National Academy of Sciences Report

A CHALLENGE TO THE COURTS

BY NANCY GERTNER

In 2009, the National Academy of Sciences issued its report on the forensic sciences. The report was entitled *Strengthening Forensic Science in the United States: A Path Forward* (2009) [hereinafter NAS report]. It was an extraordinary document. It questioned whether the underlying research justified the claims forensic scientists were regularly making in courts throughout this country, claims that they had been making for decades. It concluded that for many long-used types of forensic science, including fingerprint identification, firearms identification, handwriting, and toolmark identification, experts' conclusions were simply not supported by their methodology or their training. There was not an adequate basis for individualization, for linking crime scene evidence to a particular defendant, much less for conclusions that were announced to an exceptional degree of certainty: This bullet matches the gun associated with the defendant "to the exclusion of anyone else in the world," as one ballistics expert testified, to my astonishment. There was "a notable dearth of peer-reviewed, published studies establishing the scientific bases and validity of many forensic methods." (NAS report, at 8.) Moreover, research on proficiency, performance, and the role of bias and observer effects is "sorely needed." (*Id.*) The report did not suggest that this field could never meet scientific standards, only that the current state of the field was—"seriously wanting." (*Id.* at 13.)

But the forensic scientists were not the only target of the NAS report's authors. The courts, the report suggested, have been "utterly ineffective" in assessing the

Published in *Criminal Justice*, Volume 27, Number 1, Spring 2012.

research basis of these sciences. (*Id.* at 53.) They have routinely assumed that since this testimony has always been admitted, it should continue to be, rather than actually evaluating its validity under *Daubert v. Merrell Dow Pharmaceuticals, Inc.*, 509 U.S. 579 (1993), and *Kumho Tire Co. v. Carmichael*, 526 U.S. 137 (1999). Moreover, the NAS criticisms echoed the even more direct critiques of other scholars: Courts were regularly shirking their gate-keeping responsibilities by failing to hold an admissibility hearing or failing to give reasons for their admissibility decision; misapplying or misinterpreting *Kumho Tire*; reversing the burden of persuasion onto the challenger; conflating general acceptance *in the scientific community*—which for the most part, was problematic—with acceptance *by the courts*—in which precedent was simply recited without examination; overemphasizing the flexibility of the inquiry; relegating fundamental questions of validity to issues of weight rather than admissibility; etc. (*See* Michael J. Saks, *The Legal and Scientific Evaluation of Forensic Science (Especially Fingerprint Expert Testimony)*, 33 Seton Hall L. Rev. 1167 (2003).) It was, as one scholar described it, the "ostrich" approach—pretend that there is no problem, pretend that the academic and NAS critique didn't happen, ignore the profound implications for defendant's liberty of the uncritical admission of this testimony, and maybe the issue will disappear. (Jennifer L. Mnookin, *The Courts, the NAS, and the Future of Forensic Science*, 75 Brook. L. Rev. 1209, 1243 (2010).)

The authors of the NAS report were unimpeachable. Leaders of the forensic community had in fact requested this investigation. Congress responded by commissioning the NAS. The report was written by an interdisciplinary panel of distinguished scholars, scientists, and practitioners, who conducted their own investigation and heard days of testimony from leading forensic science professionals, researchers, and others knowledgeable in the field. Given its conclusions, one would have expected a sea change in the treatment of pattern evidence—at the very least robust hearings to evaluate the NAS criticisms, decisions excluding the testimony if scientific standards were not met, or at the minimum, decisions limiting its application.

