SYMPOSIUM
LESSONS FROM THE LAB: IMPLICATIONS OF THE 2009
NATIONAL ACADEMY OF SCIENCES REPORT ON THE FUTURE
OF FORENSIC SCIENCE

Introduction: Path Forward or Road to Nowhere?
Implications of the 2009 National Academy
of Sciences Report on the Forensic Sciences

Daniel S. Medwed

The NAS/NRC Report on Forensic Science:
A Path Fraught with Pitfalls

D. Michael Risinger

Independent Crime Laboratories:
The Problem of Motivational and Cognitive Bias

Paul C. Giannelli

Swirls and Whorls: Litigating Post-Conviction
Claims of Fingerprint Misidentification After
the NAS Report

Jacqueline McMurtrie

Will History Be Servitude?:
The NAS Report on Forensic Science
and the Role of the Judiciary

Jane Campbell Moriarty

C.S.I. Bulls#!t: The National Academy
of Sciences, Melendez-Diaz v. Massachusetts,
and Future Challenges to Forensic Science
and Forensic Experts

Joëlle Anne Moreno

Evidence Theory and the NAS Report
on Forensic Science

Michael S. Pardo

Stable Isotopes and Courts

James R. Ehleringer &
Scott M. Matheson, Jr.

COMMENT

Long Overdue: Utah’s Incomplete Approach to Eyewitness
Identification and Suggestions for Reform

Steven J. Joffee
SUBMISSIONS: The *Utah Law Review* receives submissions in March and April. Submissions should include a cover letter, providing the title of the work and the author’s contact information (phone number and e-mail address). Submissions should be double-spaced, in Times 11-point or larger font, printed on one side of the page only, contain footnotes rather than endnotes, and have numbered pages. Microsoft Word format is preferred, but other formats also will be accepted. Electronic submissions can be sent to the *Utah Law Review* Expresso account or via e-mail at utahlawreview_submit@law.utah.edu, or a hard copy can be sent to: Executive Articles Editor, *Utah Law Review*, S.J. Quinney College of Law, University of Utah, 332 S 1400 E Rm. 101, Salt Lake City, UT 84112-0730.

SUBSCRIPTIONS: The *Utah Law Review* is published quarterly. All subscriptions begin January 1 and expire on December 31. For subscriptions, renewals, or cancelations of the *Utah Law Review*, please send notice and/or payment to: *Utah Law Review*, Attn: Angela Turnbow, S.J. Quinney College of Law, University of Utah, 332 S 1400 E Rm. 101, Salt Lake City, UT 84112-0730.

PRICES: $40.00 per volume per year; United States and possessions $45.00 per volume per year; foreign (postage included)

To subscribe to more than one copy of the *Utah Law Review*, send notice of the number of copies requested along with payment. Please note that renewal payment must be received in advance before any issues are sent out to a subscriber. If you work with a subscription agency, send notice and payment to that agency and it will contact the *Utah Law Review* regarding the status of your subscription.

The *Utah Law Review* can now be accessed online at: http://epubs.utah.edu/index.php/ulr.

ISSUES AND REPRINTS: Single issues and back issues are available from William S. Hein & Co., Inc., 1285 Main Street, Buffalo, New York 14209-1987; telephone (800) 828-7571; e-mail: wsheinco@class.org. Reprints are available from Joe Christensen, Inc., 1540 Adams Street, Lincoln, NE 68521; telephone (800) 228-5030; e-mail: sales@christensen.com.
BROWN, TENELLE (2009), Associate Professor of Law. B.A., 2000, Univ. of Pennsylvania; J.D., 2004, Univ. of Michigan.
CHAHINE, KENNETH GREGORY (2009), Visiting Assistant Professor of Law. Ph.D., 1992, Univ. of Michigan; J.D., 1996, Univ. of Utah.
GUIORA, AMOS N. (2007), Professor of Law. B.A., 1979, Kenyon College; J.D., 1985, Case Western Reserve Univ.
LOCKHART, WILLIAM J. (1964), Professor of Law. B.A., 1955, J.D., 1961, Univ. of Minnesota.
McLAUGHLIN, NANCY ASSAF (2000), Associate Dean for Faculty Research and Development and Robert S. Swenson Professor of Law. B.S., 1987, Univ. of Massachusetts, Amherst; J.D., 1990, Univ. of Virginia.
MORRIS, JOHN K. (1979), General Counsel, University of Utah and Professor of Law. B.A., 1966, Univ. of California at Los Angeles; J.D., 1969, Univ. of California at Berkeley.
PoulTer, Susan R. (1990), Professor of Law Emerita. B.S., 1965, Ph.D., 1969, Univ. of California at Berkeley; J.D., 1983, Univ. of Utah.
ROSKEY, CLIFFORD J. (2008), Associate Professor of Law. B.A., 1996, Amherst College; J.D., 2001, Yale Univ.

S.J. QUINNEY LAW LIBRARY FACULTY

DARLAIS, SUZANNE (1990), Head of Information Technical Services and Adjunct Professor of Law. B.S., 1983, Univ. of Utah; J.D., 1989, Brigham Young Univ.; M.L.S., 1990, Univ. of Washington.
FULLER, RONALD (2006), Assistant Librarian and Adjunct Professor of Law. B.A., 2001, J.D., 2005, Brigham Young Univ.
MCPhail, Ross (2007), Assistant Librarian and Adjunct Professor of Law. B.A., 1990, Quincy College.; J.D., 1994, Hamline Univ.
STEPHENSON, LINDA (1986), Head of Information Access Services and Adjunct Professor of Law. M.A., 1983, Univ. of Missouri at Columbia; J.D., 1987, Univ. of Missouri-Kansas City.
# UTAH LAW REVIEW

**Volume 2010**

**Number 2**

## Board of Editors

2009–2010

<table>
<thead>
<tr>
<th>Position</th>
<th>Name</th>
</tr>
</thead>
<tbody>
<tr>
<td>Managing Editor</td>
<td>Miriah Sorensen Elliott</td>
</tr>
<tr>
<td>Executive Text Editor</td>
<td>Samuel Hood</td>
</tr>
<tr>
<td>Executive Footnote Editor</td>
<td>Steven Joffee</td>
</tr>
<tr>
<td>Executive Articles Editor</td>
<td>Tyler Hawkins</td>
</tr>
<tr>
<td>Symposium Editor</td>
<td>Tlesia Stanley</td>
</tr>
<tr>
<td>Articles Editors</td>
<td>Jacob Fonnesbeck, Sarah Goldberg, Ben Lear, Mary Ann Call May, Greg Newman, Adam Reiser</td>
</tr>
<tr>
<td>Note &amp; Comment Editors</td>
<td>Rachel Jacques, Adam Kaas, Nate Olsen</td>
</tr>
<tr>
<td>Associate Editors</td>
<td>David Call, Nicholas Craft, Ashton Hyde, Razvan Ungureanu, Ruke Wang</td>
</tr>
<tr>
<td>Senior Staff</td>
<td>Brooks Benson, Jacob Lloyd</td>
</tr>
<tr>
<td>Staff Members</td>
<td>Hannah Huegel, Brooke Johnson, Freyja Johnson, Beth Kennedy, Peter Law, Jeffrey Mann, Blakely Neilson, Christine Poleshuk</td>
</tr>
<tr>
<td>Advisory Committee</td>
<td>Eric A. Olson, Patricia M. Leith, Amy J. Wildermuth, Michael D. Zimmerman</td>
</tr>
<tr>
<td>Editorial Assistant</td>
<td>Angela Turnbow</td>
</tr>
<tr>
<td>Computing Services</td>
<td>Duc Phuong</td>
</tr>
</tbody>
</table>
# Table of Contents

## Symposium


<table>
<thead>
<tr>
<th>Article Title</th>
<th>Author(s)</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Introduction: Path Forward or Road to Nowhere? Implications of the 2009 National Academy of Sciences Report on the Forensic Sciences</td>
<td>Daniel S. Medwed</td>
<td>221</td>
</tr>
<tr>
<td>The NAS/NRC Report on Forensic Science: A Path Forward Fraught with Pitfalls</td>
<td>D. Michael Risinger</td>
<td>225</td>
</tr>
<tr>
<td>Independent Crime Laboratories: The Problem of Motivational and Cognitive Bias</td>
<td>Paul C. Giannelli</td>
<td>247</td>
</tr>
<tr>
<td>Swirls and Whorls: Litigating Post-Conviction Claims of Fingerprint Misidentification After the NAS Report</td>
<td>Jacqueline McMurtrie</td>
<td>267</td>
</tr>
<tr>
<td>Will History Be Servitude?: The NAS Report on Forensic Science and the Role of the Judiciary</td>
<td>Jane Campbell Moriarty</td>
<td>299</td>
</tr>
<tr>
<td>Evidence Theory and the NAS Report on Forensic Science</td>
<td>Michael S. Pardo</td>
<td>367</td>
</tr>
<tr>
<td>Stable Isotopes and Courts</td>
<td>James R. Ehleringer &amp; Scott M. Matheson, Jr.</td>
<td>385</td>
</tr>
</tbody>
</table>
COMMENT

Long Overdue: Utah’s Incomplete Approach to Eyewitness Identification and Suggestions for Reform

Steven J. Joffee

443
INTRODUCTION: PATH FORWARD OR ROAD TO NOWHERE?
IMPLICATIONS OF THE 2009 NATIONAL ACADEMY OF SCIENCES
REPORT ON THE FORENSIC SCIENCES

Daniel S. Medwed∗

The evolution of DNA technology has shined a light on an issue previously shrouded in darkness: the extent to which innocent prisoners are wrongfully convicted. To date, post-conviction DNA testing has succeeded in exonerating 252 criminal defendants, their innocence proven beyond a shadow of a doubt. DNA exonerations provide a dataset to study in the hopes of figuring out “what went wrong” in these cases. This process of deconstructing post-conviction innocence cases helps not only in understanding the causes of wrongful convictions, but in designing reforms to prevent future miscarriages of justice.

Notably, unreliable forensic science played a role in 52 percent of the first 250 DNA exonerations. This suggests there are problems with the manner in which forensic science is (a) initially produced and (b) later presented as evidence at criminal trials. In its pioneering 2009 report, Strengthening Forensic Science in the United States: The Path Forward, the National Academy of Sciences (NAS) tackled these problems. The NAS Report advocated the creation of a national supervisory body, the National Institute of Forensic Sciences (NIFS), to establish and enforce best practices, to develop standards for the mandatory accreditation of crime laboratories and certification of forensic scientists, and to promote scholarly, peer-reviewed research in the forensic sciences. Pursuant to the NAS Report’s vision, the NIFS also would craft standard terminology for use in reporting and testifying about forensic investigations as well as model laboratory reports. These are just some of the many recommendations contained in the NAS Report; others emphasized the need for greater funding to enhance forensic scientific research and provide better quality control. Now that the initial hubbub surrounding the issuance of the NAS Report has begun to wane, the question becomes whether the

∗ © 2010 Daniel S. Medwed, Professor of Law, University of Utah, S.J. Quinney College of Law; J.D. Harvard Law School, 1995; B.A. Yale College, 1991. I am particularly grateful to Tesia Stanley, Adam Weinacker, and the other members of the Utah Law Review for their assistance in organizing this symposium.

1 See The Innocence Project, www.innocenceproject.org (last visited June 1, 2010); see also Daniel S. Medwed, Innocentrism, 2008 U. ILL. L. REV. 1549, 1551–52 (discussing the emerging centrality of innocence-themed arguments in the criminal law discourse).


4 Id. at 21–22.

5 Id. at 24–25.
report will have any practical impact. The essays in this symposium issue of the *Utah Law Review* offer insight into this important question.

Several contributors to this symposium focus on the NAS Report’s significance regarding the initial production of forensic scientific results. Professor Paul Giannelli praises the NAS Report’s endorsement of independent crime laboratories, a reform he has long championed. Drawing upon the social scientific literature on cognitive bias, Professor Giannelli explains why the current norm in which most state crime laboratories are located within (or at least affiliated with) law enforcement agencies creates a situation where forensic scientists have subconscious motivations to side with the police and prosecutors. Professor Giannelli points out that, along with other measures, fulfilling the NAS Report’s goal of wresting “administrative control” of crime laboratories from law enforcement might help “protect the integrity” of forensic scientific findings.

Professors James Ehleringer, Scott Matheson, and Jacqueline McMurtrie explore the NAS Report through the lens of specific forensic scientific disciplines. Professors Ehleringer and Matheson team up as a biologist and law professor to set their keen eyes on the use of stable isotope ratio analysis in forensic investigation and identification. Stable isotope analysis involves evaluating chemical compounds to discern whether they share a common origin or bear some other form of relationship. This science has been used in a variety of applications in the investigative context and is likely to be used as evidence in courts. Professors Ehleringer and Matheson analyze the reliability of stable isotope analysis and caution that extensive research and rigid adherence to protocols and standards are critical to the successful adaptation of this scientific technique into the forensic realm.

Professor McMurtrie, in turn, assesses the issue of latent fingerprint identification evidence. The NAS Report strongly disputes claims made by latent fingerprint examiners about their ability to pinpoint the source of a fingerprint with a startling degree of accuracy and instead determines that this technique is riddled with holes. In her essay, Professor McMurtrie applies the NAS Report’s findings

---


8 See Gianelli, supra note 6, at 261.


10 See generally id.

to the litigation context, arguing that defendants convicted based on latent fingerprint evidence can and should cite these findings in seeking post-conviction relief.\textsuperscript{12}

The other authors in this symposium mainly examine the repercussions of the NAS Report on the presentation of forensic scientific evidence at trial. Professor Michael Risinger’s contribution affords a bird’s-eye view of the NAS Report, lauding its thorough cataloguing of the ills of forensic science.\textsuperscript{13} Professor Risinger is less sanguine, however, about its potential to cure these ills. In particular, Professor Risinger criticizes the NAS Report’s reluctance to put forth any recommendations regarding the admissibility of forensic science at trial, decrying its essential conclusion “that the litigation process and the actors in it, lawyers and judges alike, are unqualified to exert any positive influence on forensic science practices through rules of admissibility or otherwise.”\textsuperscript{14}

Professor Jane Moriarty is slightly more optimistic about the possible impact of the NAS Report on the admissibility of forensic scientific evidence going forward. She focuses on the role of the judiciary in regulating the admission of forensic science, questioning why courts have either been unwilling or unable to grasp the shortcomings of many forensic science specialties.\textsuperscript{15} She argues that because judges perceive forensic science admissibility as a relatively simple issue (as opposed to grappling with toxic tort causation), they employ intuitive decision making, which behavioral science has shown to be both error-prone and subject to cognitive bias problems. Examining post-NAS Report cases, Professor Moriarty sees small but encouraging signs that some courts may be starting to grasp the complexity of forensic science admissibility questions, which may lead to more careful judicial gatekeeping.\textsuperscript{16}

Professors Joelle Anne Moreno and Michael Pardo also analyze the NAS Report’s likely influence on the future admissibility of forensic scientific evidence in criminal trials. Professor Moreno looks at this issue principally through the United States Supreme Court’s 2009 decision in \textit{Melendez-Diaz v. Massachusetts}.\textsuperscript{17} That case addressed the question of whether a state statute geared toward admitting certified forensic reports into evidence in criminal trials without any live testimony from the forensic scientist herself denied criminal defendants their constitutional right to confront the witnesses against them. In holding that such a statute indeed violated the Confrontation Clause, Justice Scalia cited the NAS Report’s conclusion that the forensic science system is deeply flawed.

\textsuperscript{12} See generally id.


\textsuperscript{14} Id. at 234.


\textsuperscript{16} Id. at 321.

Professor Moreno envisions that *Melendez-Diaz* may prove to be a watershed moment in forensic science (as well as Confrontation Clause jurisprudence) by alluding to a possible “constitutional solution” to the problem of unreliable forensic evidence.\(^{18}\) Professor Pardo evaluates the interaction between the practical problems detailed in the NAS Report and high-level evidence theory.\(^{19}\) As described by Professor Pardo, theoretical accounts of evidence are primarily “epistemological” in that they seek to either justify or to reform evidentiary rules in light of their tendencies to yield true or false outcomes.\(^{20}\) With that in mind, Professor Pardo posits that many of the problems associated with the use of forensic evidence in criminal trials should not be treated as matters of admissibility, but as those of sufficiency of the evidence. That is, instead of jettisoning forensic evidence at the front end of the process, Professor Pardo suggests this information deserves critical examination at the back end in weighing its overall impact on the verdict.\(^{21}\)

***

It is too early to tell whether the NAS Report will lead to sweeping, fundamental change in the production and presentation of forensic scientific evidence. But it is not too early to hazard a guess about the likely consequences of the report. I am fortunate, to say the least, that some of the nation’s leading evidence scholars and scientists are willing to share their views on this critical topic for this symposium issue.

---

\(^{18}\) *Id.* at 350, 365.

\(^{19}\) *See, e.g., Michael S. Pardo, Evidence Theory and the NAS Report on Forensic Science, 2010 Utah L. Rev. 367.*

\(^{20}\) *Id.* at 369.

\(^{21}\) *See generally id.*
American forensic science as an organized field is less than a hundred years old.\(^1\) On balance, this relatively young field has almost certainly had a positive impact on the accurate determination of factual guilt and factual innocence in the criminal justice system. However, in the flush of youth, it is not uncommon for claims to outrun capabilities. For several decades now, many from the academy and some from forensic science itself have pointed to weaknesses both in various forensic fields and in the structure of forensic science practice itself—weaknesses that raised the specter of a forensic science that sometimes made unwarranted claims and that could in practice sometimes aid in the conviction of the innocent. These criticisms were generally dismissed without much examination by the bulk of the forensic science establishment, and the proponents of those claims were dismissed as well. However, the National Academy of Sciences/National Research Council (NAS/NRC) Committee Report\(^2\) has now made it untenable to treat

\(^1\) See D. Michael Risinger, Michael J. Saks, William C. Thompson & Robert Rosenthal, The Daubert/Kumho Implications of Observer Effects in Forensic Science: Hidden Problems of Expectation and Suggestion, 90 CAL. L. REV. 1, 43 n.194 (2002) (“While the product of whatever science an era might muster has made its way into the courtroom for centuries, until the early twentieth century it was the ad hoc product of individual practitioners. General forensic science laboratories in the United States have generally been set up as an adjunct to law enforcement organizations. The first laboratory worthy of the name was set up in the Berkeley (Cal.) Police Department by August Volmer around 1918.” (internal citations omitted)).

