Forensic science reform in the 21st century: a major conference, a blockbuster report and reasons to be pessimistic

JONATHAN J. KOEHLER†
Northwestern University School of Law, Chicago, IL 60611, USA

A 2009 conference at Arizona State University brought together leading scholars to discuss the future of forensic science in light of the blockbuster National Academy of Sciences report entitled ‘Strengthening Forensic Science in the United States: A Path Forward’. This paper introduces the special issue on forensic science that this conference spawned, considers the significance of the report and then offer reasons to be pessimistic about whether major reforms are forthcoming.

Keywords: evidence; forensic science; individualization; National Academy of Sciences; National Research Council.

1. Introduction

On 3–4 April 2009, the Center for the Study of Law, Science, and Technology at the Arizona State University (ASU) Sandra Day O’Connor College of Law sponsored an international conference on the future of forensic science entitled Forensic Science for the 21st Century: The National Academy of Sciences Report and Beyond. The conference, organized by ASU law professors David H. Kaye, Michael J. Saks and me, brought together some of the world’s leading scholars of law, forensic science, statistics, psychology, criminology and related areas. The purpose of the conference was to discuss the future of forensic science and its relationship to the legal system in light of the 2009 National Academy of Sciences report entitled ‘Strengthening Forensic Science in the United States: A Path Forward’ (National Research Council [NRC], 2009).

Forty-one speakers took to the stage during the 2-day conference, including the two co-chairs of the NRC panel and two other committee members. Nearly every speaker called for reform. Topics included the use of unproven methods with unknown error rates, local and institutional changes for governing the operation of the forensic disciplines, legal rules for assessing the reliability of forensic science evidence and lessons learned from cases of postconviction review and reversal. There were also nine question and answer periods.1

Recognizing that the conference was likely to spawn a number of excellent academic papers, Kaye, Saks and I arranged for two peer-reviewed legal journals—Jurimetrics Journal (JJ) and Law, Probability and Risk (LPR)—to consider publishing selected conference papers in special issues. Saks is the special issue editor at JJ. I am the special issue editor for this issue of LPR.

† Email: jay.koehler@northwestern.edu

1 Slides for the talks are available on the conference Web site at http://lst.law.asu.edu/FS09/index.html. A complete set of videotapes of the talks (which include parts of the question and answer periods) may be purchased by contacting the ASU Sandra Day O’Connor College of Law.
2. LPR special issue papers

Each of the four papers in this special issue draws on the NRC report and then reaches well beyond it. Law professor Erin Murphy focuses on the hidden role that DNA typing played in paving the way for a National Academy of Sciences panel to investigate non-DNA forensic science techniques. She underscores that many of the forensic science criticisms and suggestions aired by the NRC report apply to the more highly regarded DNA typing techniques as well.

Historian of science Simon Cole argues that the importance of the NRC report in the ongoing dispute between practicing forensic scientists and judges on the one side and scientists and various academics on the other lies in its embrace of the scientific perspective. Focusing on the controversy over the reliability of fingerprint evidence and testimony, he endorses the scientific perspective and various reforms designed to promote accuracy.

Psychologist Itiel Dror and law professor Jennifer Mnookin warn that the cognitive technologies that some forensic sciences use (such as the Automated Fingerprint Identification System, AFIS) introduce problems related to the interplay between man and machine that can affect the accuracy of forensic science conclusions. They argue that forensic scientists who rely on AFIS and other large computer databases must change the criteria that they use in their examinations.

Economists Glen Whitman and Roger Koppl describe what they refer to as a ‘rational bias’ in the interpretation of forensic science results. They argue, in part, that the prior beliefs of forensic science examiners affect their judgments about the evidence they are asked to evaluate. They recommend major structural reforms—including placing crime laboratories under the auspices of courts or local health departments—to reduce these unwanted effects.

Below, I briefly consider the significance of the NRC report and then offer reasons for my own skepticism about whether it will lead to major reforms.

3. A blockbuster report

By most accounts, the NRC report was a blockbuster. Stating that ‘In a number of forensic science disciplines, forensic science professionals have yet to establish either the validity of their approach or the accuracy of their conclusions’ (NRC, 2009, p. 53), the report calls for a ‘major overhaul’ of the U.S. forensic science system (p. 285).

From a scientific standpoint, the report is significant for its stark acknowledgment of the ‘paucity of research in forensic science’ in support of its fundamental tenets and claims (p. 186). In his Keynote address at ASU, NRC panel Co-Chair Harry Edwards bluntly offered that ‘the most important part of our committee’s report is its call for real science to support the forensic disciplines’ (Edwards, in press). From an institutional standpoint, the report is significant for its finding that the many problems that plague the forensic science community can only be addressed by creating an external national agency that oversees and directs the entire forensic science enterprise (NRC, 2009 p. 20). From a legal standpoint, the report is significant for concluding that the courts have been ‘utterly ineffective’ in pushing the forensic sciences to become more scientific (NRC, 2009, p. 53). Though the report stops short of suggesting that forensic science evidence (aside from DNA) is insufficiently reliable to be admitted at trial, it certainly opens the door to such opinions from U. S. trial judges who hold Daubert or Frye admissibility hearings in response to defense challenges.

