respect for a process in any kind of forum or arena; right? If it’s continually to be relied upon as a credible instrument of identification, that fact, I think, has some bearing on any kind of contemporary assessment of its validity in a proceeding like this; isn’t that right?

THE WITNESS: Well, respectfully I think that one of the principles of science is not to have respect for continuing institutions just because they’re continuing.

THE COURT: Well, if the courts have found that there was sufficient challenge to a process, they would say that the evidence is not admissible as in the case of the lie detector; isn’t that right? . . . If the courts over the passage of time found that a particular process of scientific information like lie detector polygraph machines, if they found that that wasn’t a valid process, they would say it’s inadmissible, they would say it’s not going to be considered, it’s not satisfying what we call a *Kelly* or *Daubert* standard, which I’m sure you’re familiar with. We don’t have that here; do we? Fingerprint analysis hasn’t been rejected by either *Daubert* or *Kelly* or *Frye*, or whatever state jurisdiction you’re relying upon; isn’t that right?

THE WITNESS: That’s correct, but to a scientist that does not trump the absence of data. The fact that courts have accepted it for nearly a century doesn’t trump the absence of data.

THE COURT: But you’re not making the proposition that all scientists have decided as a universe of individuals to reject fingerprints as a proof of anything. You’re saying a group of scientists, including yourself, have rejected fingerprints but not all scientists.

THE WITNESS: No, of course I could never say that because they have not rendered an opinion.

THE COURT: Right.

THE WITNESS: But what I am saying is that if you look at what scientists have said, almost all of them are falling on the side of this claim has not been proven. Almost none of them are saying the opposite.

THE COURT: Well, the *Llera Plaza* case you identified there were two out of five experts that testified, three said one thing and two said something different.

THE WITNESS: That’s right. And the three—the two—two of the five of those people were scientists, and they both were testifying for the defendant.⁴

THE COURT: Yes.

THE WITNESS: The other witnesses were latent print examiners.

THE COURT: Yes. Who the judge essentially found to be apparently more credible.

THE WITNESS: Apparently so.

---

³ *People v. Kelly* (1976).
⁴ The expert witnesses who testified at the *Daubert* hearing in *Llera Plaza* were Janine Arvizu, a quality assurance auditor; Ralph Haber, a psychologist; and Allan Bayle, Stephen Meagher and Kenneth Smith, who were latent print examiners (*United States v. Llera Plaza*, 2002).
The NAS Report, in a crucial passage that should not be overlooked, uses the disturbing term ‘judicially certified’ for this phenomenon (p. 86). It is important to emphasize what the Report is saying here: it is saying that judicial opinions have emerged as an alternative source of legitimacy for forensic techniques that are unable to derive legitimacy from normal scientific channels. Thus, one leading latent print practitioner, when asked on national television about the reliability of latent print identification in 2003 replied, ‘We’re winning 41 times out of 41 [admissibility] challenges. I think that says something’ (Fingerprints , 2003). This same practitioner recently co-authored an article acknowledging that ‘for many years the forensic science community has pointed to successful admissibility of its science findings, and the opportunity to cross examine expert witnesses, as support of a technique’s “general acceptance” and “reliability.”’ Remarkably, however, he appears to have now disavowed this rhetorical strategy, warning ‘philosophically we do not advocate successful admissibility as demonstrating good science’ (Budowle et al., 2009, p. 799). Perhaps this turnaround was prompted by the Rose case (see below) which rendered it no longer possible to argue that latent print identification always won admissibility challenges. However, because common law is based on precedent, this turnaround is disturbing. By the time forensic practitioners got around to disavowing the tactic of treating admissibility hearing victories as de facto evidence of both ‘general acceptance’ and ‘reliability’ in 2009, they had amassed a large body of legal opinions rendering many forensic techniques, especially latent print identification, legally admissible. Future courts asked to rule on admissibility challenges may well simply rely on precedent, despite practitioners’ belated concession that admissibility hearing victories do not constitute empirical evidence of a technique’s reliability.