\(^2\) NATIONAL RESEARCH COUNCIL, NATIONAL ACADEMY OF SCIENCES, STRENGTHENING FORENSIC SCIENCE IN THE UNITED STATES: A PATH FORWARD (2009) [hereinafter NAS/NRC COMMITTEE REPORT]. This report was issued by the Committee on Identifying the Needs of the Forensic Science Community, which is identified on the title page as a committee of the National Research Council (NRC) (a joint endeavor of the National Academy of Sciences (NAS), the National Academy of Engineering (NAE) and the Institute of Medicine). See id. at iii. In addition, the title page suggests some formal conjunction with both the NRC Committee on Science, Technology, and Law, Policy and Global Affairs, and the NRC Committee on Applied and Theoretical Statistics. Id. at i. Proper characterization of the report is complicated by the fact that the NRC and the NAE are merely administrative subdivisions of an organization whose name under its charter is simply the National Academy of Sciences. The NRC was created as a subdivision of the NAS in 1916, the NAE as a subdivision in 1964, and the Institute of Medicine in 1970. See National Academy of Sciences, History of the National Academies, http://www.nationalacademies.org/about/history.html (last visited June 1, 2010). Things are further complicated by the fact that when the NAS created the National Academy of
criticisms as simply the cavils of uninformed academics with nothing better to do.\(^3\) For that Report identifies and documents many of the weaknesses that had been pointed out by critics over the years, and it calls for the strengthening of forensic science through a process designed to address those weaknesses. As a well-documented catalogue of the problems of forensic science by a highly credentialed body, this report is hugely important. But as a blueprint for change, it is subject to some serious reservations. The official title of the report is *Strengthening Forensic Science in the United States: A Path Forward.* This Article addresses the general route envisioned by the NAS/NRC Committee (the “Committee”) in its Report and some of the more serious roadblocks and pitfalls that one may expect to encounter on that proposed path forward.

Engineering, it retained the designation “National Academy of Sciences” for the non-engineering members of the academy, and after the most recent subdivisions it rebranded itself with the umbrella appellation “The National Academies,” without amendment to its charter, as set out in its act of incorporation. See Act to Incorporate the National Academy of Sciences, ch. 111, § 1, 12 Stat. 806 (1863).

Be all these things as they may, the body of the report makes clear that the Committee was the result of a 2006 congressional charge to the National Academy of Sciences, that the Committee was formed in response to that charge to the NAS, and that the report is the work primarily of that Committee. NAS/NRC COMMITTEE REPORT, supra, at xix, 1. It has become commonplace to refer to the Committee on Identifying the Needs of the Forensic Science Community simply as the “NAS Committee” and the report as the “NAS Committee Report,” or simply the “NAS Report,” and originally I adopted that convention in this Article. However, in some circles the Committee is referred to as the “NRC Committee,” and this was the view of proper characterization adopted by the editors of a recent *Jurimetrics* issue devoted to the report in which I published an article which is something of a companion piece to this Article. D. Michael Risinger, *The NAS/NRC Report on Forensic Science: A Glass Nine-Tenths Full (This Is About the Other Tenth)*, 50 JURIMETRICS J. 21 (2009). Although I disagreed with that editorial decision, the editors of *Jurimetrics* graciously agreed to the compromise designation “NAS/NRC” in my contribution, and that convention has been followed here for the sake of uniformity between the two pieces.

\(^3\) Untenable, but not impossible. For instance, such attacks have continued to be a staple of the editorial posts at the leading website for forensic science practitioners, Crime Lab Report. See, e.g., Crime Lab Report, *Congress Must Protect Forensic Science and Taxpayers from Bad Politics*, MONTHLY REPORT, Mar. 18, 2009, at 2, available at http://www.crimelabreport.com/library/pdf/3-09.pdf (“For years, forensic science has been impugned by a network of post-conviction legal activists who are resistant to acknowledging the preventive benefits that forensic science accreditation has provided over the last twenty-five years.”). These pieces are co-written by Chief Editor John M. Collins, Director of the DuPage County Sheriff’s Office Crime Laboratory and a member of the Board of Directors of the American Society of Crime Laboratory Directors (ASCLD), and Associate Editor Jay Jarvis, an employee of the Georgia Bureau of Investigation who is a member of the Board of Directors of the ASCLD Laboratory Accreditation Board (ASCLD/LAB). Crime Lab Report, Editorial Board, http://www.crimelabreport.com/editorial_board.htm (last visited June 1, 2010).
The NAS/NRC Report was the product of a long process that began in 2005 when the Senate Appropriations Committee inserted into its version of an appropriations bill directed toward issues of both science and justice an appropriation of $1.5 million to be given to the National Academy of Sciences “to create an independent Forensic Science Committee.” The mandate of that Committee as described in the accompanying Senate committee report was very broad, specifying a number of topics but ending with the authorization to “examine additional issues pertaining to forensic science as determined by the Committee.”

The Committee process did not get started in earnest until the fall of 2006, when the NAS set up the Committee structure and selected its members. The Committee’s makeup is an interesting study in balancing constituencies. There were seventeen members. Five were longtime academics or government employees in fields of science, engineering, or mathematics with no apparent previous contacts with the world of forensic science. Another was a distinguished

---

4 See NAS/NRC COMMITTEE REPORT, supra note 2, at 1. That language was also included in the Senate report. See S. Rep. No. 109-88, at 46 (2005). Ironically, the impetus for the congressional action was a call by the forensic science community, acting through the Consortium for Forensic Science Organizations (CFSO), for a committee to examine the problem of insufficient resources for forensic science, to the end that the status quo receive more funding. See Letter from Robert J. Garrett, President of the International Association for Identification (IAI), to Its Membership (Feb. 19, 2009) available at http://www.theiai.org/current_affairs/nas_memo_20090219.pdf; see also Joseph Polski (COO, IAI), Forensic Science: A Critical Concern for Police Chiefs, 76 POLICE CHIEF, Vol. 76, No. 9 (Sept. 2009), available at http://policechiefmagazine.org/magazine/index.cfm?fuseaction=display_arch&article_id=1887&issue_id=92009 (“The study was commissioned to call attention to the general field of forensic science (other than DNA) . . . . The NAS study would be the basis for legislation to recognize and provide funding for the broader forensic service community.”). In fact, Peter M. Marone, Director of the Virginia Department of Forensic Sciences and a member of the NAS/NRC Committee (and Chair of the CFSO) asserted in his March 10, 2009 testimony to the Subcommittee on Technology and Innovation of the House Committee on Science and Technology that, even though it was not the subject of a specific recommendation in the report, first and foremost the report should be viewed as endorsing the proposition that in general “for the state and local laboratories there has been a lack of resources (money, staff, training, and equipment) necessary to promote and maintain strong forensic laboratory systems.” Hearing on the National Research Council’s Publication ‘Strengthening Forensic Science in the United States: A Path Forward’ Before the H. Comm. on the Judiciary, Subcommittee on Crime, Terrorism, and Homeland Security, 111th Cong. 8 (Mar. 10, 2009) (statement of Peter M. Marone). It should be noted that Mr. Marone went on to make a very strong case for the establishment of a National Institute of Forensic Science (NIFS) as recommended by the NAS/NRC Committee. See id.

5 NAS/NRC COMMITTEE REPORT, supra note 2, at 2.

6 See id.

7 See id. at v.

8 Dr. Constantine Gatsonis (Co-Chair of the Committee), Professor of Biostatistics, Brown University; Dr. M. Bonner Denton, Professor of Chemistry and of Geosciences, University of Arizona; Dr. Karen Kafadar, Rudy Professor of Statistics and Physics,
professor of computer science who had actually performed grant-funded work for the FBI on computerized handwriting identification.\textsuperscript{9} Two were forensic pathologists.\textsuperscript{10} Four had experience as analysts in forensic science laboratories, though their careers had taken different paths thereafter.\textsuperscript{11} Peter M. Marone had risen to become head of the Virginia Department of Forensic Sciences.\textsuperscript{12} Dr. Randall S. Murch obtained a PhD in plant pathology in 1979, then soon after joined the FBI, putting in twenty-three years, most of it with the FBI laboratory, before retiring in 2002 and becoming the director of research program development for Virginia Technical University.\textsuperscript{13} Dr. Robert Shaler obtained a PhD in biology in 1968 and spent the next thirty-three years in a variety of positions, primarily with the New York City Medical Examiner’s Office, where his specialty was serology and the forensic analysis of blood, so that he was in at the beginning of the rise of DNA typing and its forensic applications.\textsuperscript{14} In 2005, he retired and joined Penn State University as the director of its forensic science program.\textsuperscript{15} Dr. Jay A. Siegel is a PhD chemist who worked for three years as a forensic chemist analyzing drugs and trace evidence before joining the forensic science program at Michigan State University, where he retired after twenty-five years as head of that program.\textsuperscript{16} He then joined the faculty of Indiana University-Purdue University Indianapolis, where he currently chairs the Department of Chemistry and Chemical Biology and directs its forensic and investigative sciences program.\textsuperscript{17}

\textsuperscript{9} Dr. Sargur N. Srihari, University Dist. Prof. of Computer Science and Engineering, State University of New York at Buffalo. See id. at 297.

\textsuperscript{10} Dr. Marcella Fierro, Chief Medical Examiner, Commonwealth of Virginia, and Professor of Pathology and Professor of Legal Medicine, Virginia Commonwealth University (Dr. Fierro retired from those positions in 2008). See id. at 290–91; see also Associated Press, The Real ‘Kay Scarpetta’ Retires, MSNBC, Jan. 1, 2009, www.msnbc.msn.com/id/22465005/.

\textsuperscript{11} Dr. Ross Zumwalt, Chief Medical Investigator of the State of New Mexico, is a past-president of the National Association of Medical Examiners. NAS/NRC COMMITTEE REPORT, supra note 2, at 299. See also New Mexico Office of the Medical Investigator, Staff Biographies, http://omi.unm.edu/About/Staff.aspx (last visited Mar. 31, 2010).

\textsuperscript{12} NAS/NRC COMMITTEE REPORT, supra note 2, at 292.

\textsuperscript{13} Id. at 293–95.

\textsuperscript{14} Id. at 296.

\textsuperscript{15} Id.

\textsuperscript{16} Id. at 296–97.

\textsuperscript{17} For Dr. Siegel’s current positions, see Indiana University-Purdue University Indianapolis Department of Chemistry & Chemical Biology, IUPUI Chemistry Faculty and
Five members of the panel had law degrees. Of those five, one was the Hon. Harry T. Edwards, co-chair of the Committee and a distinguished senior federal court of appeals judge (formerly chief judge of the D.C. Circuit); one was also a PhD in psychology who was the project director of the Program on Scientific and Technical Evidence of the Federal Judicial Center; one was a criminal defense attorney; one was a person with ten years of experience as a federal prosecutor, followed by seven years doing primarily white collar defense before becoming Dean of Cleveland-Marshall College of Law in 2005. The sole evidence scholar on the panel was Margaret A. Berger. Professor Berger is a distinguished member of the legal academy who has a long involvement in issues of science and the law. But her primary focus has always been on the civil side of things, and although the one short article she had written looking at the criminal side of forensic expertise did register some skepticism, she can hardly be regarded as a person who spent a significant part of her scholarly efforts examining the claims of forensic science and holding them up to extended and careful scrutiny.

I have gone into this fairly lengthy description of the makeup of the NAS/NRC Committee for a reason. The Committee was hardly a hotbed of card-carrying forensic science skeptics. As it turned out, this was a good thing, but this circumstance must be kept firmly in mind in judging the NAS/NRC Report, and especially in judging the response to the report from some quarters.

---

Staff, http://chem.iupui.edu/Faculty/ (last visited June 1, 2010); Indiana University-Purdue University Indianapolis Forensic and Investigative Sciences Program, People of Forensic and Investigative Science, http://www.forensic.iupui.edu/people/ (last visited June 1, 2010).

18 See generally NAS/NRC COMMITTEE REPORT, supra note 2, app. 1 at 287–302.
19 Id. at 287–88.
20 Dr. Joe S. Cecil. See id. at 289–90.
21 Marvin E. Schechter. See id. at 295–96.
22 Goeffrey S. Mearns. See id. at 292–93.
23 See id. at 287–302.
24 Id. at 289.
26 My favorite example of charging the NAS/NRC Committee or its members with bias is contained in the April 1, 2009 editorial posting by Crime Lab Report, which contains the following: “We are deeply troubled that Judge Edwards so inadequately summarized the impressive advancements in quality control and management practices that accreditation has brought to the forensic science community. In our opinion, this omission is evidence of a bias towards the creation of a new government agency to regulate the forensic sciences. Such a bias would tend to discourage any mention of accreditation’s track-record of success.” Crime Lab Report, A Response to the Testimony of Judge Harry T. Edwards Before the Senate Judiciary Committee 1 (April 1, 2009), http://www.crimelabreport.com/library/pdf/4-09message.pdf. Because the subject of this editorial was Judge Edwards’s testimony before the Senate Judiciary Committee on March 18, 2009, concerning the NAS/NRC Report, and because the creation of NIFS was an
Congress seems to have envisioned that the Committee’s work would only take a few months, since the Senate report called for the National Academy of Sciences to issue its report to the House and Senate Committees on Appropriations by June 1, 2006. However, as previously noted, the NAS/NRC Committee was not appointed until the fall of 2006, and did not hold its first hearings on the state of forensic science and what was needed for the improvement of forensic science until January 25, 2007. It took just over two years from that first public meeting until the issuance of NAS/NRC Committee Report in mid-February 2009. In the first year, the Committee held eight days of hearings in which it heard from sixty-four presenting witnesses. The representation of the critic community was higher among the group of presenters than on the Committee itself, comprising approximately eight of the sixty-four presenters. However, most of the other fifty-six presenters were members of the forensic science establishment or representatives of law enforcement. So the presentations given before the Committee were hardly dominated by critics. This too must be taken into account in evaluating the substance of the NAS/NRC Report.

A substantial part of the Report concentrates on weaknesses in the area delimited by the umbrella term “forensic science,” many of which surprised the Committee when they discovered those weaknesses after an extensive examination of the evidence produced by the presenters and other sources. Much of the NAS/NRC Report concentrates on what might be described as the problem children of forensic science. These are, in general, forensic techniques that were explicit recommendation of the NAS/NRC Committee, of which he was co-chair, it seems disingenuous to label his support of the Committee’s recommendation as a “bias.”

28 NAS/NRC COMMITTEE REPORT, supra note 2, at 2.
29 Id. at xi–xiii.
30 To varying degrees, perhaps, the general forensic science community would probably regard the following presenters as outsider “critics”: Peter Neufeld of the Innocence Project, Dr. Itiel Dror (psychology), Dr. Roger Koppl (economics), and Law Professors Brandon Garrett, Paul Giannelli, David Kaye, Michael Saks, and myself.
31 There have been a variety of attacks on asserted biases of the Committee and its process. For instance, the letter submitted by the National District Attorneys Association (NDAA) to the Subcommittee on Technology and Innovations of the House Committee on Science and Technology asserts that “there is a need for further impartial analysis rather than the report’s almost exclusive dependence on the Innocence Project and progenys’ [sic] analysis of the state of forensic science before appropriate, impartial and scientific corrective measures can be implemented.” See Strengthening Forensic Science in the United States: The Role of the National Institute of Standards and Technology: Hearing Before the Subcomm. on Technology & Innovation of the H. Comm. on Science & Technology, 111th Cong. 128 (2009) (March 9 letter of Joseph Cassilly, President, NDAA). Additional material for the record was tendered by Rep. Smith, (R.-Neb.). Id. at 6–7. I don’t know whether the NDAA considers me such progeny or not, but certainly the bulk of the information submitted to and considered by the NAS/NRC Committee did not come from anyone even arguably so characterized.
developed more or less at the dawn of forensic science, such as fingerprint identification, handwriting identification, firearms and toolmark identification, etc. They mostly deal with “source attributions,” that is, determining the source item that left a trace in some relevant place, such as a crime scene. The principles relied upon by such techniques are not the products of science, as that term is currently understood, but rather the product of a kind of commonsense generalization derived from experience with the subject matter under examination. Neither the generalizations so derived nor the accuracy of the results arrived at by practitioners of these disciplines have ever been subject to the kind of systematic testing that has come to be expected as a part of anything calling itself “science.” This does not mean that the results arrived at are necessarily always in error, but simply that we have no very good evidence about when they are likely to be in error and when they are likely to be accurate. In this way, these disciplines are probably best understood as being like folk medicine—they may be efficacious sometimes, maybe even most times, but we don’t really know for sure.

32 One big exception is bite mark identification evidence, which operates in the same general way as the others but which was not actually developed as an asserted form of source attribution expertise until the 1970s. For a short history of bite mark attribution as an expert claim, see D. Michael Risinger, *Navigating Expert Reliability: Are Criminal Standards of Certainty Being Left on the Dock?*, 64 ALB. L. REV. 99, 135–39 (2000).

33 Again, there is one large exception reflected in the report, since one area that was notoriously weak until recently, arson investigation, does not deal primarily with source attribution.

34 This is hardly a new point. See D. Michael Risinger et al., *Exorcism of Ignorance as a Proxy for Rational Knowledge: The Lessons of Handwriting Identification "Expertise,"* 137 U. PA. L. REV. 731, 734 (1989). However, some defenders of the problem forensic science techniques have seized upon this epistemic situation as if it provided a reason to affirmatively accept their claims. Consider the following examples, which could be multiplied considerably: “But it is important to remember the absence of rigorous scientific underpinning in many forensic disciplines does not mean these methods are inaccurate or unreliable; it simply means they are in need of evaluation. Accordingly, I think it is important to recognize the enormous value forensic evidence provides to the justice system even in the absence of full scientific validation, and accordingly exercise caution to ensure we are not overly dismissive of forensic evidence.” *Strengthening Forensic Science in the United States: The Role of the National Institute of Standards and Technology: Hearing Before the Subcomm. on Technology & Innovation of the H. Comm. on Science & Technology*, 111th Cong. 7 (2009) (statement of Rep. Adrian Smith, Ranking Member, Subcomm. on Technology & Innovation of the H. Comm. on Science & Technology); see also *Strengthening Forensic Science in the United States: A Path Forward: Hearing Before the Subcomm. on Crime, Terrorism & Homeland Security of the H. Comm. on the Judiciary*, 111th Cong. 2–3 (2009) (statement of Rep. Louie Gohmert, Ranking Member, Subcomm. on Crime, Terrorism & Homeland Security of the H. Comm. on the Judiciary). Similarly, consider this from John M. Collins: “The National Academies aren’t saying that the engines driving forensic science are pushing in the wrong direction. They are simply recommending more horsepower. Much of the research they call for will further add to the validity of the most commonly practiced disciplines.” *Crime Lab Report, Legacy of Historic Document Depends on Good-Faith Collaboration*, at 1, 3 (March 4,
This is not to say that the Committee did not recognize the scientific credentials and reliability of many techniques, methods, and processes commonly used in forensic science laboratories. Indeed, it held up one such process, DNA profiling for purposes of source attribution, as (perhaps a little too uncritically) a kind of platinum standard for forensic science processes generally. And beyond DNA, many of the instrumented analytical processes used in forensic science were developed in academic science settings and are of well-documented validity and reliability, with well-known sources of possible error, which must be taken into account in their utilization. Gas chromatography works as well in the forensic setting as it does in the research setting, and it has the same potentials for misinterpretation everywhere. But even such well-credentialed processes can

---

Movies, books and TV shows tend to misrepresent the capabilities of gas chromatography and the work done with these instruments.

In the U.S. TV show CSI, for example, GCs are used to rapidly identify unknown samples. “This is gasoline bought at a Chevron station in the past two weeks,” the analyst will say fifteen minutes after receiving the sample.

In fact, a typical GC analysis takes much more time; sometimes a single sample must be run more than an hour according to the chosen program; and even more time is needed to “heat out” the column so it is free from the first sample and can be used for the next. Equally, several runs are needed to confirm the results of a study—a GC analysis of a single sample may simply yield a result per chance . . . .