But in fact, little has changed. Ironically, many courts are paying far, far more attention to the so-called CSI effect, as to which there is little or no basis, than the conclusions of the NAS report. The CSI effect, deriving from the *CSI* (crime scene investigation) television program and its spinoffs, suggests that jurors have unrealistic and preconceived notions about the availability and precision of forensic evidence in criminal trials. Some courts have chosen to voir dire potential jurors to cull those who believe that the government can never meet its burden of proof without technical evidence, and then so instructing the jury at the conclusion of the case. (Stabb v. Maryland, 31 A.3d 922 (Md. 2011); Commonwealth v. Perez, 954 N.E.2d 1 (Mass. 2011).) One went so far as to consider the CSI effect in admitting inconclusive test results on a prosecutor's motion; the jury should know, the court found, that the prosecutor had left no stone unturned. (Delaware v. Cooke, 914 A.2d 1078 (Del. 2007).) The message is troubling: The courts are more prone to monitor the CSI effect at the prosecutor's behest, but not the NAS challenge at the defense's.

The issue could not be more critical. The Innocence Project reports that in more than 50 percent of DNA exonerations, unvalidated or improper forensic science contributed to the wrongful conviction. (*See Wrongful Convictions Involving Unvalidated or Improper Forensic Science that Were Later Overturned through DNA Testing*, Innocence Project, http://tinyurl.com/7tpxtkk.) The misidentification of Brandon Mayfield, an attorney from Portland, Oregon, suspected of the 2004 Madrid train bombing, was especially troubling. Mayfield claimed that he was innocent, that he had never even been in Spain. Three senior FBI fingerprint examiners concluded his fingerprints, which were in a computerized database, were a 100 percent match with those found at the site of the explosion. Ultimately, the FBI agreed that its agents had erred. Mayfield was released, and received a substantial amount of money from the government. More recently, Itiel Dror, in an extraordinary experiment, used the Mayfield case to show how observer bias apparently affects independent fingerprint analysis. After Mayfield was exonerated, a group of international fingerprint examiners were each given a pair of prints that they were told were from the Mayfield case. In fact, they were not; the two prints were from fingerprint sets that each examiner had previously testified conclusively were a match. Sixty percent of the examiners, believing that they were reconsidering the flawed Mayfield prints, claimed the prints did not match; a fourth found them to be inconclusive. (*See* Itiel E. Dror & David Charlton, *Why Experts Make Errors*, 56 J. Forensic Identification 600 (2006).) Their analysis was wholly skewed by the information that they had received about the context of their examination, not by the prints in front of them.

Unlike DNA evidence, which evolved in an academic laboratory setting, and was robustly challenged in court before it was ultimately accepted, trace evidence evolved in the four corners of the courtroom, as evidence important to the prosecution of crime. It was rarely chal-

NANCY GERTNER *served as the judge for the US District Court of Massachusetts from 1994 to 2011 and is now a professor of practice at Harvard Law School.*

Published in *Criminal Justice*, Volume 27, Number 1, Spring 2012.

lenged, and prior to *Daubert* and *Kumho*, rarely excluded or limited. Given this history, the way to change the calculus is clear: Until lawyers fairly bring these standards to the attention of the courts, and until the judges, both district and appellate, rigorously enforce them, nothing will change.

As I have described it:

> It is a vicious cycle: The so-called "pedigree" of trace evidence, namely, the fact that it has been admitted without limitation and without challenge over the decades, creates a disincentive for advocates to challenge it. . . . If the lawyers do not tee up the issue, the evidence will be introduced without objection. If the lawyers do decide to raise these challenges, a busy trial judge can rely on the decades of case law to legitimize decisions rejecting a hearing or motions in limine. And the trial judge can count on the Court of Appeals likely concluding that rejecting the challenge was not an abuse of the judge's discretion. Then those decisions, even if they do no more than endorse the judge's discretion under an abuse of discretion standard, reinforce the view that challenges are futile. While the "no-abuse-of-discretion" decision means only that there is a range of discretionary decisions, including, arguably, admitting or excluding, the decision is typically taken to mean more—another ratification of the status quo, another endorsement of the uncritical admission of trace and pattern evidence. too often counsel cite the decision to mean much more—another endorsement of the uncritical admission of trace and pattern evidence. Judges know they can err on the side of no hearing or no exclusion and be upheld; judges always have been.

(Judge Nancy Gertner, *Commentary on the Need for a Research Culture in the Forensic Sciences*, 58 UCLA L. Rev. 789, 790 (2011).)