In this period, during which what we might call the ‘scientific’ perspective had won nearly unanimous acceptance among scientists and scholars but unanimous rejection in the courts, the National Academies may have come to seem like a ‘court of last resort’ for this controversy. It seemed clear that the courts would continue to support the claims that latent print identification had satisfied the Kumho reliability requirement and continue to construe the relevant scientific community as practitioners, rather than the broader scientific community. It also seemed that the shelf life of admissibility challenges was nearing expiration as courts were increasingly taking the position that the issue had been sufficiently litigated and defendants were no longer entitled to admissibility hearings (United States v. Mitchell, 2004, p. 246). This was disturbing because it seemed that the courts, as essentially the sole or main ‘customer[s]’ (Meagher, 2007) of latent print analysis, were the only institutions capable of forcing change or improvement upon this particular forensic practice. The National Academies had intervened effectively in prior controversies over other forensic techniques (National Research Council, 1992, 1996, 2004) and thus seemed to have a special authority in controversies over forensic science, and so it seemed that only the National Academies could put an end to scientifically and legally questionable claims like the ‘zero error rate’, ‘adversarial’ and ‘implicit’ testing, and severe criticism constituting positive peer review.

It should, however, be noted that while the NAS Committee was conducting its study, one court did hold the government to its burden to demonstrate the reliability of the evidence it proffered.

---

5 The witness in this exchange was the author of this article. As the Report notes (p. 97), it is difficult to get a ‘clear[] picture’ of how trial courts ‘handle[] . . . judicial dispositions of Daubert-type questions’ because most cases are decided without published opinions. My experience consulting on admissibility motions gives me access to some anecdotal information about what occurs on the trial court level, information that may not be available to other observers without those experiences. In the interest of better informing readers about what may be occurring in the myriad trial courts across the United States, in this article, where relevant, I draw on the transcripts of some of those hearings.
Recognizing that the government’s proffer at the admissibility hearing concerning the formation and individuality of friction ridge skin (the anatomical structure that creates finger, palm and sole prints) was unresponsive to the issue of the reliability of latent print individualization, Judge Souder deemed latent print evidence inadmissible (State v. Rose, 2007). This lone ruling, however, had no precedential value for any other court. Indeed, Souder’s ruling in Rose, which was discussed with apparent approval in the NAS Report (p. 105), was recently even further mooted when the government refiled the case in a federal court, which then deemed latent print evidence admissible and reliable without holding an evidentiary hearing (Tamber, 2009). This gave rise to the unlikely spectacle of the government forum shopping in order to use fingerprint evidence. Whether—without the NAS Report—Rose would have constituted an anomaly or the harbinger of a new trend is a counterfactual question whose answer we will never know. Similarly, whether change would have occurred in latent print practice without the NAS Report is something we will never know.

3. Who speaks for forensic science?

This was the unusual context into which the NAS Report was issued in February 2009: stark disagreement between scholars and scientists on the one hand and courts and forensic practitioners on the other. At issue was, in some sense, who—the scientific community or the courts—would have the last word on an issue that is clearly a matter of ‘science’, but science that is primarily, if not exclusively, used in legal settings.

The Report wholly endorsed what we might call the ‘scientific’ position. It repudiated the notion of ‘adversarial testing’ (p. 42), claims of infallibility (p. 87) and the claim of a ‘zero error rate’ (p. 142). It rejected the claim that the reliability of latent print individualization could be inferred from the uniqueness of friction ridge skin (pp. 43–44, 144). It stated that no forensic assay other than nuclear DNA profiling could support claims of individualization, an assessment that included latent print analysis (p. 87). And, it stated that there was only ‘limited information about the accuracy and reliability of friction ridge analyses’ (p. 142).

In terms of the ‘head counting’ exercise discussed above, the NAS Report was highly significant. Although it adds 17 or so names to the list of informed external observers who have adopted the broad ‘scientific’ position, more important was the endorsement of a scientific institution and not, of course, just any scientific institution, but the NAS. Most important of all is the fact that the Report represented the outcome of the standard NAS Report-producing process. Thus, the process of convening and vetting a committee composed from ‘the nation’s foremost scientists, engineers, health professionals, and other experts’ (National Academies, 2009) with no vested interest in the controversy and exposing them to the available data, literature and information yielded the conclusions that they did. Thus, the issuing of the NAS Report rendered the already remarked upon discrepancy between the legal and scientific communities as to the demonstrated reliability of latent print individualization even starker than it already was. It put the weight of a scientific institution, rather than merely the accumulated weight of individual scientists and scholars, behind the ‘scientific’ position. The Report would seem to speak not only for the Committee but also for the NAS and not only for the NAS for science itself (Hilgartner, 2000, p. 88). Moreover, it again demonstrated the difficulty latent print examiners have had convincing any external observers, other than judges, that they have validated their claims to be able to individualize latent prints.