Also, GC does not positively identify most samples; and not all substances in a sample will necessarily be detected. All a GC truly tells you is at which relative time a component eluted from the column and that the detector was sensitive to it. To make results meaningful, analysts need to know which components at which concentrations are to be expected; and even then a small amount of a substance can hide itself behind a substance having both a higher concentration and the same relative elution time. Last but not least it is often needed to check the results of the sample against a GC analysis of a reference sample containing only the suspected substance.

A GC-MS can remove much of this ambiguity, since the mass spectrometer will identify the component’s molecular weight. But this still takes time and skill to do properly.
result in error in the wrong hands, or in the wrong setting. So even in regard to the well-validated methods used in forensic science, the Committee found great cause for concern resulting both from some common problems of laboratory structure and from the huge variation in the ways in which forensic laboratories are organized and run in the highly fragmented world of law enforcement in our federal system. The Committee recognized that various sources of biasing information available to analysts might seriously undermine the accuracy of the forensic science product and that, in addition, there is a serious problem with misleading testimony by forensic practitioners, which sometimes, either intentionally or inadvertently, exaggerates the probative value of various results.

So the NAS/NRC Report may be taken to identify two broad types of problem in American forensic science practice: problems in regard to our knowledge concerning the accuracy of various techniques even under ideal conditions, and problems related to the less-than-ideal conditions in which forensic science is practiced, which can potentially affect the accuracy of processes that are well researched and whose accuracy and error rates under proper conditions are well established. We may refer to the first as basic validity problems, and the second as environmental problems that undermine ideal performance. The Committee addressed these issues in a set of ten recommendations, which can be summarized as follows:

Similarly, most GC analyses are not push-button operations. You cannot simply drop a sample vial into an auto-sampler’s tray, push a button and have a computer tell you everything you need to know about the sample. According to the substances one expects to find the operating program must be carefully chosen.

A push-button operation can exist for running similar samples repeatedly, such as in a chemical production environment or for comparing 20 samples from the same experiment to calculate the mean content of the same substance. However, for the kind of investigative work portrayed in books, movies and TV shows this is clearly not the case.


The Committee issued thirteen recommendations in all. NAS/NRC COMMITTEE REPORT, supra note 2, at 14–33. The eleventh recommendation dealt exclusively with forensic pathology, calling for the abolition of the coroner system where it still survives (as has oft been recommended before). See Strengthening Forensic Science in the United States: The Role of the National Institute of Standards and Technology: Hearing Before the Subcomm. on Technology & Innovation of the H. Comm. on Science & Technology, 111th Cong. 25 (2009) (statement of Dr. James C. Upshaw Downs, Georgia Bureau of Investigation). The twelfth dealt with making a priority of obtaining interoperability in state and federal automated fingerprint databases. NAS/NRC COMMITTEE REPORT, supra note 2, at 31–32. The thirteenth recommendation dealt with planning to ensure that forensic science is maximally utilized in situations presenting threats to homeland security. Id. at 33. Those three recommendations are not a focus of this Article.
First and foremost, a new independent federal agency (the National Institute of Forensic Science, or NIFS) should be established and charged with authority to establish and enforce best practices for forensic science laboratories and professionals. This would include authority to establish standards for accreditation and certification, and also authority to promote necessary and appropriate research. NIFS should fund research to determine the accuracy and reliability of those currently used techniques that lack data on these issues, and such research should examine those techniques across the conditions that present themselves in practice.

As far as laboratory organization is concerned, all forensic science laboratories should be removed from the administrative control of law enforcement agencies.

NIFS should make sure that all work in forensic laboratories is properly documented using standard procedures and terminology, and that all resulting testimony is clear and uses standard forms of expression calculated to communicate the true meaning of the results of various forensic assays. There should also be in place in every lab a set of quality control procedures designed to identify mistakes, fraud, and bias and to ensure that best practices are followed.

Accreditation of laboratories and individual certification of practitioners should be mandatory. There should also be a standard code of ethics for all forensic science practitioners with enforcement mechanisms.

NIFS should fund research on the effects of observer bias to determine whether it currently affects the results of forensic examinations and, if so, how much.

NIFS should provide money to underwrite both academic training of forensic science personnel, and the development of a normal research infrastructure in the academy.38

It is necessary to note one glaring absence from the Committee’s explicit recommendations: there is no suggestion to improve admissibility gatekeeping decisions at trial, nor to rely upon them for the improvement of forensic science practice or product. The Committee essentially concludes that the litigation process and the actors in it, lawyers and judges alike, are unqualified to exert any positive influence on forensic science practices through rules of admissibility or otherwise. This is why the Committee’s proposed solutions are exclusively structural, institutional, and upstream for the commencement of litigation. I have elsewhere come to similar conclusions, and for similar reasons.39 However, we will revisit this topic later in this Article, and it is enough here to note the general position of the NAS/NRC Committee Report in this regard.

---

38 In these points I have tried to capture the nut of each recommendation. In order to make organizational sense, at least to me, I have had to treat some of them out of order.

In the report we can see a few main themes that are elaborated upon at length and which drive the recommendations. One main theme is that techniques currently without proper validation of some sort ought to be tested to determine the contours of their accuracy. Although only one recommendation addresses this problem directly, a great part of the text deals with it in one way or another, including a whole primer on the basics of scientific validity research, which I suppose is intended at least in part to act as a guide for developing and judging validation research proposals.40 A second main theme is the presence of a significant amount of unpolicied bad practice in parts of the law-enforcement-dominated institutions that constitute our fractured and unstandardized forensic science system (if it can be called a system). While this theme is not explicit in the recommendations, it lies behind the bulk of the details in the recommendations, which deal with issues of institutional structure and independence from law enforcement control.

Let us address the first theme first: the recognition of the existence of problem-child areas in forensic science. As previously noted, the report makes clear that the Committee believes these areas create potential problems of inaccuracy of unknown dimensions, and that these problems have been basically unaddressed and unaffected by lawyers and judges making decisions in the criminal trial setting, on admissibility or otherwise. The only solution the Committee could envision is a substantial, well-funded research effort aimed at testing the claims of these areas across the whole range of issues that forensic practitioners currently claim to be able to resolve dependably, coupled with the hope that the results of such research will be incorporated into best-practice standards through laboratory accreditation and individual certification requirements operating upstream from the trial. In short, the problem children will be disciplined through research administered under the same reformed regime that will solve the problems of bad standards and lax practices in other more validated areas. It is to these envisioned institutional reforms that we will now turn.

The NAS/NRC Committee clearly would like to see a forensic science world manifesting significant uniformity of high-quality practice, with results of well-validated processes uninfluenced by partisan considerations that are communicated to fact finders as neutrally as possible, without overclaiming or obfuscation. To make progress toward this clearly desirable but perhaps utopian vision, the Committee puts its faith in central command and control by a body committed to the best standards of modern science, or at any rate it recommends as close to such a model as it thought it could recommend with any chance of success, given the political realities of our federal system. It further places its faith in the power of accreditation of laboratories and certification of individual practitioners.

Now I personally might favor such an arrangement. The emphasis on the standards of normal science and on building a university-based research community are perhaps enough to calm my fears that a single dominant authoritative agency will make the epistemically wrong decisions about which

40 NAS/NRC COMMITTEE REPORT, supra note 2, at 111–126.
practices to enshrine and which to eliminate; however, the Committee’s own punt on masking protocols to eliminate observer bias does give me some pause.

Let me expand just a bit on this latter point. I have already treated it in detail in another paper. Suffice it to say here that the NAS/NRC Committee Report took the position that steps to address the distortions of observer bias resulting from exposure of forensic analysts to domain-irrelevant information should await more research. This is totally wrongheaded, in my view, and I consider this the worst failure of the NAS/NRC Committee Report, since the research record about the distorting effects of such suggestive information (e.g., the other evidence against a suspect, or the police theory of the case) is quite overwhelming in establishing dangers of current practice, and those dangers can be eliminated by implementing masking and sequential unmasking procedures of very modest cost. Although further research might establish effect sizes in various contexts, one does not need to calculate exactly how hard it is raining to see the wisdom of using an umbrella.

However, even in the face of this extremely serious defect, I would probably be willing to accept the Committee’s recommendations as giant strides forward. But what I would be willing to accept does not define what the body politic, through its various elected representatives, is willing to accept (to say the least). What are the prospects for the recommendations made by the Committee being adopted and having a substantial effect? Consider the two centerpieces of the Committee’s recommended reforms: the establishment of a National Institute of Forensic Science (NIFS) and the removal of all American forensic science laboratories from law enforcement “administrative control.”

This latter proposal is not new. It was originally suggested in a 1974 article in the Journal of Forensic Sciences by M.A. Thompson, and Professor Giannelli again sounded the call in an article in 1997. The proposal is also undoubtedly wise, although in the end such a separation might be more formal than real. But as I and my co-authors Saks, Thompson, and Rosenthal wrote in 2002:

41 See Risinger, supra note 2.

42 See id. at 22, 33–34 and authorities there cited; Dan E. Krane et al., Response to Wells, 54 J. FORENSIC SCI. 501 (2009); Dan E. Krane et al., Response to Ostrum, 54 J. FORENSIC SCI. 1500 (2009); Dan E. Krane et al., Commentary on Budowle et al, 55 J. FORENSIC SCI. 273 (2009); Risinger et al., supra note 1, at 45–52; Dan E. Krane et al., supra note 35, at 1006.

43 M.A. Thompson, Bias and Quality Control in Forensic Science: A Cause for Concern, 19 J. FORENSIC SCI. 504, 512 (1974).


45 Even in a regime of organizational independence, there is a real danger of institutional capture by law enforcement, springing from the fact that the initial contact in most cases will come from law enforcement, and law enforcement will continue to be the primary customer of the independent laboratory. This has been a problem with medical examiner’s offices in many jurisdictions, despite their nominal independence. This likelihood was cited by Matthew F. Redle, County Prosecuting Attorney, Sheridan County Wyoming, testifying on behalf of the National District Attorneys Association (NDAA)
Historically, criminal defendants as a group benefited from the unavailability of information. It is hardly surprising that the law enforcement arm of the state organized efforts to apply science and quasi-science methods to problems of solving and proving criminal cases. In so doing, law enforcement utilized the tools available: officers trained as “technicians” by the small number of scientists with law enforcement interests. In seeking to change these historical remnants, we do not pretend that we are writing on a clean slate. In regard to some organizational reforms, as the vaudeville punch line says, “you can’t get there from here,” at least not within the foreseeable future.46

I do not believe that conditions have changed sufficiently in six years for the accomplishment of this desirable result to have much of a chance. Laboratories have been intertwined with police organizations for too long to have the police organizations surrender control over them willingly. First, just on a level of pure bureaucratic power, organizations do not easily agree to give up significant centers of budget allocation, if for no other reason than that the bigger your budget is the more important and powerful you are. In addition, there is the natural fear that the new arrangements will change the relationship and the product in ways that law enforcement will not like—it may become less predictably supportive of the desires of law enforcement to have particular outcomes ratified. Whatever the reasons—so far as I have been able to determine—few speaking for law enforcement agencies, prosecutorial agencies, forensic science organizations, or forensic science publications have supported either independence or the establishment of NIFS,47 and most that have spoken have opposed both.48
The Senate Judiciary Committee and subcommittees of the House Judiciary Committees and the House Committee on Science and Technology have all held hearings on the NAS/NRC Committee Report. Both the proponents and opponents of the report and its recommendations have been well represented in the witnesses called to testify. The members from the Democratic majority on these committees have manifested the expected dismay at the weaknesses of forensic science revealed by the NAS/NRC Report (Judiciary Committee Chair Senator Patrick pdf/AAFS_Position_Statement_for_Press_Distribution_090409.pdf. It remains to be seen whether this support will continue in succeeding AAFS administrations.

Another partial exception is the position taken by Dr. James C. Upshaw Downs, who spoke to the Committee in favor of NIFS, but was not in favor of independence. See Strengthening Forensic Science in the United States: The Role of the National Institute of Standards and Technology: Hearing Before the Subcomm. on Technology and Innovation of the H. Comm. on Science and Technology, 111th Cong. 30 (2009) (statement of Dr. James C. Upshaw Downs, Georgia Bureau of Investigation).

Another arguable exception may be the position of Carol E. Henderson. In her testimony to the Subcommittee on Technology and Innovation, Professor Henderson endorsed the establishment of NIFS as “a long term goal,” with NIST doing the work for the foreseeable future. Id. at 15 (statement by Carol E. Henderson, Professor of Law at Stetson University School of Law, Director of the National Clearinghouse for Science, Technology & the Law (which is partly funded by the Justice Department’s National Institute of Justice) and immediate past president of the American Academy of Forensic Sciences). How you view her position is much dependent on how long a term you think she had in mind. She points out that in Australia it took twenty years to establish a National Institute of Forensic Science, and that our political situation is even more fragmented than theirs was. One gets the impression that she would be happy to see it come sooner rather than later, but does not believe this is possible.

None of this is to say that there have not been a few individual (generally independent) forensic science practitioners who have supported the core recommendations of the NAS/NRC Committee. See, e.g., Brent E. Turvey, The NAS Report on Forensic Science: A Forensic Scientist’s Response, http://crimereconstruction.blogspot.com/2009/02/nas-report-no-science-in-forensic.html (last visited June 1, 2010).

48 See, e.g., Written Testimony of Matthew F. Redle, County and Prosecuting Attorney, Sheridan County, Wyoming, to the S. Comm. on the Judiciary (Sept. 9, 2009) (testifying on behalf of the National District Attorneys Association (NDAA) and opposing NIFS and laboratory independence); National Research Council’s Publication “Strengthening Forensic Science in the United States: A Path Forward”: Hearing Before the Subcomm. on Crime, Terrorism, and Homeland Security of the H. Comm. on the Judiciary, 111th Cong. 5–15 (2009) (statement of Kenneth E. Melson, Acting Director of the Bureau of Alcohol, Tobacco, Firearms and Explosives, testifying on behalf of the Department of Justice, and opposing, in relatively diplomatic terms, NIFS and laboratory independence); Meredith Mays, IACP Responds to National Academy of Sciences Report on Forensics, POLICE CHIEF, July 2009 (strongly opposing lab independence and nationally mandated standards, but supporting an alternative agency of law enforcement practitioners with “state, county, local, and tribal representation” to establish best practices and serve as a funding source for forensic science services).
Leahy referred to the report as “rather chilling.” The Republican members have manifested equal dismay at the two main recommendations of the NAS/NRC Committee, and the Democrats have not been so enthusiastic either. For instance, the agenda (aka “Hearing Charter”) of the first hearing, that of the Subcommittee on Technology and Innovation, chaired by Rep David Wu (D-Ore.), had as its main focus the idea of using the National Institute of Science and Technology (NIST) as the oversight body rather than creating a National Institute of Forensic Science.

Given the usual power in Congress of the law enforcement lobby and states’ rights proponents in law enforcement matters, it does not seem to me that the independent agency called for—with its power to set national standards and then call for their implementation, and to influence that implementation through imposing conditions on federal aid to state forensic science laboratories—is likely to be adopted anytime soon. Furthermore, any hope of congressional action to coerce or encourage the establishment of independence of forensic labs from law enforcement control is also dead on arrival. Some narrow authority may be given to the NIST or some other existing federal agency, and some money for research on the problem children will almost certainly be forthcoming. But unless this research money is administered by an agency with strong science credentials, that develops both an understanding of the problems and a commitment to their solution, any “honest broker” scientific research proposals coming from academia will almost certainly face stiff competition for the money with law-enforcement-sponsored, faux-research proposals of a kind that have been unfortunately common in the past, and which may suck up most of the money and generate mostly research pre-tailored to reaching results affirming the propriety of the status quo.


51 States’ rights has been a theme of some Republican responses to both NIFS proposal and any national mandate on either independence of laboratories or the setting of obligatory standards. See, e.g., id. at 7–8 (statement of Paul C. Broun, R-Ga., rejecting a “new department to oversee this venture” and stating that “it is unconstitutional to do so, as nowhere in the documents our Founding Fathers penned do they afford Congress that power. Instead, I believe that we should look to individual states to set uniform standards for use within their borders. . . . Any move to federalize forensic science is a move to stifle scientific freedom and in its place adopt more government control.”).

I suppose I sound pessimistic. But I am perhaps a bit more hopeful than I sound. If some NIST-administered research money makes it into the hands of honest-broker researchers who design real studies to test the limits of expertise and the error rates of common subtasks of various currently undervalidated areas, then

The Mecklenburg Report was a report to the Illinois Legislature about a study comparing the results in actual practice of current procedures for administering line-ups and photo spreads with the results of procedures incorporating reforms suggested by research. Among other deficiencies, the design conflated the effects of sequential presentation of subjects with the effects of administrator-blind presentation of subjects. The results were used to resist any change in regard to either sequential or blind presentation. For a full examination of the many weaknesses of the Mecklenburg Report, see Daniel Schacter et al., Policy Forum: Studying Eyewitness Identifications in the Field, 32 LAW & HUM. BEHAV. 3 (2007). This evaluation was the product of a blue ribbon panel assembled by John Jay College of Criminal Justice. The members of the panel were Daniel Schacter, Professor of Psychology, Harvard University; Robyn Dawes, Queenan Distinguished University Professor, Carnegie Mellon University, Fellow, American Statistical Association; Henry L. Roediger III, James S. McDonnell Distinguished University Professor at Washington University, former President, Association for Psychological Science; Larry L. Jacoby, Professor at Washington University; Daniel Kahneman, Professor of Psychology, Princeton University, 2002 Nobel Laureate in Economics; Richard Lempert, Distinguished Professor, University of Michigan School of Law, and Division Director for the Social and Economic Sciences at the National Science Foundation, 2002-2006; Robert Rosenthal, Distinguished Professor, University of California, Riverside, and Pierce Professor of Psychology emeritus, Harvard University, Co-Chair, Task Force on Statistical Inference, American Psychological Association. For another evaluation of the Mecklenburg Report, see Richard A. Wise, Kirsten A. Dauphinais & Martin A. Safer, A Tripartite Solution to Eyewitness Error, 97 J. CRIM. L. & CRIMINOLOGY 807 (2007).

As for the Hall & Player study, it was apparently intended to “ascertain if the normal working practices employed by the Metropolitan Police Fingerprint Bureau introduce an emotional bias,” Hall & Player, supra note 52, at 37, assertedly using as a model previous research by Itiel Dror. However, the Hall & Player design was staggering deficient to produce data on that or virtually any other question of interest, although the authors claimed to have answered the study question in the negative. Id. For a catalogue of the deficiencies of the study, see Itiel E. Dror, Letter to the Editor, On Proper Research and Understanding of the Interplay Between Bias and Decision Outcomes, 191 FORENSIC SCI. INT’L e17 (2009), and Michael J. Saks, Letter to the Editor, Concerning L. J. Hall, E. Player, “Will the Introduction of an Emotional Context Affect Fingerprint Analysis and Decision-Making?” Id. at e19 (which ends, “[D]id anyone read this manuscript thoughtfully (before it was published)?”). The ineffectiveness of Hall & Player’s rejoinders may be judged by examining Lisa J. Hall & Emma Player, Letter to the Editor, The Value of Practitioner Research in the Field of Fingerprint Analysis, id. at e15 (responding to Dror) and Lisa J. Hall & Emma Player, Response to Letter to the Editor, The Value of Practitioner Research in the Field of Fingerprint Analysis, id. at e21 (responding to Saks).