Consider for a moment the content of the abuse of discretion standard: It virtually ensures that there will be no meaningful review of admissibility decisions—NAS report or no NAS report. The Supreme Judicial Court of Massachusetts, one of the finest courts in the country, defines abuse of discretion in the ordinary criminal case as follows: To show an abuse of discretion, the defendant has the burden of showing that "no conscientious judge, acting intelligently, could honestly have taken the view expressed by [her]." (Commonwealth v. Cruz, 926 N.E.2d 142, 153 (Mass. 2010) (internal quotation marks omitted).) What then does it take to prove trial error—that the judge was not smart, dishonest, and perfunctory?

To be sure, it is not entirely the court's fault. Advocacy in this area, even after the NAS report, is poor. Take, for example, *United States v. Pena*, No. 1:05-cr-10332-NG (D. Mass. 2008), *aff'd*, 586 F.3d 105 (1st Cir. 2009), *cert. denied*, 130 S. Ct. 1919 (2010), a case before me. The defendant moved for a hearing on fingerprint testimony. There was no expert affidavit, only a single citation to a student note. Nevertheless, I scheduled a hearing and ordered the government to produce its witnesses. At the appointed time, the government represented that its witnesses were in the hall, ready to be examined. To my amazement, the defense announced it had no witnesses, no experts, and did not even wish to take advantage of the opportunity to examine the government's witnesses. The motion, he insisted, had only been brought "for the record." In fact, he added, in what can only be described as an historic understatement: "Well, with all due respect, Judge, I appreciate the Court gave me more credit than I deserve."

I was astonished. I indicated how very seriously I take these challenges, long before the NAS report. But to justify exclusion, the defendant had to do some work—produce some data or expert testimony, real evidence suggesting the limitations of fingerprinting. I even suggested what the problems might be, the problems later highlighted by the NAS report: the lack of uniform standards; the problem that the proficiency rates, error rates, are basically determined without a control; and the fact that the premise of uniqueness (of fingerprints) had never been empirically tested. And I noted my cases in which such a presentation had been made with success, namely, *United States v. Hines*, 55 F. Supp. 2d 62 (D. Mass. 1999) (handwriting), and *United States v. Green*, 405 F. Supp. 2d 104 (D. Mass. 2005) (ballistics).

The case was appealed, and I was upheld—in a decision that did not remotely reflect what had happened below: First, my decision was characterized as one that refused to grant the defendant a *Daubert* hearing, like the hundreds of cases before it in the district courts of this country. The fact is that I *had* offered a *Daubert* hearing. Defense counsel waived it. Second, the First Circuit went so far as to provide a justification for my decision not to have a hearing—that the defendant had failed to bring a novel challenge: "A district court does not abuse its discretion by dispensing with a *Daubert* hearing if no novel challenge is raised. Here, Pena raised no new favorable case law or expert testimony to challenge the admissibility of the fingerprint identification evidence. . . ." (*Pena*, 586 F.3d at 111 n.4 (citation omitted).) Not so. The issue was not one of novelty. I would have permitted counsel to offer testimony of experts on fingerprinting, even if they were doing no more than presenting long-standing critiques of the limitations of fingerprint testimony—such as the critiques in the NAS report. Third, the decision suggests that no hearing was justified because of the particular reliability of the ACE-V fingerprint method, something I never

addressed and would not have accepted without scrutiny: "The district court did not abuse its discretion. Numerous courts have found expert testimony on fingerprint identification based on the ACE-V method to be sufficiently reliable under *Daubert*." (*Id.* at 110.)

The fact is, I would have welcomed a hearing on the ACE-V method and its bona fides. ACE-V stands for "analysis, comparison, evaluation, and verification," nothing more. (*Michele Triplett's Fingerprint Terms*, Northwest Lean Networks, http://www.nwlean.net/fprints/a.htm (last updated Jan. 25, 2012).) Jennifer Mnookin calls the ACE-V relationship to the scientific method "tenuous at best." It is just a way of describing "a set of procedures to describe the careful comparison of a latent print with a potential source print by an initial examiner and a subsequent verifier." She likens it to glorifying "the methodology for fixing a car by the acronym DACT—Diagnose, Acquire, Conduct, and Test." (Mnookin, *supra*, at 1219.)