Nonetheless, it raises the question: does it matter? I am not merely being facetious. By my estimate, the NAS Committee included only one member who has probably ever analysed a latent
print. Since previous interventions by scientists and scholars had been dismissed by practitioners as the efforts of ‘a handful of critics, mostly academics and law professors who, beyond literature searches, have not acquired any practical experience in comparing prints’ (Moenssens, 2005), should or will the NAS Report be dismissed on the same grounds? If not, we must ask hard questions about why scientists and scholars were not heeded until the NAS Report. We must take note of the fact that the courts left here a record of ignoring scientists and scholars until the intervention of the NAS, a record that led the Committee to characterize the courts as ‘utterly ineffective’ in dealing with the lack of validity and accuracy data in many forensic disciplines (p. 109). This raises discomfiting questions about judicial dealings with science—since surely we cannot rely on the NAS to correct all such situations. What other scientific matters are being decided by courts in this manner without remediation through an NAS Report? Indeed, even the present Report had a long and tortured pre-history before it was officially commissioned (Kennedy, 2003), and it might easily never have seen the light of day. What would be the legal status of latent print identification and other forensic disciplines had this Report never been commissioned?

If, on the other hand, the NAS Report is dismissed on the same grounds as previous scientific interventions, what then? We will still have a standoff between law and science, but now we will have not just lawyers, judges, scientists and scholars involved but scientific institutions as well. The supposed ‘court of last resort’ will turn out not to have been a court of last resort at all. Further adjudication will be necessary. And where will this adjudication occur? In the courts, as defendants leverage the Report in challenges to forensic evidence. Thus, the ‘court of last resort’ for forensic science may again be the courts themselves.

4. Speaking for evidence

Assuming then that the controversy does return to the courts and that the Report is treated with the authority that it deserves, what will happen in the courtroom? There will be admissibility challenges to be sure. But there will also be challenges to the nature of latent print testimony itself. Many evidence scholars have taken the position that courts and scholars alike have focused too much on the underlying validity of forensic techniques and perhaps not enough on an equally important issue: the nature of forensic testimony (Beecher-Monas, 1999, p. 1062; Berger, 2003, p. 1140; Black, 2003; Friedman, 2003, p. 1063; Gross & Mnookin, 2003, p. 143; Nance, 2003, p. 253; Faigman, 2004, p. 258; Imwinkelried, 2004, p. 277). ‘Forensic’ science, after all, is literally ‘science that speaks’—in a courtroom, to a factfinder, which is usually, in the United States, a jury of laypersons. I have elsewhere proposed putting the ‘forensic’ back into forensic science and devoting some serious thought not merely to the dichotomous question of whether particular expert witnesses should or should not be permitted to testify before a jury but also to the continuous question of what evidentiary claims they should be permitted to make before that jury (Cole, 2007). The Report takes a similar position in its Recommendation 2: that the proposed National Institute of Forensic Science ‘should establish standard terminology to be used in reporting on and testifying about the results of forensic science investigations’ (p. 22).

Let us consider how this recommendation would play out for latent print evidence. Latent print analysis is unusual in that, unlike many other trace evidence disciplines, it has taken a quite clear position on the way in which testimonial conclusions are framed, at least in the United States.  

---

6 I am thinking of Jay Siegel.
According to both the Scientific Working Group on Friction Ridge Analysis, Science & Technology (SWGFAST) and the IAI, only three testimonial conclusions are permitted: exclusion, inconclusive and individualization (IAI, 1979, 1980; SWGFAST, 2003b). ‘Individualization’, therefore, is the only permissible conclusion that implicates a defendant. Individualization has been defined as ‘The determination that corresponding areas of friction ridge impressions originated from the same source to the exclusion of all others (identification)’ (SWGFAST, 2003a). As numerous commentators (e.g. Stoney, 1997, p. 70; Champod and Evett, 2001, p. 113; Epstein, 2002, p. 612; Cole, 2004, p. 1196; Zabell, 2005, p. 155; Champod, 2008, p. 117; Mnookin, 2008) and the Report (p. 142) have noted, such categorical conclusions are problematic.