To be sure, the report does not say that all forensic science work to date lacks scientific rigor, that most forensic science researchers are incompetent or that the conclusions reported are commonly
inaccurate. But it does say, with substantial justification, that the traditional forensic sciences cannot be trusted to move forward with the foundational work that remains without substantial assistance and guidance from the outside. The central problem is not, as some forensic scientists would argue, simply a lack of resources. The problems that the report identifies—which include inconsistent requirements for training, certification and accreditation, insufficiently rigorous protocols, lack of oversight, inadequate testing and validation of elementary principles, unaided subjective interpretation of data, poor acknowledgment, understanding and measurement of potential sources of bias and error and so forth—will not disappear on their own when funding for the forensic sciences increases. As the report suggests, substantial progress can only be made if there is a genuine commitment to scientific reform within the forensic science community and structural support outside of it. Unfortunately, there are reasons to be skeptical about the will of the requisite communities.

4. Reasons to be pessimistic

4.1 The Federal Bureau of Investigation

The forensic laboratory at the U.S. Federal Bureau of Investigation (FBI) is one of the largest and best funded forensic laboratories in the world. As such, FBI laboratory practices and policies affect the outcomes of many cases and guide the policies of other laboratories. The Director of the FBI Laboratory, Dr. Chris Hassell, was one of the invited speakers at the ASU conference. Dr. Hassell also participated in one particularly significant exchange during the question and answer session following his talk. This exchange, documented in detail by Simon Cole in this special issue (Cole, this issue), served as a kind of microcosm of the gulf that divides forensic scientists from forensic science critics.

Dr. Hassell was asked whether the FBI will continue its practice of instructing its fingerprint examiners to testify that they can determine the source of a latent fingerprint with absolute certainty and to the exclusion of all other possible sources in the world. This question is especially important in light of the NRC report’s conclusion that no area of forensic science technique (other than DNA) ‘has been rigorously shown to have the capacity to consistently and with a high degree of certainty support conclusions about “individualization”’ (NRC, 2009, p. 87). Dr. Hassell invited a more knowledgeable ‘top’ FBI examiner to answer the question. She said, in part, ‘as far as still individualizing, yes we still do it’ (Cole, this issue). This is a stunning nondevelopment. If there is one consistent theme forensic scholars of all types have embraced in recent years, it is the idea that individualization claims cannot be scientifically justified based on existing data. Yet, the most important lab of them all chooses to ignore this fact in favour of a policy of exaggeration.

The FBI also resists the NRC report’s recommendations pertaining to the identification of error rates. The NRC report states that the estimation of error rates from rigorously designed proficiency tests ‘are key components of the mission of forensic science’ (NRC, 2009, p. 122). However, the FBI’s leading scientists have repeatedly stated that forensic science error rates should not be computed. For example, a recent Journal of Forensic Science article by FBI scientists argued that ‘suggesting that a specific error rate must be presented [at trial] adds little value to the discussion on reliability’ (Budowle et al., 2009, p. 801). Likewise, the FBI’s Firearms-Toolmarks Unit argues that proficiency tests should not be used to estimate error rates (Bunch et al., 2009). Significantly,

2 To be clear, ‘individualization’ refers to the practice of identifying the source of a print or marking to the exclusion of all other possible sources in the world.
the four reasons that Bunch et al. (2009) offer in support of their position (anyone can take the tests, the tests are not blind, participants are not anonymous, returns are not anonymous) have more to do with shortcomings of previous tests rather than why properly constructed proficiency tests cannot provide useful error rate estimates.

In short, the FBI is unmoved by two key scientific recommendations in the NRC report. It has no intention of curtailing its practice of making source identifications in various forensic disciplines even though scientific support for this practice is lacking. And it has no intention of trying to scientifically estimate the rate at which it makes mistakes, even though this rate is crucial to understanding the probative value of the evidence (Koehler et al., 1995; Thompson et al., 2003).

4.2 Forensic science organizations

Some professional forensic science organizations have also rejected large portions of the NRC report. Consider, for example, the responses of two important fingerprint organizations—the International Association for Identification (henceforth ‘IAI’) and the Scientific Working Group on Friction Ridge Analysis, Study, and Technology (henceforth ‘SWGFAST’)—to the NRC report’s conclusions on individualization, error rate data and the need for oversight.

Regarding individualization, the IAI hints that the burden of scientific proof should lie with the critic rather than the proponent: ‘There is no research to suggest that properly trained and professionally guided examiners cannot reliably identify whole or partial fingerprint impressions to the person from whom they originated’ (Garrett, 2009, p. 1). Of course, the question begged by individualization claims is whether fingerprint examiners have amassed sufficient scientific proof that the individualization claims made by examiners are accurate. As in all scientific endeavours, the burden of proof lies with the proponent to make the case for the claim based on a body of relevant and rigorous scientific studies.