The reader can perhaps begin to see why I am suspicious of law-enforcement or ingroup sponsored research on these issues.
progress can be made, although it will be slow because no single study can map the contours of reliability for any area of claimed expertise. And perhaps more importantly, it will be slow because any such well-designed testing regime will be threatening, and there is likely to be little voluntary cooperation with such studies unless it is made a condition of employment, which nobody currently anticipates doing even under the regime envisioned by the NAS/NRC Committee. Absent some effective incentive to cooperate, progress in validating these areas will take many years, if not decades.

Still, glacial progress would still be progress. And there is no doubt forensic science culture is slowly changing, as more science-trained individuals replace technicians who were police officers first. But it is easy to exaggerate the actual shift. Perhaps half of the forensic service “providers,” as Peter Marone refers to them, do not work in forensic science laboratories, so laboratory accreditation efforts, whatever their virtues, don’t touch them. There are perhaps 11,000 of these, operating out of the 17,000 local police agencies and departments—police officers who have taken a variety of instruction in long or short courses and obtained certificates to attest to their supposed competence in one or more of the “problem child” areas such as fingerprint identification or handwriting analysis or firearms examination. These people testify in court, but we don’t even know how many of them there are. They will be unaffected by even the most expansive application of an individual certification policy if it is tied to laboratory employment. Unless somebody provides the impetus for cooperation with real research (as opposed to friendly faux research) and for individual certification

53 Resistance by practicing forensic scientists to participation in research studies, especially studies not undertaken by “in-group” researchers, has a decades-long history. See D. Michael Risinger, Appendix: Cases Involving the Reliability of Handwriting Identification Expertise Since the Decision in Daubert, 43 TULSA L. REV. 477, 480–81 (2007) (a separately published appendix to Risinger, supra note 39). After all, when you think of it, as far as cooperation with research is concerned, there is really nothing in it for them under current arrangements. They are already, with only rare exceptions, allowed to testify to whatever they want, and the results of research can only make their situation the same or worse; it can’t make it better, from their perspective.

54 Consider, for example, the research that has been done on the validity of handwriting identification. The first of the formal studies directed at these issues was undertaken by Dr. Moishe Kam and his colleagues in the early 1990s, and thereafter such research was undertaken by the Australians Dr. Brian Found, Dr. Doug Rogers, and Dr. Jodi Sita, as well as Kam. But in all the intervening years, no forensic task beyond signature authentication has been studied. See the discussion and analysis in Risinger, supra note 53, at 486–94.


56 Id. These numbers are clearly very rough estimates. The problem of knowledge concerning these numbers is so great that these local police forensic science “providers” are sometimes referred to as the “dark matter” of forensic science.
beyond the laboratory setting, most of us will not live to see substantial improvement in much of the forensic science that makes it into court.

And that is where the courts will have to come in.

Earlier in this Article, I explained that the NAS/NRC Committee had indicated that the courts were useless as a check on bad forensic science, and I intimated that I had come to that conclusion myself (in writing) before the NAS/NRC Report was issued.\(^57\) In a way, this lets lawyers and judges off the hook. But in the aftermath of the NAS/NRC Report, given the low likelihood of its suggested reforms being adopted and actually generating reform, perhaps lawyers and judges should not get off the hook so easily.

It is undoubtedly true that in the main, control of unreliable forensic expertise in the courtroom has been abysmal. And there is enough blame to go around. Prosecutors, defense lawyers, and both trial and appellate judges all bear part of the blame. Many prosecutors regularly push forensic science witnesses to testify in terms that are overly certain and misleading (though it must be said that many do not have to be pushed). However, given the nature of the adversary system, and the prosecutor’s own lack of qualifications in most cases to tell what is well warranted and what is bogus, prosecutors may have the least to apologize for.

Criminal defense lawyers, on the other hand, are supposed to be the people who recognize bogus expert claims, challenge them, move to get them excluded, and undermine those that survive exclusion by knowledgeable, thorough, and telling cross-examination. On the whole, they don’t do any of these things very well.

There are reasons, of course. Most criminal defense lawyers are greatly overworked and have many fewer resources than the prosecution, including virtually no access to forensic experts of their own in most cases. They are constantly “putting out fires” and as a result potential expert reliability issues are seen late and processed under intense time pressure. In addition, most criminal defense lawyers have no background in science or in anything else that would prepare them to evaluate the strengths and weaknesses of the expert claims being made by a particular witness. When they do mount a reliability challenge under such conditions, it is usually underresearched and overgeneralized. The latter circumstance is especially pernicious, for a number of reasons. First, in a federal setting, *Kumho Tire v. Carmichael*\(^58\) requires that all challenges to expert reliability pursuant to Federal Rule of Evidence 702 be directed specifically to the “task at hand.”\(^59\) Most federal defense attorneys don’t understand this, although the opinion

---

\(^{57}\) See supra note 39 and accompanying text.

\(^{58}\) 526 U.S. 137 (1999).

is clear enough, because they have never read the opinion. Rather, they have generally taken their sole knowledge from the conventional headnote-style wisdom that *Kumho Tire* extends the judge’s gatekeeping responsibility under Rule 702 to “nonscientific” expertise. But that is in many ways the least of what it does. And even in a state setting, whether operating under a state analogue to Rule 702 as interpreted by the U.S. Supreme Court or under the so-called *Frye* test or some other standard, the task-specific approach to attacking expert reliability should be used. A more global attack may be more dramatic and may garner more headlines (Lawyer tries to have fingerprints excluded!), but it is by far the least likely to prevail for the client in front of you. Sure, most defense attorneys would love to go down in history as the person who killed handwriting identification expertise (or what have you). And on top of that, the downstream payoffs to future clients from such a global victory are obvious. But it almost certainly isn’t going to happen, and, more importantly, it really shouldn’t happen. Even the least validated of the problem child forensic identification techniques has some well-warranted applications, and global victories, if won, would eject those sometimes rare babies with the bath water.

Of course, global attacks are easier because you don’t really have to learn anything about the real strengths and weaknesses of an area of claimed expertise. And a proper task-at-hand attack has a couple of practical drawbacks. First, it might not be effective if your case just happens to involve one of the more defensible applications, and second, it requires the lawyer to identify and understand the particular expert claim at issue in the case and where it fits in with other kinds of expert claims in the area. It further requires assembling the sources of empirical data that might be said to bear on task-specific reliability. This is a tall order to fill on the fly and under pressure. And so global claims get made, usually backed by very little, which invites global rejection by judges.

It isn’t just the on-the-fly challenges that are too global. Some of the more well-prepared attacks, such as the attack on fingerprints in the famous *Llera Plaza* case, have been overly global, and I am afraid that some academics are responsible for this. There is one school of thought in the academy that holds, essentially, that without quantified random match probability data, all expert claims concerning identification or source attribution should be regarded as

and bar to read and take *Kumho Tire* seriously (or even read it at all), see Risinger, *supra* note 39.

60 *United States v. Llera Plaza* was a federal prosecution in which the defense launched a global attack asserting the failure of fingerprint identification to meet the requirements of Rule 702 as interpreted by the Supreme Court. 179 F. Supp. 2d 492, 494 (E.D. Pa. 2002) [*Llera Plaza I*]. Judge Pollak originally ruled that fingerprint identification could not meet the required standards. *Id.* at 515. He then reversed himself and found fingerprint identification globally acceptable. United States v. Llera Plaza, 188 F. Supp. 2d 549, 576 (E.D. Pa. 2002) [*Llera Plaza II*]. For an extended discussion of why he was wrong both times, see Denbeaux & Risinger, *supra* note 59, at 66–74.
invalid. But I believe this goes too far. Random match probability modeling for phenomena like toolmarks or handwriting is many times more complicated than for DNA, where we are blessed with an underlying system that comes in basically binary units, and DNA has been hard enough to get control of. But there is an alternative to random match probability data—black box testing of the practitioners in the field. Of course, little of this has been done either, but the fact remains that it is a viable research alternative to statistical probability modeling.

Global attacks encourage a “circle the wagons” attitude within the forensic science community and may impede any hope of general reform by the hardening of battle lines. In the end, it is well to remember that in a simple us-or-them fight with law enforcement, law enforcement will almost certainly win.

And some of the expert applications are so likely to be valid that attacking them seems feckless even without specific testing of either the random match probability modeling or of the black box variety. In principle, if we were to absolutely require random match probability data as a condition of admissibility, we would be forced to exclude rolled ten-print matches as a means of identifying an unknown cadaver. Suspicion of such a result can easily drive a judge to regard a global challenge as some sort of trick.

Not to say that judges have performed well in dealing with challenges to prosecution-proffered forensic science. As a group, they most certainly have not. Like lawyers in general, judges rarely have technical backgrounds sufficient to understand the validity issues at stake in a challenge on their own. In addition, there seems to be a clear institutional bias against excluding prosecution-proffered expertise. There are many possible reasons for this. In jurisdictions where judges are elected, a “tough on crime” reputation is often a significant asset. In addition, in all systems a high percentage of judges are former prosecutors. Aside from simple identification bias, it is easy to understand resistance to concluding that expertise you yourself have used in trying and convicting criminal defendants is fundamentally unsound and unreliable. But the problems go deeper. It is easy to see why defense lawyers make global challenges, but why do judges not require and foster more tailoring to the specific issues of the case, particularly in federal courts where such an approach is mandated by *Kumho Tire v. Carmichael*?

First, I believe judges have generally been as lax about actually bothering to read *Kumho Tire* as criminal defense attorneys have been. But even those few who have read it don’t really want to believe it. As I have recently written:

---

61 In this instance, I decline to list my friends who fall into this category. They know who they are.


63 For a detailed explanation of why there is a proper belief warrant for such ten print matches independent of statistical modeling or random match probability data, see Denbeaux & Risinger, *supra* note 59, at 67–70.
I believe that the lower courts have instinctively understood that such an approach would be a lot of work, would involve them in complex issues of empirical evaluation of claims beyond their comfort level if not their competence, and most importantly after all that work, would not yield results with much dispositive carry-over application to the next case arising, except in regard to narrow task issues that they were not even sure how to formulate. What they instinctively wanted was a system that yielded broad dispositive precedent, thus obviating the need for repetitively examining an area (like handwriting) in order to map out the real reliability contours of claimed expertise. In order to achieve that, they had to fail to see, never mind understand, the task-at-hand mandate of Kumho Tire, because only then could the legal questions be formulated and treated globally enough to be domesticated to a precedent-based system. And so they did. Then, when decisions, especially appellate decisions, were rendered, they were treated as strong authority in resolving the global issues of reliability (“Is handwriting expertise reliable enough to be admitted?” “Does handwriting expertise pass muster under Rule 702?”), even though the context of those decisions deprived them of any binding authority or claim to significant precedential weight under the official doctrines and assumptions of the precedent system. By doing this, the courts turned a process of reliability evaluation which the Supreme Court seems clearly to have envisioned as being empirical, data driven, and specific to the task being performed by the expert in the case at bar, into a single global issue resolvable at a swoop by reference to precedent for all the tasks undertaken [in a particular area].

Perhaps we can hope that the NAS/NRC Report can energize both defense attorneys and judges and change the current reality. Because the NAS/NRC Report

---

64 This is consistent with the response given to me at a conference some years ago by a respected federal district court judge, who came from a background in academia (not Jack Weinstein), whose response to me (on the topic of Kumho Tire and forensic science) was approximately, “If the Supreme Court thinks I’m going to have a hearing on every little question about the reliability of this stuff, they’re crazy.”

65 This is not a trivial issue, and it is reasonable for the courts to have this concern. The question is, how does a court define a case-specific task in such a way that it is narrow enough to meet the policies behind Kumho Tire but broad enough to yield at least some stare decisis carry-over effect. I believe that there are ways of systematically resolving the problem if it is faced explicitly, which Professor Denbeaux and I have set out in detail in Denbeaux & Risinger, supra note 59, at 35–55, particularly at 54–55. See also Roger C. Park, Signature Identification in Light of Science and Experience, 59 HASTINGS L.J. 1101 (2008) (recognizing the problems of defining the specific task sufficiently broadly to have some appropriate carry-over effect).

66 Risinger, supra note 39, at 472.
identifies problems that can’t be ignored, and supplies solutions that are not likely to work, judges and lawyers will have to do better in the trial setting, or we are unlikely to see any real reform in the foreseeable future. But nothing would do more to incentivize the forensic science community to accept and cooperate with honestly designed research efforts like the actual exclusion of a few of the weakest applications of the problem areas. Exclusion is a blunt instrument to try to coerce forensic science to reform, but in the end it may be the only one we are left with.
INDEPENDENT CRIME LABORATORIES:
THE PROBLEM OF MOTIVATIONAL AND COGNITIVE BIAS

Paul C. Giannelli*

INTRODUCTION

One of the most controversial recommendations in the National Academy of Sciences’ report on forensic science, *Strengthening Forensic Science in the United States: A Path Forward* (NAS Report), 1 concerns the removal of crime laboratories from the administrative control of law enforcement agencies. 2 According to the NAS Report:

The best science is conducted in a scientific setting as opposed to a law enforcement setting. Because forensic scientists often are driven in their work by a need to answer a particular question related to the issues of a particular case, they sometimes face pressure to sacrifice appropriate methodology for the sake of expediency. 3

For decades, scholars have commented on the “inbred bias of crime laboratories affiliated with law enforcement agencies” 4—as have courts, 5

---

1 National Research Council of the National Academies, *Strengthening Forensic Science in the United States: A Path Forward* (2009) [hereinafter NAS Report]. The report’s recommendation for an independent federal entity, the National Institute of Forensic Science, is also controversial. *Id.* at 19–20 (Recommendation 1).

2 *Id.* at 24 (Recommendation 4). The report also states: “Scientific and medical assessment conducted in forensic investigations should be independent of law enforcement efforts either to prosecute criminal suspects or even to determine whether a criminal act has indeed been committed. Administratively, this means that forensic scientists should function independently of law enforcement administrators.” *Id.* at 23.

3 *Id.* at 23–24.

legislators, prosecutors, investigators, and reporters. The NAS Report is not the first to acknowledge the problem of bias. The National Academy of Sciences’ 1996 DNA Report observed that “[l]aboratory procedures should be designed with safeguards to detect bias and to identify cases of true ambiguity.” Similarly, the ABA Standards on DNA Evidence contain a provision on bias.  

The problem of bias in crime laboratories is not unique to the United States. According to a British court:

Forensic scientists may become partisan. The very fact that the police seek their assistance may create a relationship between the police and the forensic scientists. And the adversarial character of the proceedings tends to promote this process. Forensic scientists employed by the government may come to see their function as helping the police.

---

6 See Rodney Ellis, Editorial, Want Tough on Crime? Start by Fixing HPD Lab., HOUS. CHRON., Sept. 5, 2004 (“When crime labs are operating within a police department, examiner bias can undermine the integrity of scientific results.”). Ellis was a Texas state senator at the time he wrote the editorial. See id.  
7 See Scott Bales, Turning the Microscope Back on Forensic Scientists, 26 LITIG. 51, 55 (2000) (“But whether nefarious or innocent, too close a connection between scientists and the law enforcement officers with whom they work creates a real danger of biased testimony.”). As an assistant U.S. attorney, Justice Bales served on the team that produced the 1997 I.G. Report on the FBI lab. See infra text accompanying notes 38–39. He is now a justice on the Arizona Supreme Court.  
8 See M.A. Thomson, Bias and Quality Control in Forensic Science: A Cause for Concern, 19 J. FORENSIC SCI. 504, 509–10 (1974) (“Is the witness who has his job and salary controlled by the State completely free from pressure, conscious or unconscious, to be entirely impartial?”). Captain Thomson was an Air Force investigator at the time he wrote this article. See id. at 504 n.1.  
11 AMERICAN BAR ASSOCIATION, ABA STANDARDS FOR CRIMINAL JUSTICE: DNA EVIDENCE 67 (3d ed. 2007) [hereinafter ABA DNA STANDARDS] (“Cognitive bias (e.g., observer effects) occurs because people tend to see what they expect to see, and this typically affects their decision in cases of ambiguity.”), available at http://www.abanet.org/crimjust/standards/dnaevidence.pdf.
One commentator concluded that the miscarriages of justice in Britain constituted “unequivocal evidence that the pro-prosecution orientation of government scientists . . . had not adequately been countered in England.”

Some commentators have proposed independent laboratories as the remedy for this problem, and in 2002, the Illinois Governor’s Commission on Capital Punishment proposed the establishment of an independent state crime laboratory. The Commission majority believed that “the overall quality of forensic services would be improved if the laboratory personnel were truly independent.” In contrast, the Department of Justice and the National District Attorneys Association oppose the NAS recommendation of independent laboratories.

This Essay examines the issue of independent crime laboratories. Part I documents the problems that triggered the NAS Report’s recommendation, while Part II explores the counterarguments. Part III examines the NAS

---

13 Ian Freckelton, Science and the Legal Culture, 2 EXPERT EVID. 107, 112 (1993); see also David E. Bernstein, Junk Science in the United States and the Commonwealth, 21 YALE J. INT’L L. 123, 171 (1996) (“Many reformers in the United Kingdom believe that a large percentage of the problems that have arisen in the forensic science context are attributable to the fact that English forensic science is almost solely the province of the state.”); Paul Roberts, Forensic Science Evidence After Runciman, 1994 CRIM. L. REV. 780, 784 (commenting that “forensic scientists who run with the hounds cannot be expected to give a savaged fox the kiss of life”) (citing Russell Stockdale, Running with the Hounds, NEW L.J. 772 (June 7, 1991)).

14 See BARRY SCHECK ET AL., ACTUAL INNOCENCE: FIVE DAYS TO EXECUTION, AND OTHER DISPATCHES FROM THE WRONGLY CONVICTED 257 (2000) (stating laboratories should “function as an independent third force within the criminal justice system”); Giannelli, supra note 4, at 457–62 (arguing for labs associated with a medical examiner system); see also Ellis, supra note 6 (stating “crime labs should operate as a separate and independent third party force in the criminal justice system”).

15 REPORT OF THE GOVERNOR’S COMMISSION ON CAPITAL PUNISHMENT 52 (2002), available at http://www.idoc.state.il.us/ccp/ccp/reports/commission_report/chapter_03.pdf [hereinafter CAPITAL PUNISHMENT COMM.] (“An independent state forensic laboratory should be created, operated by civilian personnel, with its own budget, separate from any police agency or supervision.”). The proposal was never adopted.

16 Id.

17 Strengthening Forensic Science in the United States: A Path Forward: Hearing Before the Subcomm. on Crime, Terrorism, and Homeland Security of the H. Comm. on the Judiciary, 111th Cong. 13 (2009) (statement of Kenneth E. Melson, Acting Dir., Bureau of Alcohol, Tobacco, Firearms, and Explosives) (“DOJ also questions whether full independence of laboratories from law enforcement is advisable or feasible . . . . To be separated completely from interaction with investigative partners would likely cause missteps in decision-making that could result in either loss and/or destruction of evidence, or important analyses left undone.”).