The main issue is that such conclusions require knowledge about the rarity of observed features in the relevant population. Even if estimates of rarity are available, as, e.g. in the case of DNA profiling, a conclusion of individualization involves a process of rounding off a small probability into a probability of zero that is controversial even for DNA profiling (Buckleton, 2005). However, in the case of latent prints, individualization becomes even more problematic because rarity estimates are not derived from formal, systematically collected data, but rather intuited by the examiner (Champod et al., 2004; IEEGFI II, 2004; Thompson & Cole, 2007).

Recognizing this, the Report clearly states that claims of individualization are not supported for any forensic assay other than DNA profiling (p. 87). Thus, claims of latent print individualization are not supported. In this, the Report is not alone. At least one court that admitted latent print evidence precluded testimony of individualization (State v. Pope, 2008). Here one wishes that the Report had gone further and noted the bankruptcy of individualization as a concept, no matter what the assay (Biedermann et al., 2008; Champod, 2008; Saks & Faiqman, 2008; Saks & Koehler, 2008; Champod, 2009; Cole, 2009). DNA profiling has shown that it is possible to have highly probative evidence without rounding down the probability of the alternative hypothesis for the factfinder, and this renders individualization ‘at least conceptually, needless’ (Biedermann et al., 2008, p. 130).

What, however, does the future hold for latent print testimony? We now have a situation in which latent print examiners are restricted by professional guidelines to giving testimony which, according to the NAS Report, they cannot support. Among the high points of the Symposium that generated this journal issue was the following exchange:

**WILLIAM THOMPSON**\(^8\): There’s some questions about how forensic scientists should be characterizing their findings at this point in time in light of the Report. For a long time, latent print examiners and tool mark examiners and others who examine marks have been coming into court and testifying that they can identify the source of the mark to the exclusion of all other sources and with complete certainty. Do you think that kind of testimony is justified at this point in time pending further research, and, if not, what should forensic scientists be saying about that kind of examination at this point in time? And what’s the FBI going to be saying about it?

**CHRISTIAN HASSELL**\(^9\): Those absolute individualization statements are no longer a part of the practice, especially the tool marks, the pattern based areas.

**WILLIAM THOMPSON**: What about latent prints?

\(^7\) Granting individualization to DNA profiling was an odd decision given that many DNA scientists do not even claim to be able to achieve individualization, and, even for those who do, DNA individualizations must always be accompanied by the caveat, ‘barring monozygous twins’, as SWGFAST (2009, p. 3) points out.

\(^8\) William Thompson is a Professor of Criminology, Law & Society at the University of California, Irvine.

\(^9\) Christian Hassell, an analytical chemist, had been named Director of the FBI Laboratory in June 2008.
CHRISTIAN HASSELL: I have two of my examiners here in the room so I... I'm not sure, I can't remember what we do exactly. Like I say, I'm not the expert in these areas (Forensic Science for the 21st Century, 2009, Disk 4, Q&A section, at 0:00–01:20).

MELISSA GISCHE: As far as fingerprinting is concerned, we do still identify, we will effect individualizations, to the exclusion of all others based on the science behind our examinations.

WILLIAM THOMPSON: And does the NRC Report’s statement about the lack of science in this area change your impression about how you should be—

CHRISTIAN HASSELL: No, because we don’t agree with the lack of science. We’ve got—

MICHAEL RISINGER: For individualization?

UNIDENTIFIED: Individualization?

MICHAEL RISINGER: For perfect individualization, you don’t agree with the lack of science?

CHRISTIAN HASSELL: Oh, not perfect, but—

MICHAEL RISINGER: Well, that’s what she said.

UNIDENTIFIED: She used the words ‘to the exclusion of all others’.

HENRY LEE: We do have chances of perfect individualization. Just now I give an example, in [inaudible] case I have, they leave the bumper behind with the license plate. In my opinion, that’s a perfect individualization. I have no other cases with that license plate, with the bumper, that the car lost the bumper and license plate. That’s a perfect individualization example. Of course, it does not address the fingerprint issue.

WILLIAM THOMPSON: Right.