We, the courts, can do better. In fact, we already do, albeit in civil, not criminal, cases. (*See generally* D. Michael Risinger, *Navigating Expert Reliability: Are Criminal Standards of Certainty Being Left on the Dock?*, 64 Alb. L. Rev. 99 (2000) (noting that courts have been less rigorous in their scrutiny of forensic evidence in criminal cases rather than civil).) In civil cases, the parties put considerable effort into pretrial discovery; on the eve of trial, motions in limine and requests for hearings abound. Despite the best efforts of judges, it is still common for settlements to occur only on the eve of trial. There is every incentive to raise challenges to expert testimony on summary judgment, in a pretrial hearing, or on the eve of trial. In researching an arson case in *United States v. Hebshie*, 754 F. Supp. 2d 89 (S. D. Mass. 2010) (granting habeas relief based on flawed arson testimony), I found no criminal cases critically evaluating arson canine evidence, as the academic literature did; I found only civil cases involving insurance companies and damages.

The pacing of criminal prosecutions, the pressures, the unequal and limited resources make it particularly difficult to raise forensic challenges. The pressure is to plead guilty as quickly as possible, often with little or no information. Defendants race to cooperate with the government; the first one in the door gets the best bargain. *Daubert* challenges, if they are raised at all, are raised at the last minute, an afterthought. And the implications of exclusion are not lost on the judge. It will undermine the state's prosecution and even cast doubt on other older convictions.

If the lawyers have wrongly ignored *Daubert* challenges, there is an arguable challenge under *Strickland v. Washington*, 466 U.S. 668 (1984). Counsel should be presumed to be familiar with the kind of issues the NAS report raised. Courts need to make it clear that such familiarity may be one of the benchmarks in evaluating when assistance of counsel is constitutionally ineffective. The best cross-examiner, with the best skills in the usual driving-under-the-influence case, may not be up to par when complex forensic evidence is involved.

We, the courts, can do better. In fact, we already do, albeit in civil, not criminal, cases.

But I am not sanguine that the courts will enforce this standard. *Strickland* is a notoriously difficult test to meet. While there are some suggestions that advocacy standards should be higher when technical forensic evidence is involved, *Hebshie*, 754 F. Supp. 2d at 113, recent Supreme Court case law points in a different direction. (Cullen v. Pinholster, 131 S. Ct. 1388 (2011); Harrington v. Richter, 131 S. Ct. 770 (2011).) If there has been no meaningful review of counsel's performance on the trial level of appellate levels, there is less and less likely to be effective review on habeas.

The *Pena* case, and the First Circuit's decision affirming it, led me to try a different approach. I issued a procedural order addressing trace and pattern evidence in all criminal cases before me. The order provided that in the wake of the NAS report, admissibility of trace evidence "ought not to be presumed; that it has to be carefully examined in each case, and tested in the light of the NAS concerns, the concerns of *Daubert*/*Kumho* case law, and Rule 702 of the Federal Rules of Evidence." The order also described the procedures governing such a challenge. (Procedural Order: Trace Evidence at 3, No. 1:08-cr-10104-NG (D. Mass. Mar. 8, 2010).) Its impact remains to be seen.

It makes no sense to decry wrongful convictions based on flawed forensic evidence, on the one hand, and to continue to allow it to be admitted without examination, on the other. It makes no sense to be concerned that jurors are holding prosecutors to a higher standard than they should, the so called *CSI* effect, which may not exist, and being less than attentive to when the evidence they offer fails to meet the ordinary standards, *Daubert* and *Kumho Tire*. It makes no sense to ignore *Daubert* and *Kumho* when liberty is at stake, but apply these cases rigorously when all that is involved is money. We must do better. ■