SWGFAST likewise rejects the NRC report’s individualization conclusion with a hand wave: ‘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). SWGFAST offers a similarly vacuous argument against the call for studies to establish rates of error in forensic science: ‘Although current practices and procedures will not facilitate the calculation of error rates in actual casework because of varying factors and limited information, history demonstrates that the actual error rate in practice is very low. It may be possible to arrive at a generic error rate that considers methodological and practitioner errors through the use of an appropriately designed study. However, determining the reliability of the practice and not error rates would be a better metric in assessing its value as evidence’ (SWGFAST, 2009, p. 4–5). As many commentators have noted, ‘history’ is a poor methodology for assessing error rates due to lack of ground truth. And the notion offered in the final quoted sentence that ‘error rates’ are not part of a determination of the reliability of a forensic practice is just plain wrong. To its credit, SWGFAST (and most other forensic science agencies) endorses the NRC’s call for accreditation of laboratories and certification of examiners. But it rejects formation of an independent oversight agency on grounds that ‘such an entity would require an unprecedented and problematic management model’ (SWGFAST, 2009, p. 5).

The Association of Firearm and Tool Mark Examiners (henceforth ‘AFTME’) also recoils from the Report’s conclusion that the existing science does not support the strong claims made by firearms and tool mark examiners at trial. ‘There is an extensive body of research, extending back over one
hundred years, which establishes the accuracy, reliability, and validity of conclusions rendered in the field of firearm and tool mark identification’ (AFTME, 2009, p. 2). Unlike some other responses, the AFTME cited scientific studies that could lay a statistical foundation for the firearms and tool marks field. For example, one of the cited papers identifies a method for computing the probability that a target bullet and a random bullet of the same type would contain a corresponding set of impression marks (Howitt et al., 2008). The computation Howitt et al. (2008) describe requires measuring the width of the land, the magnification used to examine the bullets using a comparison microscope and the number and width of the striae on the bullets being compared. Howitt et al. (2008) should be commended for doing what they say has never been done in the field: attempting to identify the chance that similar bullet striation markings will occur by chance. But the fact remains that the NRC report has it exactly right: the field of firearms and tool mark examinations has ‘a heavy reliance on the subjective findings of examiners rather than on the rigorous quantification and analysis of sources of variability’ (p. 5–21). As Howitt et al. (2008) note, firearms examiners do not quantify the significance of reported bullet matches in any rigorous way: they simply conclude that a target bullet must have been fired from the same gun as the one that fired a known bullet when the bullets have consistent markings. Given this fact, it is puzzling—and depressing—that the AFTME professional organization failed to take this opportunity to embrace the substance of the NRC report and announce that it would discontinue its practice of reporting matches as absolute source identifications.

4.3 Congress

Finally, the congressional response to key recommended reforms in the NRC report has been tepid. For example, recent hearings in the U.S. House of Representatives and U.S. Senate related to the NRC’s call for an independent federal agency (tentatively called the National Institute of Forensic Science) to oversee and direct the forensic science community reveal little enthusiasm for the proposal among U.S. lawmakers. Some of these lawmakers simply do not get it. They continue to believe the very myths about the forensic sciences that the NRC report so eloquently and forcefully demolishes. Consider, for example, the following comment made by Senator Jeffrey Sessions: ‘I don’t accept the idea that they seem to suggest that fingerprints is not a proven technology. . . . I don’t think that we should suggest that those proven scientific principles that we’ve been using for decades are somehow uncertain and leaving prosecutors as having to fend off challenges on the most basic issues in a trial’ (Senate Judiciary Committee, 2009, p. 35–36).3 Ironically, it is this very type of hand-waving approach to the validity of the forensic sciences that brought us to the point a few years ago when Congress directed the National Academy of Sciences to investigate the forensic sciences.4

5. Conclusions

The good news is that everyone is talking. Forensic scientists, lawyers, judges, legislators, statisticians, chemists, psychologists, criminologists, historians of science and many others are now in the conversation. Some people are listening to opposing points of view and trying to find common ground for purposes of figuring out what the next steps should be. From a scientific standpoint,

3 For more on reactions to the NRC report by the legislative and forensic science communities, see Murphy (this issue, including footnotes 106–112).

4 ‘Recognizing that significant improvements are needed in forensic science, Congress directed the National Academy of Sciences to undertake the study that led to this report’ (NRC, 2009, p. xix).
this type of inclusive process worked to resolve some (but not all) disagreements that arose in the so-called DNA wars of the 1990s (Lander and Budowle, 1994; Saks and Koehler, 2005; Thompson, 1993). This fact alone should make me an optimist, and Saks and I expressed substantial optimism in the Science paper cited above. But in light of the defensive response to the NRC report by the forensic science community, the poor understanding of the scientific issues among legislators and jurists, and the ordinary human tendency to prefer the status quo, I am pessimistic about the prospects for major structural reforms in the traditional forensic sciences.

REFERENCES


GARRETT, R. J. (2009). Letter from International Association for Identification President to IAI Members.