[other questions asked of the panel]

JAY KOEHLER: I also have a question for Dr Hassell. I just want to get clarification on what Bill Thompson was asking. Is it the FBI’s position that an examiner, a fingerprint examiner, can identify the source of a fingerprint with 100% certainty? Is it their position that they can individualize a latent print? . . .

CHRISTIAN HASSELL: Ok, I don’t mean to weasel out of this, but I am admitting my naiveté in some of these areas, just being so new. That’s why I have two of the top people here with me. So I’m going to ask Melissa to answer the first part of your question because I think it’s very important. However, I will say . . . the statements made earlier that the areas of fingerprint science has no scientific foundation, that was in error and, in fact we talked about that at the break. I mean who

---

10 The remainder of the dialogue presented here was not recorded on the publicly available video record of the conference proceedings. With the help of Laurie Ralston, Sandy Askland and Jay Koehler, I obtained the raw footage of the ‘Question & Answer’ section of this panel discussion. Because not everyone involved in the discussion spoke into microphone and some people spoke simultaneously, some portions of the discussion were difficult or impossible to hear. What is presented here represents my best effort at a faithful transcription, with some assistance from Jay Koehler, edited for flow and clarity.

11 Melissa Gische is a latent print examiner for the FBI Laboratory.

12 Michael Risinger is John J. Gibbons Professor of Law at Seton Hall University.

13 Henry Lee is the Chief Emeritus for the Connecticut Department of Public Safety, Division of Scientific Services.

14 Jay Koehler is a Professor of Law and Business at Arizona State University.
made that statement and the basis for it? That’s just wrong. There’s quite a bit of work done on areas
of uniqueness, persistence, and others, so I’m going to, Melissa has the . . .

MELISSA GISCHE: If I can clarify, as far as the absolute certainty argument, it’s still the exam-
iner’s opinion that there is enough information present that they can identify the source of that print
to a single area of friction ridge skin. Now, as far as—I may as well cover this now—the zero error
rate claim, I’m not sure why this keeps getting so misinterpreted. Fingerprint examiners have not,
nor do we currently, claim to have a zero error rate. We know that errors have occurred. I think where
some of the confusion in the past has come is in trying to distinguish between a methodological ver-
sus a practitioner error.\textsuperscript{15} I think things have gotten all twisted up in that, but as examiners, errors
have been publicized. Any time the human is involved we know there is a potential for error, which
is why we have all the various quality assurance measures, and so on . . . it’s too long for me to
list here. But, as far as still individualizing, yes we still do it, based on the underlying principles of
friction ridge skin. Jay, does that answer your question?

JAY KOEHLER: Well, I still don’t know whether . . . so the opinions are made with 100% certainty?
Source opinions, source identification? You would say ‘I’ve identified this latent to Jay Koehler with
100% certainty’. Is that still what you would say?

MELISSA GISCHE: Again, it is my opinion as far as the certainty, I don’t know that I would ever
attribute a percentage to that. But in my opinion, when I identify a print, I am saying that there is, in
fact, enough information present to determine the source of that print.

WILLIAM THOMPSON: To the exclusion of all others in the universe?

MELISSA GISCHE: That is correct.

MICHAEL RISINGER: You might want to think about modifying that. You won’t lose that much,
and you won’t be fighting a losing battle over what the science can do.

Thus, by April 2009, the NAS Report by itself had not persuaded the FBI to modify the testimo-
nal claim of individualization, and a recent FBI publication continues to defend individualization
(Peterson et al., 2009). Similarly, SWGFAST—a body that is sponsored, though not controlled—
by the FBI (Grieve, 1999, p. 145) has continued to defend individualization. It issued a statement
‘respectfully disagree[ing]’ with the NAS’s assertion that only nuclear DNA analysis could support
claims of individualization, stating ‘History, practice, and research have shown that fingerprints can,
with a very high degree of certainty, exclude incorrect sources and associate the correct individual
to an unknown impression’ (SWGFAST, 2009, p. 3). The statement confuses the issue of reaching
correct conclusions with the issue of making a scientific claim about the size of the potential donor
pool, and no data are cited to support or further specify the claimed ‘high degree of certainty’.