18 National District Attorneys Association, NDAA Comments Provided to the Consortium of Forensic Sciences Regarding the National Academy of Sciences Report [hereinafter NDAA Statement] (“NDAA does not believe, as some have suggested, that all forensic labs must be ‘independent,’ that is, housed outside of a law enforcement or prosecution agency.”).
proposal as well as an alternative approach. Part IV sets forth additional measures that should protect forensic analyses from improper influence.

I. THE PROBLEM

A. Organizational Structure

Crime laboratories are “the oldest and strongest link between science and technology and criminal justice.” 19 In the United States, crime laboratories developed in the 1920s as an adjunct of police departments. 20 A survey of approximately three hundred crime laboratories revealed that “[s]eventy-nine percent of all laboratories responding . . . are located within law enforcement/public safety agencies” 21 and “[f]ifty-seven percent . . . would only examine evidence submitted by law enforcement officials.” 22 Thus, it is not surprising that police norms would influence the laboratory culture. As one scholar observed: “[T]he police agency controls the formal and informal system of rewards and sanctions for the laboratory examiners.” 23

B. Types of Bias

Commentators have identified both motivational and cognitive bias as a concern in the forensic setting. 24 These classifications are not mutually exclusive, and cognitive bias comes in several forms. 25

---

20 See Richard Saferstein, Criminalistics: An Introduction to Forensic Science 6 (5th ed. 1995) (“The oldest forensic laboratory in the United States is that of the Los Angeles Police Department, created in 1923 by August Vollmer, a police chief from Berkeley, California.”); John I. Thornton, Criminalistics—Past, Present and Future, 11 Lex et Scientia 1, 23 (1975) (“In 1923, Vollmer served as Chief of Police of the City of Los Angeles for a period of one year. During that time, a crime laboratory was established at his direction.”); see also Bales, supra note 7, at 55 (“The tie between crime labs and law enforcement agencies is not inevitable. In part, it is a product of history: rudimentary crime labs were first established near the turn of the century by law enforcement agencies when officials began to recognize the possible application of science to criminal investigations. Since that time, the relationship between labs and law enforcement has flourished because of practical benefits—for example, streamlining tasks such as close and timely communication, the transfer of evidence, and record-keeping.”).
22 Id. at 13.
1. Motivational Bias

Motivational bias “is close to the popular notion of bias (the referee is biased because he wants one side to win).”26 Several notorious examples seem to fit within this category. For example, Fred Zain, who became infamous because of his misconduct at the West Virginia state crime laboratory, routinely reported results that favored the prosecution.27 An investigation by the American Society of Crime Laboratory Directors/Laboratory Accreditation Board (ASCLD/LAB), found that, “when in doubt, Zain’s findings would always inculpate the suspect.”28 His replacement as director of serology described Zain as “very pro-prosecution.”29 Zain was such a treasured witness that, even after he left the state to accept a position in a San Antonio crime laboratory, West Virginia prosecutors sent evidence to him for retesting.30 The prosecutors relied on Zain because the remaining West Virginia serologists were incapable, in their view, of reaching the “right” results.31

While working at the Oklahoma City Crime Laboratory for nearly twenty years, Joyce Gilchrist repeatedly overstated test results, withheld evidence, and provided critical evidence for the prosecution.32 The Court of Appeals for the Tenth Circuit criticized Gilchrist for “provid[ing] the jury with evidence

several types of cognitive bias, including observer effects, anchoring effects, role effects, conformity effects, and experimenter effects).

26 REDMAYNE, supra note 24, at 14.
27 Zain falsified test results in as many as 134 cases from 1979 to 1989. See In re Investigation of the W.Va. State Police Crime Lab., Serology Div., 438 S.E.2d 501, 510–11 (W.Va. 1993). In reviewing a judicial report on Zain’s decade of misconduct, the West Virginia Supreme Court spoke of “shocking and . . . egregious violations” and the “corruption of our legal system.” Id. at 508. The judicial inquiry concluded that “as a matter of law, any testimonial or documentary evidence offered by Zain at any time in any criminal prosecution should be deemed invalid, unreliable, and inadmissible.” Id. at 520; see generally Paul C. Giannelli, Wrongful Convictions and Forensic Science: The Need to Regulate Crime Labs, 86 N.C. L. REV. 163, 172–74 (2007) (discussing Zain’s conduct).
28 In re Investigation of W.Va., 438 S.E.2d at 512 n.9.
29 Id. at 514 n.23.
30 His work in Texas also proved troublesome: “In the case of Gilbert Alejandro, the expert, Fred Zain claimed a DNA match when in fact Zain had never conducted any testing beyond initial inconclusive testing, and final DNA testing conducted after the trial excluded Alejandro.” Brandon L. Garrett, Judging Innocence, 108 COLUM. L. REV. 55, 84 n.109 (2008).
31 According to Zain’s replacement, “several prosecutors expressed dissatisfaction with the reports they were receiving from serology and specifically requested that the evidence be analyzed by Zain.” In re Investigation of W. Va., 438 S.E.2d at 513 n.16 (referring to deposition of T.S. Smith). “[Serologist] Myers also testified that after he had been unable to find blood on a murder suspect’s jacket, it was sent to Texas, where Zain found a bloodstain which tested consistent with the blood of the victim.” Id. at 512. “[Serologist] Bowles also testified that at least twice after Zain left the lab, evidence on which Bowles had been unable to obtain genetic markers was subsequently sent to Texas for testing by Zain, who again was able to identify genetic markers.” Id.
32 See Mitchell v. Gibson, 262 F.3d 1036, 1064 (10th Cir. 2001).
implicating [a defendant] in the sexual assault of the victim which she knew was rendered false and misleading by evidence withheld from the defense. As one author commented: “If [Gilchrist] were simply incompetent, her mistakes would have been all over the map. Instead, her mistakes benefited the prosecution.”

2. Cognitive Bias: Role Effects

If the motivation is subconscious, the bias can be classified as a type of cognitive bias called “role effect” bias. In short, people’s perception of their role can influence their decisions, especially in cases of ambiguity. “Given what is known about reference group phenomena, the need that people have for social support of attitudes and conduct, and the process of socialization in occupational settings, it strains credulity to believe that these experts do not identify with prosecutors.” According to a former laboratory director, “Many forensic scientists at the state police labs . . . saw their role as members of the state’s attorney’s team. ‘They thought they were prosecution witnesses[,]’ . . . ‘They didn’t understand they were just scientists.’”

In 1997, the Inspector General of the Department of Justice issued a report on the FBI laboratory’s explosives unit. This report documented numerous deficiencies, including inaccurate testimony, testimony beyond the competence of examiners, improperly prepared laboratory reports, insufficient documentation of test results, inadequate record management and retention, and failure to resolve serious and credible allegations of incompetence.

---

33 Id.; see generally Giannelli, supra note 27, 174–82 (discussing Gilchrist’s conduct).
34 MARK FUHRMAN, DEATH AND JUSTICE: AN EXPOSE OF OKLAHOMA’S DEATH ROW MACHINE 223 (2003). Fuhrman also wrote that Gilchrist “appears to have used her lab tests to confirm the detectives’ hunches rather than seek independent scientific results. . . . She treated discovery requests with contempt and kept evidence from the defense. She systematically destroyed evidence at the very time when she knew that much of that evidence might be retested.” Id. at 232.
35 See Risinger et al., supra note 25, at 18–19.
37 Mills et al., supra note 9 (quoting Don Plautz, a former director in the Illinois crime lab system); see also Teichroeb, supra note 9 (explaining that crime labs are often biased in favor of the prosecution).
39 Id.; see also JOHN F. KELLY & PHILLIP K. WEARNE, TAINTING EVIDENCE 2 (1998) (concluding that FBI examiners “had given scientifically flawed, inaccurate, and overstated testimony under oath in court; had altered the lab reports of examiners to give them a pro-prosecutorial slant, and had failed to document tests and examinations from which they drew incriminating conclusions, thus ensuring that their work could never be properly checked”; Bales, supra note 7, at 53 (“[T]he [1997 I.G. Report] did contain deeply disturbing findings of inadequate procedures, insufficient
Oklahoma City bombing case, the I.G. Report found that an examiner’s conclusion about the identity of the explosive charge was “speculation” and “tilted in such a way as to incriminate the defendants.”

3. **Cognitive Bias: Contextual Bias**

Another type of cognitive bias is contextual bias, which occurs when extraneous information influences a decision, typically in cases of ambiguity. When clinical trials for a new drug are conducted, “double blind” procedures are used—i.e., randomized clinical trials. Neither the patient nor the physician knows whether the patient is receiving the new drug or a placebo (the control). Numerous studies have demonstrated that physicians who know that patients are receiving a new drug tend to see positive results, even when there are none. In short, extraneous knowledge alters our expectations, which in turn affects our perceptions.

There is no shortage of examples: “[Professor] Peter DeForest has described investigators who responded to inconclusive results by saying to forensic examiners: ‘Would it help if I told you we know he’s the guy who did it?’” One laboratory examiner “said she tried not to be swayed by detectives’ belief that they had a strong suspect. ‘We’re all human,’ she said. ‘I tried not to

---


---

Redmayne, supra note 24, at 15 (“It also appears that extraneous information supporting a hypothesis will affect our judgement of that hypothesis, and of the evidence for it, even when we know we should not take the extraneous information into account.”).

---

See Robert J. Levine, *Ethics and Regulation of Clinical Research* 185 (2d ed. 1986) (“When it is feasible, a double-blind technique is employed. That is, neither the investigator nor the subject knows until the conclusion of the study who is in the treatment or control group. The purpose of double-blinding is to overcome biases on the part of both subjects and investigators . . . .”).

---

Risinger et al., supra 25, at 45 (“The simplest, most powerful, and most useful procedure to protect against the distorting effects of unstated assumptions, collateral information, and improper expectations and motivations is blind testing. An examiner who has no domain-irrelevant information cannot be influenced by it. An examiner who does not know what conclusion is hoped for or expected of her cannot be affected by those considerations.”).

---

See id. at 39. The psychological literature on lineups provides another illustration. Eyewitnesses with reservations about their identifications often become positive after learning that the person they identified was the prime suspect in the case. See Report of the ABA Criminal Justice Section’s Ad Hoc Innocence Comm. to Ensure the Integrity of the Criminal Process, Achieving Justice: Freeing the Innocent, Convicting the Guilty 37 (Paul C. Giannelli & Myrna Raeder eds., 2006) (“Ideally, the witness should never be told whether he selected the ‘right man’ so that his confidence is not artificially inflated by the time of trial.”).
let it influence me. But I can’t say it never does.”

Joyce Gilchrist often received detectives’ views on suspects before she conducted her examinations. In another case, an FBI examiner identified a substance as being consistent with an explosive “based in part on the fact that pieces of cut detonation cord had been found in a garbage can outside the suspect’s house.”

4. Cognitive Bias: Confirmation Bias

Another type of cognitive bias known as “confirmation bias” concerns “the tendency to test a hypothesis by looking for instances that confirm it rather than by searching for potentially falsifying instances.” Confirmation bias played a role in the FBI’s misidentification of Brandon Mayfield’s fingerprints in the Madrid terrorist train bombing investigation. According to an FBI review, the “power” of the automated fingerprint correlation “was thought to have influenced the examiner’s initial judgment and subsequent examination.” Three other experts, one of whom was court-appointed, also confirmed the initial misidentification. These reviews were not conducted blind—i.e., the reviewer knew that a positive identification had already been made—and thus were subject to the influence of contextual/confirmation bias.

---


46 See FUHRMAN, supra note 34, at 91 (“When Cook and other homicide detectives gave Gilchrist hair samples from a suspect, they would often let her know that this was the person they wanted to arrest.”).

47 Bales, supra note 7, at 55 (“Of course, where the cord was found was irrelevant to the scientific examination of the residue on the knife and to the examiner’s conclusions. . . . Based on recommendations by the OIG, the FBI has instructed its examiners not to base forensic conclusions on unstated assumptions or information that is collateral to the examinations performed.”); see also id. at 52 (The 1997 I.G. Report “concluded that an examiner from the lab’s explosives unit had erred by purporting to identify the particular explosives used in the [1993] World Trade Center and Oklahoma City bombings. The error stemmed from the examiner’s reliance on information that was tied to suspects but not relevant to his scientific analysis.”).

48 Risinger et al., supra note 25, at 7; see also REDMAYNE, supra note 24, at 15 (“We tend to look for confirming, rather than disconfirming, evidence; we may judge evidence of better quality if it agrees with our theory, or worse quality if it does not; and our beliefs can persevere even after being discredited.”).

49 See Sara Kershaw, Spain and U.S. at Odds on Mistaken Terror Arrest, N.Y. TIMES, June 5, 2004, at A1 (Spanish authorities cleared Brandon Mayfield and matched the fingerprints to an Algerian national).


51 Id. at 709–11.

52 Id. at 713.
5. **Cognitive Bias: Reconstructive Effects**

Another type of cognitive bias involves “reconstructive effects.” The when people rely on their memory, they tend to fill in gaps with what they believe should have occurred. One of the Inspector General’s reports on the FBI laboratory addressed this issue: “[C]ontemporaneous documentation is important to ensure that the case file accurately reflects the work performed on each evidence item that is tested. . . . [S]taff members may be unduly influenced by protocol requirements when relying on memory, and document what they know should have occurred when their recollection is vague.”

6. **Research**

Although the psychological literature on cognitive bias is well developed, research in forensic science has lagged. One researcher performed a rudimentary experiment involving handwriting comparisons in 1984 and then followed up with a study on hair examinations in 1987. Although Professor Jonakait mentioned the topic in a 1991 law review article, the issue was thrust to the forefront when Professor Risinger and his colleagues published an extensive article on the subject in 2002. As a result of the Mayfield case, British researchers devised a covert experiment to test contextual bias. Five fingerprint examiners who were unfamiliar with the Mayfield prints were asked by colleagues to compare a crime scene print and suspect print. “They were told that the pair of prints

---

53 Risinger et al., *supra* note 25, at 15–16 (providing the example of a “forensic scientist who takes poor notes during an examination and prepares a skimpy report, but then goes back to ‘spruce them up’ shortly before trial”).


57 Larry S. Miller, *Procedural Bias in Forensic Science Examinations of Human Hair*, 11 *L. & Hum. Behav.* 157, 161 (1987). In the conventional method of hair examination, the examiner is given hair samples from a known suspect along with a report including information relating to the guilt of the suspect. In the study on hair examinations, the findings “raise some concern regarding the amount of unintentional bias among human hair identification examiners . . . . A preconceived conclusion that a questioned hair sample and a known hair sample originated from the same individual may influence the examiner’s opinion when the samples are similar.” *Id.* at 161.

58 Jonakait, *supra* note 4, at 160–64.

59 Risinger et al., *supra* note 25.


61 *Id.* at 75.
was the one that was erroneously matched by the FBI as the Madrid bomber, thus creating an extraneous context that the prints were a non-match. The participants were then instructed to ignore this information. The prints, in fact, were not from the Mayfield case; they were from cases that each of the participants had previously matched. Of the five examiners, only one still judged the print to be a match. The other four changed their opinions; three directly contradicted their prior identifications, and the fourth concluded that there was insufficient data to reach a definite conclusion. “This is striking given that all five experts had seen the identical fingerprints previously and all had decided that the prints were a sound and definite match.”

A follow-up covert study, which also involved experts, showed that fingerprint examiners could be biased toward a finding of identification if informed that the suspect confessed or toward a finding of exclusion if told that the suspect had an alibi. Another investigation focused on the effects of emotions on decision making.

Because the research in the forensic field is in its nascent stage, the NAS Report recommends further investigation of observer bias and other sources of

---

62 Id. at 76.
63 Id.
64 Id. at 75.
65 Id. at 76.
66 Id.
67 Id. The authors of the study concluded the “study shows that it is possible to alter identification decisions on the same fingerprint, solely by presenting it in a different context. This does not imply that fingerprint and other forensic identifications are not a science, but it does highlight problems of subjectivity, interpretation, and other psychological and cognitive elements that interact and may distort any scientific inquiries.” Id. at 77.

69 Itiel Dror et al., When Emotions Get the Better of Us: The Effect of Contextual Top-Down Processing on Matching Fingerprints, 19 APPL. COGNIT. PSYCHOL. 799, 806–07 (2005) (“The results of this study demonstrated that emotion and subliminal messages did influence decision making[,]” but not in clear-cut cases).

human error in forensic examinations. Cognitive bias is most likely a far greater danger than motivational bias precisely because it is a subconscious influence. Forensic techniques that have a substantial subjective component should be a special concern—e.g., fingerprint identifications, firearms (ballistics) identifications, and handwriting comparisons.

C. The Prosecutor

Of course, the police are not the only ones who may influence government experts. Prosecutors also have pressured experts to slant their testimony.

---

71 NAS REPORT, supra note 1, at 24 (Recommendation 5) (“Such programs might include studies to determine the effects of contextual bias in forensic practice (e.g., studies to determine whether and to what extent the results of forensic analyses are influenced by knowledge regarding the background of the suspect and the investigator’s theory of the case).”).

72 See REDMAYNE, supra note 24, at 14 (“Cognitive biases are potentially more problematic, for these result from unconscious reasoning strategies that can lead us to unwarranted conclusions.”); Dror & Cole, supra note 70, at 162 (“Errors committed by well-intentioned experts are more problematic and dangerous . . . .”); Risinger et al., supra note 25, at 11 (finding cognitive bias “far more pervasive but generally unnoticed” and “a problem in some respects more troublesome and troubling than the intentional misconduct”).

73 See Commonwealth v. Patterson, 840 N.E.2d 12, 16–17 (Mass. 2005) (“In the evaluation stage, . . . the examiner relies on his subjective judgment to determine whether the quality and quantity of those similarities are sufficient to make an identification, an exclusion, or neither.”); Sandy L. Zabell, Fingerprint Evidence, 13 J. L. & POL’Y 143, 158 (2005) (“In contrast to the scientifically-based statistical calculations performed by a forensic scientist in analyzing DNA profile frequencies, each fingerprint examiner renders an opinion as to the similarity of friction ridge detail based on his subjective judgment.”).

74 See United States v. Glynn, 578 F. Supp. 2d 567, 571 (S.D.N.Y. 2008) (“[T]he Government did not seriously contest the Court’s conclusions that ballistics lacked the rigor of science and that, whatever else it might be, its methodology was too subjective to permit opinions to be stated to ‘a reasonable degree of ballistic certainty.’”); United States v. Monteiro, 407 F. Supp. 2d 351, 372 (D. Mass. 2006) (“Because an examiner’s bottom line opinion as to an identification is largely a subjective one, there is no reliable statistical or scientific methodology which will currently permit the expert to testify that it is a ‘match’ to an absolute certainty, or to an arbitrary degree of statistical certainty.”).

75 See United States v. Starzecpyzel, 880 F. Supp. 1027, 1048 (S.D.N.Y. 1995) (“Such [overly fine] distinctions are certainly improper in forensic document examination, where it is conceded that conclusions are drawn, in large part, on subjective criteria.”).