Whether courts will continue to accept such arguments remains to be seen. Unless latent print ex-
aminers and the litigants who employ them want to be cross-examined on the NAS Report, they may
need an alternative testimonial claim. Already a number of emerging alternatives can be discerned.
Let us examine them in turn.

\textsuperscript{15} Note that, as discussed above, the Report, as well as many scholars, reject this distinction. The FBI’s conception of this
distinction has been further elucidated by Peterson et al. (2009), who state:

Because latent print examinations do not employ instrumentation that can introduce systematic or random errors, the only
general type of scientific error in the latent print discipline is human error, also commonly referred to as practitioner error . . .

It is more accurate to say that the ACE-V methodology does not have a calculable error rate because it has no inherent error.
4.1 *Redefine individualization*

SWGFAST’s response to NAS Recommendation 2 states that SWGFAST ‘uses the development of our glossary to standardize terminology’ (SWGFAST, 2009, p. 6). The new SWGFAST Glossary removes the term ‘to the exclusion of all others’ and substitutes the term ‘conclusion’ for ‘determination’, thus redefining ‘individualization’ as ‘The conclusion that corresponding areas of friction ridge impressions originated from the same source’ (SWGFAST, 2008). Although removing the term ‘to the exclusion of all others’ may make the testimony less hyperbolic, the analyst is still attesting to something she does not know (or, in Bayesian terms, to a posterior probability)—that the two impressions derive from the same source. Redefining the statement as a conclusion, rather than a determination is perhaps a slight improvement, but still conveys the impression that the statement derives from an analysis of data rather than being what it really is: a ‘decision’ (Biedermann et al., 2008).

4.2 *Identification*

Another potential alternative is substituting the term ‘identification’ for ‘individualization’. In a recent admissibility hearing on latent print evidence, a member of SWGFAST disavowed the notion of ‘individualization’, conceded that it was unsustainable and offered in its stead a conclusion of identification, defined as ‘a decision given the relevant population . . . that the chance that someone could have left [the latent print] is so remotely small that he is willing to dismiss it and say yes, I believe that this latent print in my opinion was produced by that individual’ (Testimony of Glenn Langenburg, *State v. Hull*, 2008, p. 149). There ensued some confusion about whether individualization and identification are in fact distinct concepts given that the entry for ‘Identification’ in the SWGFAST (of which the witness was a member) Glossary reads ‘See Individualization’ (SWGFAST, 2002, 2003a). If they are synonymous, then nothing has changed. If they are distinct, then the witness violated the SWGFAST and IAI guidelines which ban inclusive conclusions other than individualization.

This confusion about whether these two terms are synonymous or distinct is common. For many forensic practitioners, ‘identification’ is used in a manner tantamount to ‘individualization’, as, e.g. in toolmark examiners’ definition of identification as ‘the likelihood another tool could have made the mark is so remote as to be considered a practical impossibility’ (Association of Firearms & Toolmark Examiners, 1998; Schwartz, 2005). However, ‘identification’ has been defined by some forensic scientists to indicate that the potential donor pool of a trace is greater than one, whereas ‘individualization’ is reserved for situations in which the donor pool has been reduced to one (Kirk, 1963; Inman and Rudin, 2001, p. 115; Thornton & Peterson, 2002, p. 8). The problem with this notion of ‘identification’ is that it is vague: has the donor pool been reduced to a class of a million potential sources or only two? In any case, this does not appear to be the usage proposed by the witness in *Hull*, which appears to be drawn more from the reasoning of Biedermann *et al.* (2008), in which the relevant donor pool contains only one member but the reduction to this pool is conceived as a decision rather than as, say, a conclusion. If this is indeed the case, the testimony’s status as decision would need to be made clear to the factfinder. The term ‘identification’ would not seem to do this job well; laypersons would seem to be apt to construe ‘identification’ to mean precisely what some forensic scientists claim is meant only by ‘individualization’, a possibility that is given added plausibility by the fact that, as discussed above, SWGFAST itself presents ambiguous information about whether the terms are synonymous.
Moreover, of course, questions might be raised about the empirical basis upon which the expert witnesses draws conclusions about the probability of false association. Currently, there is no basis other than the expert’s intuition based on her experience in casework (Thompson & Cole, 2007). This seems inadequate, given that experts do not systematically record and compile data about the rarity of features during casework and then retrieve that data to generate an estimate of the probability of false association. Even if such data were recorded, there would be difficulties extrapolating from the relatively small populations observed to the world population invoked by the decision described above (Zabell, 2005; Saks & Koehler, 2008). Resorting to ‘identification’ as a testimonial claim allows latent print practitioners to elude what, presumably, is one of the main objectives of efforts at forensic reform, as embodied by the NAS Report: generating rarity data and encouraging its use in formulating defensible reports.

4.3 Opinionization

A related possibility is that latent print reports could be rehabilitated by couching them as ‘opinions’, rather than ‘determinations’, ‘conclusions’ or ‘facts’. This is the solution urged by a recent report from the United Kingdom (Nuffield Council on Bioethics, 2007). We can also see signs of the ascendance of this approach in statements to the NAS Committee by the IAI (2007a). This solution also seems undesirable. For one thing ‘opinionization’ (Cole, 2008a) seems to function like a ‘universal solvent’ (Risinger, 2009), excusing all sins. (‘I was wrong? Sorry, it was only my opinion.’) Like the identification solution, opinionization allows forensic scientists to elude developing rarity data that would allow them to more precisely convey the probative value of the evidence.

4.4 Accuracy rates

One possible alternative that has yet to show signs of acceptance in the practitioner committee is the notion of attaching estimated accuracy rates to latent print examiners’ conclusions. As discussed above, historically latent print have conveyed accuracy rates to factfinders but these were assertions not derived from measurements: the claim that the ‘methodological error rate is zero, and the practitioner error rate’ is ‘vanishingly small’, ‘essentially zero’, ‘negligible’ or ‘microscopic’. Such claims were founded upon a supposed separation of the causes of error into two categories, ‘methodological’ and ‘human’ or ‘practitioner’. This separation was dubious, first, because, as the Report notes, there is no method without a practitioner (p. 143) (Giannelli, 2009); second, because the category of ‘methodological error’ was constructed in such a way that no errors would ever be assigned to it and, thus, it would remain, eternally, definitionally fixed at ‘zero’ (Peterson et al., 2009); third, because if the ‘methodological error rate’ is fixed as zero, it seems misleading to state it to the jury as if it were some sort of meaningful finding; and, fourth, because even if there were two categories of error the most meaningful single piece of information to convey to the consumer of the evidence would be the total error rate (Cole, 2005).

The NAS Report’s criticism of the zero error rate claim should render it extinct in the courtroom, and the IAI has now recommended that latent print examiners not make such claims in their testimony (Garrett, 2009). While all rational thinkers should rejoice at the disappearance of the zero error rate claim, it is important not to forget that it really was made in the first place (testimony of Stephen Meagher and Bruce Budowle United States v. Mitchell, 1999b), and it really was accepted by some courts (United States v. Havward, 2000). It is also important to note that the zero error rate was showing no signs of disappearing before the NAS intervened. I have seen affidavits asserting
a zero error rate as recently as 2006 (Declaration of Erik Carpenter, *United States v. Mikhel et al.*, 2006, p. 3), I have heard the claim made in admissibility hearings as recently as 2008 (Letter from Elisa J. Macken, *State v. Sheehan*, 2008), and I have heard it defended by practitioners even after the publication of the NAS Report and the IAI response.

If latent print examiners can no longer tell jurors that the error rate of latent print analysis is zero, can they tell them something else? The NAS Report states that there is only ‘limited information about the accuracy and reliability of friction ridge analyses’ (p. 142), so it does not seem like latent print examiners can reasonably offer either quantitative or qualitative estimates of their accuracy.

That does not, of course, mean that researchers will not eventually be able to generate responsible estimates of the accuracy of latent print examiners’ conclusion of common source that might be conveyed to jurors. This is one potential solution to the dilemma of latent print testimony that was discussed at the symposium under the label ‘black-box validation’ (Risinger *et al.*, 1998).

Another potential solution is the likelihood ratio approach being explored by Champod and his colleagues (Egli *et al.*, 2006; Meuwly, 2006; Neumann *et al.*, 2006, 2007). This approach does attempt to estimate the rarity of observed features in the relevant population. Although there was some debate at the symposium over these two approaches, it would seem that they are complementary, and neither piece of information (accuracy or rarity) would be sufficient without the other. It will be important to try to estimate the rarity of features, but, if analyses are to be carried out by human examiners, it will also be important to qualify that estimation with data on actual performance rates of comparable human examiners.