76 ABA Criminal Justice Standards state that “[a] prosecutor who engages an expert for an opinion should respect the independence of the expert and should not seek to dictate the formation of the expert’s opinion on the subject. . . . [T]he prosecutor should explain to the expert his or her role in the trial as an impartial expert . . . .” ABA STANDARDS FOR CRIMINAL JUSTICE, PROSECUTION FUNCTION AND DEFENSE FUNCTION 58 (Standard 3-3.3(a)) (3d ed. 1993), available at http://www.abanet.org/crimjust/standards/prosecutionfunction.pdf. A comparable
For more than a decade, a Texas pathologist worked closely with prosecutors and police “‘shad[ing] things to follow along with the police theory of the case.’”78 As the special prosecutor remarked: “If the prosecution theory was that death was caused by a Martian death ray, then that was what [the pathologist] reported.”79

In one of Joyce Gilchrist’s cases, an appellate court wrote: “[W]e are greatly disturbed by the implications that the Oklahoma County District Attorney’s Office may have placed undue pressure upon Ms. Gilchrist to give a so-called expert opinion, which was beyond scientific capabilities . . . .”80 In Troedel v. Wainwright,81 a capital murder case, the court found that a FBI expert shaped his testimony in a way that was “at the very least, . . . standard applies to defense counsel. See id. at 188 (Standard 4-4.4(a)). The commentary to this standard elaborates: “Statements made by physicians, psychiatrists, and other experts about their experiences as witnesses in criminal cases indicate the need for circumspection on the part of lawyers who engage experts. Nothing should be done by a lawyer to cast suspicion on the process of justice by suggesting that the expert color an opinion to favor the interests of the client the lawyer represents.” Id. at 189.

77 See generally Paul C. Giannelli & Kevin C. McMunigal, Prosecutors, Ethics, and Expert Witnesses, 76 FORDHAM L. REV. 1493, 1520–27 (2007) (discussing the problems associated with prosecutors and experts). Experts often are pressured by attorneys to “push the envelope”—not a surprising occurrence in an adversary system. See SCHECK ET AL., supra note 14, at 31 (“Most attorneys . . . like to let you know what their opinions of the facts of the case are — irrespective of the scientific conclusions.”) (quoting Dr. Robert Shaler, former head of N.Y.C. Medical Examiner’s DNA unit).

78 Roberto Suro, Ripples of a Pathologist’s Misconduct in Graves and Courts of West Texas, N.Y. TIMES, Nov. 22, 1992, at A22 (quoting Tommy J. Turner, appointed by a state district judge to investigate Dr. Ralph R. Erdmann). “[A]ll the while [Dr. Erdmann] worked in close collaboration with many prosecutors and police officials, some of whom are now prominent in politics.” Id.; see also Roy Bragg, Autopsy Record of Pathologist Who Quit Raises Many Eyebrows, HOUS. CHRON., Mar. 8, 1992, at A1 (Dr. Linda Norton, a former Dallas County assistant medical examiner, stated: “It’s as though there’s some sort of collusion between Dr. Erdmann and the DA.”).


80 McCarty v. State, 765 P.2d 1215, 1219 (Okla. Crim. App. 1988). There, the court ultimately held that despite these concerns, it “could not conclude . . . that appellant has established the prosecution’s knowing use of false or misleading evidence.” Id.; see also Bank of Nova Scotia v. United States, 487 U.S. 250, 258 (1988) (“The District Court further concluded that one of the prosecutors improperly argued with an expert witness during a recess of the grand jury after the witness gave testimony adverse to the Government.”).

misleading.”82 The expert claimed that the prosecutor had “pushed” him to enhance his testimony, a claim the prosecutor substantiated.83 Consequently, removing the crime laboratory from police control still leaves a problem of prosecutorial influence, albeit perhaps lessened.

II. COUNTERARGUMENTS

There are several criticisms of the proposal for establishing independent laboratories, which are discussed in this Part.

A. Integration with Police Investigative Function

A forensic laboratory may play an important role in the early stages of a criminal investigation. As two commentators have noted: “Increasing the laboratory’s geographical or organizational remoteness . . . can limit the effectiveness of the laboratory’s participation in the investigative phases of a case, when its scientific input may have the greatest chance of contributing to justice.”84 This argument raises a serious concern. However, homicide

82 Id. at 1459. The expert’s report of a gunshot residue test concluded that swabs “from the hands of Troedel and Hawkins contained antimony and barium in amounts typically found on the hands of a person who has discharged a firearm or has had his hands in close proximity to a discharging firearm.” Id. at 1458. The expert testified in accordance with this report at Hawkins’s trial but enhanced his testimony at Troedel’s trial, where he testified that “Troedel had fired the murder weapon.” Id. In contrast, during federal habeas proceedings, the same expert testified in a deposition that “he could not, from the results of his tests, determine or say to a scientific certainty who had fired the murder weapon” and “the differences in the amount of barium and antimony on the hands of Troedel and Hawkins were basically insignificant.” Id. at 1459. In granting habeas relief, the court “conclude[d] that the opinion Troedel had fired the weapon was known by the prosecution not to be based on the results of the neutron activation analysis tests, or on any scientific certainty or even probability. Thus, the subject testimony was not only misleading, but also was used by the State knowing it to be misleading.” Id. at 1459–60.

83 Id. at 1459 (“[A]s Mr. Riley candidly admitted in his deposition, he was ‘pushed’ further in his analysis at Troedel’s trial than at Hawkins’ trial. . . . [A]t the . . . evidentiary hearing held before this Court, one of the prosecutors testified that, at Troedel’s trial, after Mr. Riley had rendered his opinion which was contained in his written report, the prosecutor pushed to ‘see if more could have been gotten out of this witness.’”).

84 Jan S. Bashinski & Joseph L. Peterson, Forensic Sciences, in LOCAL GOVERNMENT: POLICE MANAGEMENT 559, 581 (William Geller & Darrel Stephens eds., 4th ed. 2004). Bashinski and Peterson state: “Remoteness also makes the police department less able to direct the efforts of the laboratory toward the cases that the department considers most important . . . .” Id.; see also Bales, supra note 7, at 55 (“[T]he relationship between labs and law enforcement has flourished because of practical benefits—for example, streamlining tasks such as close and timely communication, the transfer of evidence, and record-keeping.”); NDAA Statement, supra note 18 (“We believe that laboratories housed within government agencies and whose mission is focused on public safety are likely to be more responsive and accountable to those community needs than those situated otherwise.”).
detectives work closely with medical examiner officials in death investigations, and medical examiner offices are typically independent of the police.

B. Practicability

According to a 2005 census, there are now 389 publicly funded crime laboratories in the United States: 210 state or regional laboratories, eighty-four county laboratories, sixty-two municipal laboratories, and thirty-three federal laboratories. Some of these laboratories are quite small: “The median staff size in 2005 was 16.” This suggests that some laboratories could probably not exist as an independent entity.

C. Funding

Because underfunding of crime laboratories in this country is chronic, resources are always an issue. The minority report of the Illinois Capital Punishment Commission argued that funding for the state laboratory would be jeopardized if it were separated from the police:

This new agency will have to compete with other, larger agencies for scarce state resources. Retaining the forensic laboratory system as part of the Illinois State Police provides an opportunity for achieving economies of scale and administration, as well as security in funding and accountability that might not otherwise be available for a much smaller, stand-alone agency left to fend for itself.

In contrast, the NAS Report assumed that laboratory independence would protect a laboratory’s budget. According to the report, law enforcement control “leads to significant concerns related to the independence of the laboratory and its budget.” Under this view, independence would mean “the forensic science laboratories would be able to set their own budget priorities and not have to compete with the parent law enforcement agencies.”

---

86 Id. at 2.
87 See PRESIDENT’S COMMISSION ON LAW ENFORCEMENT AND ADMINISTRATION OF JUSTICE, THE CHALLENGE OF CRIME IN A FREE SOCIETY 255 (1967) (“[T]he great majority of police department laboratories have only minimal equipment and lack highly skilled personnel able to use the modern equipment now being developed . . . .”); NATIONAL ADVISORY COMMISSION ON CRIMINAL JUSTICE STANDARDS AND GOALS, REPORT ON POLICE 304 (1974) (“Too many police crime laboratories have been set up on budgets that preclude the recruitment of qualified, professional personnel.”).
88 CAPITAL PUNISHMENT COMM., supra note 15, at 54.
89 NAS REPORT, supra note 1, at 183–84.
90 Id. at 184; see also SHECK ET AL., supra note 14, at 257 (“Crime laboratory budgets should be independent from the police . . . .”).
Because of the diversity of crime laboratories and their funding structures, it is almost impossible to predict how funding would be affected if laboratories became independent.

D. Efficacy of Reform

The minority report of the Illinois Capital Punishment Commission also argued that an independent laboratory would not solve the problem of police influence. Because police and prosecutors use crime laboratories far more than defense attorneys do, the minority believed that close relationships were inevitable.91 There is some merit in this position. Yet there is a difference between working with someone, even extensively, and working with someone who is a superior (or works for a superior) within the same organization.

III. THE NAS PROPOSAL

The NAS Report recommends only that “administrative control” of the laboratory be removed from law enforcement agencies or prosecutors.92 The report went on to explain:

Ideally, public forensic science laboratories should be independent of or autonomous within law enforcement agencies. In these contexts, the director would have an equal voice with others in the justice system on matters involving the laboratory and other agencies. The laboratory also would be able to set its own priorities with respect to cases, expenditures, and other important issues.93

In other words, the goal is for a laboratory to have sufficient autonomy to protect the integrity of the laboratory’s findings. As a byproduct of a laboratory controversy,94 the Virginia legislature in 2005 made the Division of Forensic Science a separate department under the Secretary of Public Safety.95 The

91 CAPITAL PUNISHMENT COMM., supra note 15, at 53 (“The reality is that no matter how ‘independent’ this separate state agency is, the bulk of its work will still be for police agencies and prosecutors. As is true today for the vast majority of cases, the forensic experts will be called to testify by the prosecution and these experts will undoubtedly continue to be subject to cross-examination for that testimonial history. As a result, an ‘independent’ laboratory will be subject to criticism as a ‘police/prosecutor’ lab even if it is not under the direct control and management of a police agency, because of the nature of its day to day work.”).

92 NAS REPORT, supra note 1, at 24 (“Congress should authorize and appropriate incentive funds . . . for the purpose of removing all public forensic laboratories and facilities from the administrative control of law enforcement agencies or prosecutors’ offices.”).

93 Id. at 184.

94 See Giannelli, supra note 27, at 192–95 (discussing the Earl Washington Jr. case where a mentally retarded man had been convicted of a rape-murder and spent seventeen years in prison, only to be pardoned based on DNA evidence that was erroneously interpreted by experts).

laboratory had previously been under the Department of Criminal Justice Services. Although perhaps not a major change, this reorganization did increase laboratory autonomy.

The American Society of Crime Laboratory Directors proposed a different approach. That organization recommended that crime laboratories should not be removed “from parent agencies if the parent agency is required to document how crime laboratories have scientific autonomy with the freedom to conduct testing and report results without pressure from [external] activity, interest, or influence.”

In sum, the critical issue is for law enforcement and crime laboratories to acknowledge the problem and then to take steps to insulate the laboratory from improper influence.

IV. ADDITIONAL MEASURES

The problems raised by the law enforcement-crime laboratory relationship should also be addressed by the implementation of additional measures, many of which appear as other recommendations in the NAS Report. As one commentator noted: “To the extent that we are aware of our vulnerability to bias, we may be able to control it. In fact, a feature of good scientific practice is the institution of processes—such as blind testing, the use of precise measurements, standardized procedures, statistical analysis—that control for bias.”

First, case files need to document the laboratory analysis. The lack of bench notes was a significant problem in the laboratory scandals. For example, the Chicago, Houston, and FBI explosives unit investigations all found inadequate documentation in forensic case files.

Advisory Committee were created at the same time. VA. CODE ANN. § 9.1-1109 & 1111 (2005).


97 See Strengthening Forensic Science in the United States: Hearings Before the Comm. on the Judiciary, U.S. Senate, 111th Cong., Sept. 9, 2009 (statement of Matthew Redle, County and Prosecuting Attorney, Sheridan County, Wyoming) (discussing the importance of implementing quality control measures in laboratories such as “laboratory accreditation and personnel certification programs . . . ; internal peer review procedures; maintenance of appropriate testing documentation to facilitate internal and external peer review of individual case testing; external and internal performance audits; regular proficiency testing as a check on both personnel and protocol performance; and corrective action procedures when proficiency testing or casework errors are discovered”).

98 REDMAYNE, supra note 24, at 16 (footnote omitted).

99 Letter from Professor George F. Sensabaugh, University of California at Berkeley, to Locke E. Bowman, The MacArthur Justice Center, University of Chicago Law School 5 (Oct. 16, 2003) (on file with author) (“Overall, the documentation of the lab work as described in the three pages of lab notes is inadequate and incomplete. Moreover, the formal lab reports describe results of testing for which there is no record
Second, bench notes should be recorded contemporaneously with the examination. Otherwise, the examiner is subject to “reconstructive effects.”

Third, protocols should address contextual bias by shielding examiners from information that is not germane to the examination.

Fourth, comprehensive laboratory reports are necessary. Currently, laboratory reports often are “terse to the point of being indecipherable.” For example, some laboratory reports provide only a brief statement of the results in the lab notes. In short, the documentation in this case falls short of accepted scientific standards.”

See Michael R. Bromwich, Third Report of the Independent Investigator for the Houston Police Department Crime Laboratory and Property Room 28 (June 30, 2005) (“Among other problems it identified, the 2002 DPS audit found that no such written procedures [for case notes and lab reports] existed and identified numerous deficiencies in the documentation contained in the lab reports.”), available at http://www.hpdlabinvestigation.org/reports/050630report.pdf.

See 1997 I.G. REPORT, supra note 38 (recommending the preparation of adequate case files to support reports); Bales, supra note 7, at 57 (noting that one FBI examiner “testified that he had performed certain tests that were not described in his notes”).

See Law v. State, 307 S.E.2d 904, 908 (Ga. 1983) (Smith, J., dissenting) (“It is an insult to intelligent people to say that a scientific test was conducted from which absolutely no notes or records survive. . . . A basic principle of scientific testing is that careful records of test procedure and results are to be scrupulously maintained. A scientific test without an accompanying report of the testing environment, number of trials, raw results and analyzed data is in reality no test at all.”).

See supra notes 53–54 and accompanying text.

ABA DNA Standard 16-3.1(a)(v) recommends laboratories “follow procedures designed to minimize bias when interpreting test results.” ABA DNA STANDARDS, supra note 11, at 6. Additionally, “[c]ognitive bias (e.g., observer effects) occurs because people tend to see what they expect to see, and this typically affects their decisions in cases of ambiguity,” Id. at 67. See also NAS REPORT, supra note 1, at 26 (Recommendation 8: “Forensic laboratories should establish routine quality assurance and quality control procedures to ensure the accuracy of forensic analyses and the work of forensic practitioners. Quality control procedures should be designed to identify mistakes, fraud, and bias; confirm the continued validity and reliability of standard operating procedures and protocols; ensure that best practices are being followed; and correct procedures and protocols that are found to need improvement.”).

ABA DNA Standard 16-3.3 recommends the use of comprehensive laboratory reports. ABA DNA STANDARDS, supra note 11, at 7. The Journal of Forensic Sciences, the official publication of the American Academy of Forensic Sciences, published a symposium on the ethical responsibilities of forensic scientists in 1989. Symposium, Ethical Conflicts in the Forensic Science, 34 J. FORENSIC SCI. 717 (1989). One article discussed a number of laboratory reporting practices, including (1) “preparation of reports containing minimal information in order not to give the ‘other side’ ammunition for cross-examination,” (2) “reporting of findings without an interpretation on the assumption that if an interpretation is required it can be provided from the witness box,” and (3) “[o]mitting some significant point from a report to trap an unsuspecting cross-examiner.” Douglas M. Lucas, The Ethical Responsibilities of the Forensic Scientist: Exploring the Limits, 34 J. FORENSIC SCI. 719, 724 (1989). Lucas was the Director of the Centre of Forensic Sciences, Ministry of the Solicitor General, Toronto, Ontario. Id. at 719.

Bales, supra note 7, at 56.
"e.g., ‘The green, brown plant material in item #1 was identified as marijuana." In its recent decision, *Melendez-Diaz v. Massachusetts*, the Supreme Court noted that the report in that case contained only the bare-bones statement that ‘[t]he substance was found to contain: Cocaine.’ At the time of trial, petitioner did not know what tests the analysts performed, whether those tests were routine, and whether interpreting their results required the exercise of judgment or the use of skills that the analysts may not have possessed.

Fred Zain, Joyce Gilchrist, and Pam Fish, among others, omitted critical information from their reports.

Fifth, the reporting of test results should be accompanied by an explanation of the significance of any finding. A recent investigation of forensic testimony in DNA exoneration cases revealed that some experts gave misleading testimony by omitting critical information.

---

107 NAS REPORT, *supra* note 1, at 186.
108 129 S. Ct. 2527 (2009). The Court held that admission of a laboratory certificate identifying a substance as cocaine, in the absence of an opportunity to cross-examine the analyst, violated the Confrontation Clause. *Id.* at 2532.
109 *Id.* at 2537 (citations omitted).
111 Mitchell v. Gibson, 262 F.3d 1036, 1064 (10th Cir. 2001) (“Ms. Gilchrist thus provided the jury with evidence implicating Mr. Mitchell in the sexual assault of the victim which she knew was rendered false and misleading by evidence withheld from the defense.”); McCarty v. State, 765 P.2d 1215, 1218 (Okla. Crim. App. 1988) (“[T]he forensic report was at best incomplete, and at worst inaccurate and misleading. . . . Gilchrist admitted at trial, however, that she failed to include her conclusion . . . in the forensic report given to Mr. Wilson. This significant omission, whether intentional or inadvertent, resulted in a trial by ambush . . . .”) (citations omitted).
112 See SCHECK ET AL., *supra* note 14, at 125 (“Fish’s misleading testimony in the Willis case, which led to the conviction of an innocent man and allowed a predator to continue roaming the streets, shows why the state should have turned over all of Fish’s laboratory notes and data, rather than merely presenting her final report.”).
113 See FORENSIC ANALYSIS: WEIGHING BULLET LEAD EVIDENCE 110–11 (2004) (“The conclusions in laboratory reports should be expanded to include the limitations of compositional analysis of bullet lead evidence. . . . Moreover, a section of the laboratory report translating the technical conclusions into language that a jury could understand would greatly facilitate the proper use of this evidence in the criminal justice system.”).
114 See Brandon L. Garrett & Peter J. Neufeld, *Invalid Forensic Science Testimony and Wrongful Convictions*, 95 VA. L. REV. 1 (2009). The study identified several different types of invalid testimony: (1) presenting non-probative evidence as probative, (2) discounting exculpatory evidence, (3) using inaccurate frequencies or statistics, (4) providing a statistic without support, (5) providing non-numerical statements without empirical support, and (6) attributing the source of evidence to the defendant. *Id.* at 16–20.
Sixth, examiners should be prohibited from testifying beyond the laboratory report (unless a supplemental report is issued), a requirement that would protect against overreaching by prosecutors\textsuperscript{115} and preclude the opportunity for improper embellishments.

Finally, an enforceable code of ethics should be adopted.\textsuperscript{116}

Enforcement of these procedures can be effectuated through accreditation. For example, the American Society of Crime Lab Directors/Laboratory Accreditation Board requires quality assurance programs—i.e., proficiency testing, technical reviews, audits, and corrective action procedures.\textsuperscript{117} The NAS Report recommends mandatory accreditation of laboratories and the certification of examiners.\textsuperscript{118}

Legal procedures such as full pretrial discovery\textsuperscript{119} and the availability of defense experts also are important protections.\textsuperscript{120} Not only do they serve due process norms, they also are quality control mechanisms. Laboratory personnel should understand that the required documentation generated by the examination will be turned over to the defense and may be reviewed by defense experts.

CONCLUSION

Law enforcement influence over laboratory decisions is a serious problem. In an ideal world, independent crime laboratories would be the solution. Crime laboratories, however, have historically developed within police agencies, and

\textsuperscript{115} See supra notes 76–83 and accompanying text.

\textsuperscript{116} See NAS REPORT, supra note 1, at 26 (Recommendation 9 urges the establishment of “a national code of ethics for all forensic science disciplines and encourage[s] individual societies to incorporate this national code as part of their professional code of ethics.”).


\textsuperscript{118} NAS REPORT, supra note 1, at 25 (Recommendation 7: “Laboratory accreditation and individual certification of forensic science professionals should be mandatory, and all forensic science professionals should have access to a certification process.”).


\textsuperscript{120} The minority report of the Illinois Capital Punishment Commission believed that instead of an independent lab, a better solution “would be provided by state funding for the creation of a permanent cadre of forensic experts available to defense attorneys for consultation and review of forensic and scientific evidence.” CAPITAL PUNISHMENT COMM., supra note 15, at 54. “Such a group of permanently retained experts would provide a ready and consistent resource for information and assistance to defense attorneys (both privately retained and publicly appointed) about complicated areas of science that are not usually taught in law schools or easily understood.” Id.; see generally Paul C. Giannelli, Ake v. Oklahoma: The Right to Expert Assistance in a Post-Daubert, Post-DNA World, 89 CORNELL L. REV. 1305 (2004) (discussing the need to bolster the right to defense experts).
decades of entrenchment make it difficult to remove laboratories completely from law enforcement control.121

This does not mean, of course, that the status quo should be preserved. If located within law enforcement agencies, forensic laboratories should be as autonomous as possible and should be run in accordance with scientific norms, including procedures to protect against all types of bias. The NAS Report was not the last messenger. Within months of the report’s release, the Supreme Court wrote that “[f]orensic evidence is not uniquely immune from the risk of manipulation.”122

121 See Risinger et al., supra note 25, at 43 (“The establishment of freestanding government forensic laboratories, though occasionally advocated, would require such a revolution in thinking and organization, and diminish so many established bureaucratic empires, that it would take a generation of patient lobbying to have a chance of success.”) (citation omitted).

SWIRLS AND WHORLS: LITIGATING POST-CONVICTION CLAIMS OF FINGERPRINT MISIDENTIFICATION AFTER THE NAS REPORT

Jacqueline McMurtrie*

I. INTRODUCTION

The National Research Council of the National Academies’ recent report, *Strengthening Forensic Science in the United States: A Path Forward* ("NAS Report"), was heralded as “‘a blockbuster that will completely change the legal landscape regarding forensic evidence.’”¹ As the NAS Report notes, “[t]he number of exonerations resulting from the analysis of DNA has grown across the country in recent years, uncovering a disturbing number of wrongful convictions—some for capital crimes—and exposing serious limitations in some of the forensic science approaches commonly used in the United States.”² Those of us in the legal community representing individuals claiming they are innocent and that their convictions were based on flawed forensic science awaited the report with great anticipation.

It is undisputed that people are convicted for crimes they did not commit. To date, there have been 255 post-conviction DNA exonerations in the United States.³ A study of the first 200 exonerations identified 113 cases (57 percent) where forensic evidence was presented against the defendant during the original trial.⁴ As DNA technology has continued to improve, individuals convicted on the basis of other types of expert forensic testimony—including comparisons of bite marks, hairs, voiceprints, earprints, and fingerprints—were freed when post-conviction DNA tests proved the earlier forensic identifications wrong.⁵ However, DNA

---

* © 2010 Jacqueline McMurtrie, Associate Professor of Law; Director, Innocence Project Northwest Clinic—University of Washington School of Law; J.D., University of Michigan Law School; B.G.S., University of Michigan. The author wishes to thank Simon A. Cole and Keith A. Findley for their helpful comments, Kelly Canary and John Pantazis for their research assistance, and Tesia Stanley and other members of the *Utah Law Review* for convening the symposium.


⁵ Jacqueline McMurtrie, *The Role of the Social Sciences in Preventing Wrongful Convictions,* 42 Am. Crim. L. Rev. 1271, 1272–73 & nn.4–8 (2005) (referencing post-conviction DNA exonerations in which bitemark, hair, voiceprint, earprint, and fingerprint comparisons were presented at trial).
testing cannot provide a remedy for all wrongful convictions because in the vast majority of cases the perpetrator does not leave biological material at the crime scene and, therefore, there is no evidence to test.\footnote{Garrett, \textit{supra} note 4, at 116 (citing Protecting the Innocent: Proposals to Reform the Death Penalty; Hearing Before the S. Comm. on the Judiciary, 107th Cong. 221 (2002) (statement of Professor Barry Scheck, Co-Director of the Innocence Project) (“The vast majority (probably 80%) of felony cases do not involve biological evidence that can be subjected to DNA testing.”)); Nina Martin, \textit{Innocence Lost}, S.F. MAG., Nov. 2004, at 78, 105 (noting that “only about 10 percent of criminal cases have any biological evidence—blood, semen, skin—to test”).}

Hence, non-DNA forensic evidence will continue to play a critical role in the criminal justice system, and the work of forensic science practitioners is “wide-reaching and important.”\footnote{NAS REPORT, \textit{supra} note 2, at xix.} This Article’s focus on fingerprint evidence may at first glance appear to be an inquiry in search of a problem. It is true that only one of the 255 post-conviction DNA exonerations involved an erroneous fingerprint identification.\footnote{Simon A. Cole, \textit{The Prevalence and Potential Causes of Wrongful Conviction by Fingerprint Evidence}, 37 \textit{Golden Gate U. L. Rev.} 39, 41 (2006) (identifying Stephan Cowans as the “first—and thus far the only—person to be exonerated by DNA evidence for a wrongful conviction in which fingerprint evidence was a contributing factor”).} However in 2004, the highly publicized Brandon Mayfield case (discussed in Part IV \textit{infra}) brought national attention to the question of whether latent fingerprint identifications are reliable. Other documented cases of fingerprint misattributions are less renowned, but exist.\footnote{Simon A. Cole, \textit{More than Zero: Accounting for Error in Latent Fingerprint Identification}, 95 J. CRIM. L. & CRIMINOLOGY 985, 991 (2005) (documenting twenty-two cases of fingerprint misattribution that have been reported in the public record and concluding that the cases show “most likely only the tip of the proverbial iceberg of actual cases of fingerprint misattribution”).} In the end, it is impossible to assess the prevalence of error of latent fingerprint identifications. First, no records document how many criminal prosecutions in federal and state courts in the United States are based totally or partially on fingerprint evidence.\footnote{Lyn Haber & Ralph Norman Haber, \textit{Scientific Validation of Fingerprint Evidence Under Daubert}, 7 \textit{Law, Probability & Risk} 87, 87 (2008); see also Cole, \textit{supra} note 9, at 1017 (“Although there is no information on how many times latent print identification has been used in crime investigation, the number is clearly large.”).} Second, fingerprint misattributions go largely unnoticed, as there is “[n]o mechanism for recording, compiling, reviewing, or analyzing [the] cases.”\footnote{Cole, \textit{supra} note 9, at 997.}  

My own interest in the question of whether latent fingerprint evidence is prone to error is more than academic. I direct the Innocence Project Northwest (IPNW) Clinic at the University of Washington School of Law. The IPNW Clinic represents indigent people in Washington State who are serving long prison terms, who proclaim their innocence, and who no longer have a right to court-appointed counsel. Although the IPNW Clinic is particularly active in pursuing DNA testing on physical evidence to demonstrate innocence, we also represent clients in non-
DNA cases. We are currently investigating two cases where latent fingerprint “matches” were the only evidence linking the defendant to the crime. Because neither case is resolved, I will provide only minimal detail about the cases. In the first case, the defense presented expert witnesses to refute the latent print identification of the government’s expert witnesses. In the second, the defense did not offer an expert to rebut the government’s expert witness. In the first case, the government’s expert testified that he had found nine “points of comparison” (see Part II, infra) between the latent print and the defendant’s fingerprint. In the second case, the government’s expert testified that she used the Analysis, Comparison, Evaluation, and Verification (ACE-V) (see Part III, infra) method of comparison to match the latent print to the defendant’s fingerprint. Although both cases have the potential for post-conviction DNA testing, the chance of achieving conclusive results is stronger in one than the other. And so, I read the NAS Report with these cases in mind to assess what impact, if any, the report would have on post-conviction proceedings involving latent fingerprint identifications.

Part II of this Article provides a brief outline of latent fingerprint evidence as it is currently presented in courts. Numerous practical, scientific and legal articles, and books have been written on the subject of fingerprints and this Article cannot do justice to the vast body of literature. However, it will introduce three tenets at the heart of latent fingerprint evidence: (1) “uniqueness,” i.e., that every person possesses a set of unique and permanent fingerprints; (2) “individualization,” i.e., the determination that a latent print can be matched to its source to the exclusion of all others; and (3) “infallibility,” i.e., that when the latent print comparison is properly conducted by a trained examiner, its error rate is zero.

Part III, in reverse chronological order, provides a condensed history of latent print individualization, describing its rapid and largely unquestioned acceptance in the courts as forensic identification evidence and its ascendancy to the “gold standard” of identification. Like other types of “trace evidence” left at a crime scene, fingerprint evidence gained court acceptance prior to its being validated through scientific research.12

Part IV briefly explores the legal challenges to latent print individualization, which only began in the last decade of the twentieth century. The challenges were based upon scientific studies and scholarly research that questioned the validity and reliability of latent fingerprint individualization. Early on, the debate split into two factions, with latent print examiners and courts on one side, and legal and scientific scholars on the other, resulting in the nearly uniform admissibility of latent fingerprint comparison evidence.

12 Professors Wells and Loftus have argued that eyewitness identification evidence is another type of “trace” evidence that was accepted by the courts prior to its scientific validation. See Gary L. Wells & Elizabeth F. Loftus, Eyewitness Memory for People and Events, in 11 HANDBOOK OF PSYCHOLOGY 149, 149–50 (Alan M. Goldstein & Irving B. Weiner eds., 2003) (suggesting that one of the reasons the criminal justice system has failed to adopt a scientific model for eyewitness evidence is because eyewitness testimony was commonly used in criminal investigations long before any scientific studies of eyewitnesses had been conducted).
Part V examines the NAS Report’s discussion of latent fingerprint identification (or “friction ridge analysis”) and its conclusions regarding latent fingerprint examiners’ claims of (1) uniqueness, (2) individualization, and (3) infallibility. It discusses the reaction of the latent fingerprint community to the report, as well as trial court decisions issued after the NAS Report was published.

Part VI sets forth the procedural requirements of bringing post-conviction motions based upon a claim of newly discovered evidence. It identifies the obstacles petitioners face in bringing such claims. However, it also discusses other areas, notably Comparative Bullet Lead Analysis and Shaken Baby Syndrome, where petitioners have successfully raised post-conviction claims based upon new developments in forensic science.

Part VII returns to the question of whether a petitioner who claims to have been wrongly convicted on the basis of latent fingerprint evidence can obtain relief from the courts based upon the findings of the NAS Report. Part VII concludes that petitioners will continue to face substantial hurdles despite the questions raised by the NAS Report regarding data collection, the ACE-V method of interpretation, and issues surrounding the conclusions in reporting results of fingerprint comparisons. The primary challenge will be to educate courts that the NAS Report findings are substantive evidence that challenge the underlying principles that have long been the claim of fingerprint identifications: uniqueness, individualization and infallibility.

II. LATENT FINGERPRINT EVIDENCE

The first premise of fingerprint identification is that of “uniqueness,” i.e., every individual possesses a unique and permanent set of fingerprints. Fingerprint identifications in criminal cases are generally made from traces of fingerprint fragments detected at crime scenes. These fragments are commonly referred to as “latent fingerprints,” the word “latent” meaning hidden because the print is not readily visible. Every ridge of the fingers, palms, and soles bears sweat pores, which in the average person exude perspiration. Additionally, if the ridges come into contact with other parts of the body, such as hair or the face, or other objects, a film of grease or moisture may be deposited on the ridges. When fingers or palms touch an object, the film, moisture, or grease may be transferred to the object, leaving an outline of the ridges. Various powders and chemicals are used to develop the print to make it visible so that it may be preserved and compared. After a latent fragment is detected at a crime scene, it is then

15 Id.
16 Id.
17 Id. at 170, 173–86.
compared by a fingerprint examiner with inked or digitally scanned fingerprints taken directly from a suspect’s fingers.

Prior to the introduction of computer technology, the fingerprint identification process was conducted by clerks and fingerprint technicians who sifted through thousands of cataloged paper fingerprint cards searching for a match.\(^{18}\) In the late 1970s and early 1980s, law enforcement agencies in the United States began adopting Automated Fingerprint Identification Systems (AFIS) to improve efficiency and to reduce the amount of time it took to identify (or not exclude) a given individual from a latent fingerprint. Fingerprint examiners use computer workstations to mark the features of a scanned fingerprint image, encode the resulting data in a machine-readable format, and then search for similar fingerprints in an associated database of known fingerprints and records. The searches conducted by AFIS are rapid and allow examiners to search across the large pool of candidates in the database.\(^{19}\) Latent print searches are inherently more difficult than ten-print (or criminal identification) searches because the examiner may only have a partial print from the crime scene, the prints are regularly of poor quality, and often the examiner does not know from which finger a given latent print came.\(^{20}\)

When an optical image of a latent print is entered into the database, AFIS will retrieve optical images of candidate matches and display them in descending order of likelihood of matching.\(^{21}\) It is then up to the human examiner to compare the prints generated by AFIS to determine whether there is a match. Although “AFIS systems are very good at quickly winnowing an enormous database into a small group of candidate matches,” the systems are “relatively poor at selecting which, if any, of this small group is the actual match.”\(^{22}\)

It is then up to the fingerprint examiner to compare the latent fingerprint, or partial print, taken from the crime scene to the known exemplars generated by AFIS.\(^{23}\) Fingerprint examiners base their analyses upon the ridge outlines that appear on the inside of the end joints of the fingers and thumbs.\(^{24}\) The ridges have been categorized into several different general groups of ridge patterns: arches, loops, and whorls, which are further divided into multiple subgroups.\(^{25}\)

\(^{18}\) NAS REPORT, supra note 2, at 269.

\(^{19}\) Id.

\(^{20}\) Id. at 270.

\(^{21}\) COLE, supra note 13, at 254–55.

\(^{22}\) Id. at 255 (noting that in this particular context, “computers are fast but dumb”).

\(^{23}\) Id.

\(^{24}\) SCIENCE OF FINGERPRINTS, supra note 14, at iv.

\(^{25}\) Id. at 5. However, “there is no standard agreement among fingerprint examiners as to either the precise number or nomenclature of the different characteristics.” Robert Epstein, Fingerprints Meet Daubert: The Myth of Fingerprint “Science” is Revealed, 75 S. CAL. L. REV. 605, 608 (2002) (citing JAMES F. COWGER, FRICTION RIDGE SKIN: COMPARISON AND IDENTIFICATION OF FINGERPRINTS 143 (1983) (“The terms used to define and describe these characteristics vary markedly among writers in the field and differ even among examiners depending upon the organization in which they were trained.”); see also
The current technique used to examine and compare latent fingerprints to known print exemplars is described by the acronym ACE-V: Analysis, Comparison, Evaluation, and Verification, which, as its name implies, consists of four different stages.26

First, the examiner conducts an analysis of the latent print, which is most often viewed as a digital image.27 The examiner looks at several different factors28 to determine whether the latent print contains sufficient quantity and quality of detail to continue the analysis.29 If the detail is insufficient, the prints are called “of no value” or “not suitable” for comparison.30 If the examiner decides that there is sufficient detail, “the comparison of the latent print to the known print begins.”31

Second, the visual comparison of the prints consists of “visually ‘measuring,’ and comparing—within the comparable areas of the latent print and the known prints—the details that correspond.”32 The details could include:

[T]he overall shape of the latent print, anatomical aspects, ridge flows, ridge counts, shape of the core, delta location and shape, lengths of the ridges, minutia location and type, thickness of the ridges and furrows, shapes of the ridges, pore position, crease patterns and shapes, scar shapes, and temporary feature shapes . . . .33

Third, after the examiner completes the comparison, “the examiner performs an evaluation of the agreement of the friction ridge formations in the two prints and evaluates the sufficiency of the detail present to establish an identification (source determination).”34

Fourth, verification, the final step, occurs when a “qualified examiner repeats the observations and comes to the same conclusion, although the second examiner may be aware of the conclusion of the first.”35

Fingerprint examiners, according to the standards promulgated by the Scientific Working Group on Friction Ridge Analysis, Science and Technology (SWGFAST) are only allowed to report one of three acceptable conclusions

John Berry & David A. Stoney, History and Development of Fingerprinting, in ADVANCES IN FINGERPRINT TECHNOLOGY 2–3 (Henry C. Lee & Robert E. Gaensslen eds., 2d ed. 2001) (illustrating seven basic ridge characteristics, while acknowledging that “[s]ome authorities consider that only two types of characteristics are present”).

26 NAS REPORT, supra note 2, at 137 (internal citations omitted).
27 Id.
28 Id. at 137–38 (listing factors).
29 Id.
30 Id at 138.
31 Id.
32 Id.
33 Id.
34 Id.
35 Id.
resulting from their analysis: “individualization (or identification), exclusion, or inconclusive.”

Thus, the second premise of fingerprint identifications is one of “individualization.” When fingerprint examiners conclude that there is a ‘match’ between the latent print and the suspect’s print, it is expressed in absolute terms as an “individualization,” meaning that “[t]he determination that corresponding areas of friction ridge impressions originated from the same source to the exclusion of all others (identification).” The fingerprint literature suggests that examiners testify as follows:

Q: How sure are you that those two prints were made by the same finger?
A: Absolutely sure! I don’t testify to probabilities.

This concept of “individualization” is most commonly understood to mean the narrowing of possible sources of a forensic specimen to a particular source, “to the exclusion of all other possible sources.” Fingerprint examiners, along with those who analyze crime scene evidence, including shoe and tire impressions, toolmarks, firearms, and handwriting, have “individualization” of specific types of evidence as their goal. “Individualization” is a much stronger claim of source attribution than “identification” or “classification,” where the potential source of the specimen is narrowed to a group (or “class”) of similar items. Examples of the latter include hair comparisons, which match a hair to a particular ethnic group, or analysis of paint marks that identifies a class of vehicle.