Any such efforts will need to be accompanied by educational efforts. On the one hand, one might need to train analysts accustomed to reporting conclusions in categorical terms to report in probabilistic terms. It will be important, however, to avoid the temptation to use computer software to enable expert witnesses to give testimony using impressive probability figures without a thorough understanding of those probabilities. Future analysts may require much greater education in statistics than in the past.

Finally, it should be noted that developing a scientifically supportable testimonial claim is only half the task, or perhaps even less than half, that lies before the justice system in order to make appropriate use of forensic evidence. There remains the task of figuring out how to effectively convey the probative value of forensic evidence to jurors. Numerous psychological studies have shown that jurors are innumerate and prone to fallacies in reasoning even when presented with testimonial claims that are themselves appropriately stated (e.g. Koehler, 2001). Thus, even the development of likelihood ratios for latent print comparisons, e.g. will not fully solve our problem because we know that jurors have great difficulty understanding likelihood ratios.

5. Conclusion

I have focused my remarks here on latent print identification, but they raise implications that apply to the rest of forensic science and, indeed, to the use of science in the courts more generally. The issue is not necessarily the quality of particular techniques, but rather the irrationality of the arguments that were mounted on their behalf—the fact that courts were, in the Report’s words, ‘utterly ineffective’ at detecting that irrationality, and the fact that scholars and scientists were largely deemed irrelevant to discussions of validity.

My concerns about the need for validation from some external scientific community, rather than relying merely on ‘self-validat[ion]’ (Black, 1988, p. 633) by a practitioner community or ‘judicial
certification’ by the courts apply to other disciplines as well. The NAS, the scientific community and indeed everyone should be concerned by the counterfactual that this Report might never have been commissioned. In that case, who, ultimately, would have spoken for science?

Likewise, the problem of how to report the probative value of conclusions about forensic evidence applies to other disciplines as well. Although other disciplines are less strict than latent prints about insisting only on unsustainable claims of ‘individualization’, many of them currently make unacceptably vague claims of ‘match’ or ‘identification’. All disciplines should abandon individualization as a telos (Cole, 2009) and strive towards using accuracy and rarity to report defensible claims about the probative value of forensic associations, fulfilling their mandate to speak, with as much ‘transparency’ (Champod and Evett, 2001) as possible, for science.

Conflict of interest disclosure

The author was expert witness and/or consultant in several of the cases discussed in this article, including United States v. Mitchell, United States v. Llera Plaza, People v. Carradine, State v. Rose, State v. Hull, State v. Pope, State v. Sheehan and State v. Sullivan. The author was a signatory to the amicus curiae brief filed in Commonwealth v. Patterson that is discussed in this article.

Funding

National Science Foundation (SES-0115305).

Acknowledgements

I am grateful to D. Michael Risinger, Jay Koehler and an anonymous referee for comments on this paper and to Jay Koehler, Michael Saks and David Kaye for inviting me to the Symposium on Forensic Science in the 21st century. I am grateful to Ian Tingen, Laurie Ralston, Sandy Askland and Jay Koehler for assistance obtaining video footage of the conference. Any opinions, findings and conclusions or recommendations expressed in this material are those of the author and do not necessarily reflect the views of the National Science Foundation.

REFERENCES

Brim v. State (1997) 695 So.2d 268 (Fla.).
WHO SPEAKS FOR SCIENCE?


FINGERPRINTS (2003). 60 Minutes (5 Jan.).


*People v. Jennings* (1911) 96 N.E. 1077 (Ill.).

*People v. Kelly* (1976) 549 P.2d 1240 (Cal.).

*People v. Leahy* (1994) 882 P.2d 321 (Cal.).


*United States v. Havvard* (2001) 260 F.3d 597 (7th Cir.).


*United States v. Mikhail et al.* (2006) CR 02-220 (B) - NM (C.D. Cal.), on file with the author.

*United States v. Mitchell* (1999a) Cr. No. 96–407 Memorandum of Law in Support of Mr. Mitchell’s Motion to Exclude the Government’s Fingerprint Evidence (E.D. Pa.).