And finally, the third premise of fingerprint identifications is one of “infallibility.” Many in the latent fingerprint community also testify that the ACE-V comparison method has a “zero error rate.” They claim that when the method is used by well-trained and experienced examiners, no errors are ever made, so that the method itself is error free. Thus, the claim is that erroneous identifications are only made by poorly trained or inexperienced practitioners. In other words, the

---

39 NAS REPORT, supra note 2, at 43–44.
40 Id. at 43.
41 Id. at 117–18.
42 Id.
“methodological” (sometimes called “scientific”) error rate is zero while the “practitioner” (sometimes called “human”) error rate is unknown.44

Backed by the claims of uniqueness, individualization, and infallibility, it is no wonder that “[f]ingerprint evidence is so powerful that erroneous fingerprint evidence is likely to convict, convict securely, and never be exposed.”45

III. LATENT FINGERPRINT’S RAPID ACCEPTANCE AS IDENTIFICATION EVIDENCE BY THE COURTS

The claim that latent fingerprint examiners’ conclusions of “individualization” are infallible is remarkable for many reasons; beginning with the fact that fingerprint evidence was initially viewed as inferior to the Bertillon system of anthropomorphic identification, which measured the size and proportions of the human body.46 Cloaked with the aura of sophistication that only a French nomenclature can impart, Bertillonage was used by criminal identification bureaus for the first decade of the twentieth century.47 Early fingerprint examiners did not view fingerprinting as forensic evidence of “individualization” that could link criminals to evidence left at a crime scene.48 Instead, fingerprinting served the record-keeping purpose of criminal identification, to link persons in custody to their criminal records.49 Latent fingerprint individualization, which came about as a fringe benefit of criminal identification, requires endorsing the principle that fingerprint examiners can accurately match crime-scene fingerprints, even if they are partial prints, to one and only one source finger.50 This principle came to be accepted despite the fact that latent prints, unlike prints taken from suspects using ink or scanners, which are generally of good quality, are “typically partial, smudged, or otherwise distorted.”51

The earliest published decision in the United States addressing forensic evidence of fingerprint “individualization” is the 1911 case of Thomas Jennings.52

44 See Cole, supra note 9, at 1034–43 (describing the latent fingerprint community’s “parsing of errors” to support its claim that the method of fingerprint comparisons is free from error).

45 Id. at 1021 (citing Tamara F. Lawson, Can Fingerprints Lie?: Re-Weighing Fingerprint Evidence in Criminal Jury Trials, 31 AM. J. CRIM. L. 1, 3 (2004) (“From my practical experience and scholarly research of the topic, the reliability of fingerprint identification evidence routinely goes unquestioned at all levels of the criminal process and by both sides of the litigation, prosecution, and defense.”)).

46 Cole, supra note 13, at 150–52.

47 See id. at 152.

48 Id. at 168.

49 Id.

50 Id.

51 Cole, supra note 9, at 991.

52 People v. Jennings, 96 N.E. 1077, 1081 (III. 1911); see also Cole, supra note 13, at 159–160.
Jennings was tried for the murder of Clarence Hiller during a home invasion burglary. There was some circumstantial evidence linking him to the crime. Jennings was found by the police a few hours after the murder with a freshly fired revolver containing cartridges that were claimed to match those found near the dead body. He was also identified at trial by eyewitnesses as the intruder in several other home invasions in the neighborhood on the same evening as the murder. But the strongest evidence against Jennings came from four government witnesses who testified that fingerprints left in wet paint at the crime scene belonged to Thomas Jennings. The witnesses testified that they had, in the course of their work with various Bureaus of Identification, examined thousands of fingerprints, and Jennings was convicted. On appeal, Jennings’s argument that the prints were not properly admitted was rejected. The court, citing such authorities as the Encyclopedia Britannica and a treatise on handwriting identification, emphasized that “standard authorities on scientific subjects discuss the use of fingerprints as a system of identification, concluding that experience has shown it to be reliable.” Thus, the court concluded “that there is a scientific basis for the system of fingerprint identification . . . and that this method of identification is in such general and common use that the courts cannot refuse to take judicial cognizance of it.”

Fingerprint evidence, once it was admitted, was quickly embraced by the judicial system. Courts engaged in little substantive analysis when admitting the testimony of fingerprint matches, instead relying upon Jennings and other precedent to admit the evidence. The first legal challenges to fingerprint evidence focused upon the question of whether uncorroborated fingerprint evidence was sufficient to support a conviction. Other challenges raised claims that fingerprint

53 Id. at 1079. Although too late for Mr. Jennings, a recent study on ballistics concludes: “The validity of the fundamental assumptions of uniqueness and reproducibility of firearms-related toolmarks has not yet been fully demonstrated.” NATIONAL RESEARCH COUNCIL, BALLISTICS IMAGING 3 (Daniel L. Cork et al. eds., 2008).
54 Jennings, 96 N.E. at 1079–1080. None of the witnesses saw the intruder’s face in the dark and “could only describe his build and say he was ‘a colored man.’” COLE, supra note 13, at 178. Nonetheless, they all testified that Jennings looked like the man who broke into their houses on the night of the murder. Id. We now know that mistaken eyewitness identification is the leading cause of wrongful convictions. Garrett, supra note 4, at 78 (noting that a study of the first 200 DNA exonerations found that 158 (79 percent) of the convictions involved mistaken eyewitness identifications). Furthermore, cross-racial identification is particularly prone to error. Id. at 79.
55 Jennings, 96 N.E. at 1080–82.
56 Id. at 1082.
57 Id. at 1081.
58 Id. at 1082.
60 See, e.g., State v. Minton, 46 S.E. 2d 296, 298 (N.C. 1948) (“The fact that fingerprints corresponding to those of an accused are found in a place where a crime was
Evidence was obtained in violation of a person’s constitutional rights. In response, courts ruled that fingerprinting an arrested person prior to conviction does not violate that individual’s right to due process or the privilege against self-incrimination.\(^{61}\) Nor does it represent a critical state of the proceedings requiring the presence of counsel.\(^{62}\) However, the use of fingerprints taken from a person after an illegal arrest violates that person’s right against unlawful search and seizure.\(^{63}\)

Scholars have addressed the question of why fingerprinting was not subjected to more scrutiny when it was first offered as a new system of identification. Professor Michael Saks posits that fingerprint identifications, like many other forms of early forensic evidence, were not critically examined because courts primarily relied upon an informal and implicit “marketplace test” to determine the admissibility of expert testimony.\(^{64}\) Under this test, individuals who had succeeded in making a living from marketing their services were presumed to have sufficient expertise to testify in court.\(^{65}\) As Saks argues, courts failed to recognize that fingerprint evidence needed to be subjected to more serious scrutiny because the latent fingerprint discipline had no external constituency that would provide the safeguards imposed by the commercial market.\(^{66}\) In addition, Saks notes that the judicial habit of relying upon precedent created a snowball effect: once a number of courts admitted fingerprint evidence, later courts merely followed their lead rather than conducting their own investigation of the validity and reliability of the evidence.\(^{67}\) Professor Jennifer Mnookin has offered three additional explanations for why courts did not rigorously scrutinize the admissibility of fingerprint evidence.\(^{68}\) First, the claim that every person’s fingerprint is distinctive, just as every snowflake is unique, has “inherent cultural plausibility.”\(^{69}\) Second, fingerprinting was easily assimilated as legal evidence because it is a visible means of identification, allowing jurors to examine the fingerprints offered into evidence themselves, and to even compare them against their own prints.\(^{70}\) Third, a fingerprint examiner’s strong claim of “certain, incontestable knowledge” made fingerprinting especially appealing to prosecutors and judges.\(^{71}\)


\(^{62}\) \textit{E.g.}, Pearson v. United States, 389 F.2d 684, 686 (5th Cir. 1968).

\(^{63}\) \textit{E.g.}, Bynum v. United States, 262 F.2d 465, 466–67 (D.C. Cir. 1958).


\(^{65}\) \textit{Id.} at 1074.

\(^{66}\) \textit{Id.} at 1104.

\(^{67}\) \textit{Id.} at 1105.

\(^{68}\) Mnookin, \textit{supra} note 59, at 32–36.

\(^{69}\) \textit{Id.} at 32–33.

\(^{70}\) \textit{Id.} at 33–35.

\(^{71}\) \textit{Id.} at 36.
describes, courts were searching for “the light of scientific truth” to provide scientific certainty and objectivity to authoritative judgments.\(^72\) Fingerprint examiners rarely disagreed with each other and the lack of the “battles of the experts” apparent in other disciplines gave fingerprinting an air of objective authoritative judgment.\(^73\)

One of the reasons that fingerprint examiners rarely disagreed with each other’s conclusions, Professor Cole asserts, is that they “articulat[ed] norms of method and conduct that would preclude disagreement between experts.”\(^74\) Prior to the ACE-V method of comparison described in Part II, examiners were able to uniformly reach similar conclusions through what became known as the “point counting” method of matching corresponding ridge characteristics between the latent and inked prints.\(^75\) Even though fingerprint examiners could in theory disagree about whether a corresponding point matched, and occasionally they did, the point counting method generated relatively little disagreement between examiners about their ultimate conclusions regarding identification. If an examiner was able to generate a comparison with a large number of matching points, there was very little chance that another examiner would disagree with the identification.\(^76\)

However, one of the problems with “point counting” is that there were no standard professional norms as to how many matching points were necessary to conclude there was a match between a latent print and a known suspect’s exemplar. Historically, the presence of twelve or more points of identity between the latent print and the known print was deemed sufficient for a positive identification, so long as there were no dissimilarities noted between the prints.\(^77\) And yet, many authorities believed that eight point comparisons were sufficient for the purposes of individualization,\(^78\) so long as the examiner paid attention to uniqueness, relative position, as well as to the number of identifying points.\(^79\) Thus, latent fingerprint evidence was held admissible in cases where the examiner found five points of similarity\(^80\) and even as few as four,\(^81\) with the courts holding

---

\(^72\) Id. (quotation omitted).

\(^73\) Id. at 38–39.

\(^74\) COLE, supra note 13, at 201.

\(^75\) In the United States, “point counting” was eventually replaced by the ACE-V method of comparison. See Lisa J. Steele, The Defense Challenge to Fingerprints, 40 CRIM. L. BULL. 213, 221–22 (2004).

\(^76\) COLE, supra note 13, at 201.


\(^78\) Id.


that the minimal points of comparison went to the weight to be accorded to the examiner’s testimony, and not to its admissibility.82

Latent fingerprint evidence came to be seen as so venerated by the courts that one described the evidence in deific terms: “[L]atent fingerprints are a corporeal signature of their maker, capable of conclusively identifying the individual who impressed them.”83

IV. PRETRIAL CHALLENGES TO FINGERPRINT EVIDENCE

Questions about the scientific validity and admissibility of fingerprint evidence only began to arise in the last decade of the twentieth century. According to Professor Mnookin, early critics focused on three main arguments: First, fingerprint examiners lacked objective and proven standards for evaluating what supported a conclusion of “individualization”; second, the error rate for fingerprinting as a technique had not been adequately studied; and third, there was no “statistical foundation for assessing whether two people might have the same prints with any given number of corresponding characteristics.”84 Professor Cole has argued that the focus on uniqueness should be abandoned as irrelevant to the question of individualization and that it cannot, in and of itself, support individualization.85 Instead he urges a focus on diagnosticity, or “the ability to assign traces of these objects to their correct source with a certain degree of specificity under certain parameters of detection and under certain rules governing such assignments.”86

Several factors combined to bring about a change in the way latent fingerprint evidence was viewed outside the latent fingerprint community. During the twenty-first century, legal and scientific articles have criticized the lack of validation of latent fingerprint individualization.87 Studies have been published documenting

82 Id.
83 People v. Clark, 214 P.3d 531, 536 (Colo. App. 2009), cert. granted, 2009 WL 2489064 (Aug. 17, 2009) (No. 09SC358) (examining “[w]hether the Court of Appeals erred in finding that crime scene DNA evidence, without further significant corroborative proof, is sufficient to identify [the defendant] beyond a reasonable doubt as the perpetrator in this case and therefore sustain his sexual assault conviction”).
84 See Mnookin, supra note 59, at 57–61 (discussing the three arguments in more detail).
86 Id. at 246.
that the error rate of fingerprint comparisons was “more than zero.” 88 In addition, the effect of cognitive biases upon forensic sciences in general, and latent fingerprint identification in particular, began to emerge through the publication of scholarly articles and studies. 89 The studies revealed that the fingerprint examiner community was not immune from the common cognitive bias that results in a tendency for conclusions to be affected by the way that information is presented to an individual. As an example, a study asked experienced fingerprint examiners to analyze fingerprints that, unbeknownst to them, they had analyzed in previous years. 90 Half of the examiners were given contextually biasing information before the examinations. If the examiners had previously declared an exclusion, they were told the “suspect confessed to the crime,” but in cases where examiners had previously declared a match, they were told the “suspect was in police custody at the time of the crime.” 91 In four of the twenty-four examinations where examiners were provided with the biasing information prior to the examinations, they reached conclusions different from the ones they had previously reached. 92

One of the most highly publicized cases of erroneous fingerprint identification, which also involved potential contextual bias, occurred when Brandon Mayfield was arrested for his purported involvement in the Madrid train


91 Id. at 608.

92 Id. at 610.
bombing in 2004. 93 Mayfield was identified as the source of a latent print left at the crime scene by the FBI, and the “match” was confirmed by an independent defense fingerprint examiner.94 Mayfield served time in custody before Spanish authorities cleared him and matched the fingerprints to an Algerian national.95 As part of the aftermath generated by the misidentification, an FBI examination of the case concluded that once the first FBI fingerprint examiner reached a conclusion that the latent print matched Mayfield, the result was shared with other examiners, who were influenced by their supervisor’s conclusion and by the inherent pressures of such a high-profile case.96

The Mayfield case was subsequently used in a study that reaffirmed the effect contextual bias can have upon the results fingerprint examiners reach in their cases. After the Mayfield misidentification was revealed, researchers gave five expert latent fingerprint examiners a pair of prints, which, although unrevealed, they had previously identified as a match five years earlier in the normal course of their work.97 Prior to examining the prints, the examiners were given the contextually biasing information that the set of prints was the one the FBI had erroneously matched to Mayfield in the Madrid bombing case.98 The examiners “were asked to decide whether there was sufficient information” to make a decision and, if so, to assess whether there was a match or exclusion.99 Only one participant made a decision consistent with the one made five years earlier by declaring a match between the prints.100 The other four participants changed their identification decision; three determined that there was an exclusion, and one decided that there was insufficient information to declare a match or an exclusion.101

Because of the growing body of scientific and scholarly discourse questioning the scientific validity of the basic principles underlying fingerprint comparisons, the 1993 Supreme Court decision in Daubert v. Merrell Dow Pharmaceuticals, Inc.,102 was seen as opening a door to challenging the admissibility of the evidence, which many thought was closed under the previous Frye v. United States103

93 NAS REPORT, supra note 2, at 123.
94 Cole, supra note 9, at 985–86.
95 Id. at 986; Sara Kershaw, Spain and U.S. at Odds on Mistaken Terror Arrest, N.Y. TIMES, June 5, 2004, at A1.
98 Id. at 76.
99 Id.
100 Id.
101 Id.
103 293 F. 1013 (D.C. Cir. 1923).
standard. The admissibility of fingerprint evidence in the *Jennings* case was shortly followed by the 1923 decision of *Frye v. United States*, which became the leading standard for the admissibility of scientific evidence. In *Frye*, the defendant sought to admit the results of a favorable polygraph examination to prove his innocence. The court rejected the evidence stating:

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

For 70 years *Frye*’s “general acceptance” test “[was] the dominant standard for determining the admissibility of novel scientific evidence at trial.”

The “general acceptance” standard of *Frye* was generally assumed to preclude arguments against the admission of latent print comparisons because the “relevant scientific community” of latent print examiners “accepts” latent print evidence. Moreover, *Frye* specifies that its standard applies only to novel scientific evidence. Thus, forms of expert evidence which either: “(1) pre-date *Frye* altogether (as in the case of latent print individualization evidence), or (2) post-date *Frye*, but are not challenged until after they have become familiar enough to the criminal justice system to no longer be regarded as ‘novel,’ would not be challengeable under *Frye*.” For these reasons, the first case in which a court explicitly considered the admissibility of latent print individualization evidence under *Frye* did not occur until 2007. In that year, a Maryland state trial court ruled in *State v. Rose* that latent print evidence in a capital murder trial was inadmissible under the state’s *Frye* standard. After reconsideration of the state

104 *Id.*
105 *Id.* at 1013.
106 *Id.* at 1014.
107 *Daubert*, 509 U.S. at 585.
108 *See, e.g.*, United States v. Gary, 85 F. App’x 908, 909 (4th Cir. 2004) (“[F]ingerprint analysis is one of those forms of evidence where the reliability of the science and its general acceptance is apparent without a full reexamination of the science.”).
109 *See Frye*, 293 F. at 1014.
110 Cole, *supra* note 87, at 464–65. Cole posits that arguments to exclude evidence of latent fingerprint individualization should be reconsidered under the *Frye* standard of admissibility given its lack of general acceptance among the relevant scientific community of scholars, scientists and practitioners. *Id.* at 466.
111 *Id.* at 530.
court ruling was denied in 2008, the United States Attorney’s Office for the District of Maryland filed an indictment in federal court.\(^\text{113}\) The federal trial court, applying the \textit{Daubert} factors (discussed below), then ruled that the fingerprint identification evidence was admissible.\(^\text{114}\)

In 1993, in \textit{Daubert v. Merrell Dow Pharmaceuticals, Inc.}, the Supreme Court clarified that Federal Rule of Evidence 702,\(^\text{115}\) not \textit{Frye}, governed the admissibility of expert evidence in federal court.\(^\text{116}\) As a result, it identified factors that a trial court could consider when deciding whether to admit expert testimony based upon scientific evidence: (1) whether the evidence “can be (and has been) tested” using the scientific method; (2) whether it has “been subjected to peer review and publication”; (3) the “known or potential rate of error” of the technique in question; (4) the “existence and maintenance of standards controlling the technique’s operation”; and (5) the “general acceptance” of the technique within the relevant scientific community.\(^\text{117}\) However, the Supreme Court specified that the inquiry under Rule 702 was a “flexible one.”\(^\text{118}\) In a later decision, \textit{General Electric Co. v. Joiner}, the Court made clear that a trial court’s decision was subject to review under the narrow “abuse-of-discretion” standard of review.\(^\text{119}\) And finally, in \textit{Kumho Tire Co. v. Carmichael}, the Court stated that \textit{Daubert}’s factors might be applicable in a trial court’s assessment of non-scientific expert testimony.\(^\text{120}\) \textit{Kumho} went on to state that “whether \textit{Daubert}’s specific factors are, or are not, reasonable measures of reliability in a particular case is a matter that the law grants the trial judge broad latitude to determine.”\(^\text{121}\)

Initially, \textit{Daubert} was viewed by many legal scholars to have “opened a door to reconsideration of the admissibility of latent print evidence, a door that had been closed under \textit{Frye},” because latent print individualization was seen to have trouble meeting the first four criteria of \textit{Daubert}.\(^\text{122}\) However, many scholars initially conceded that latent fingerprint evidence probably satisfied the fifth \textit{Daubert}

\(^{114}\) \textit{Id.} at 725–26.
\(^{115}\) FED. R. EVID. 702, as amended after the \textit{Daubert} decision, provides:

\begin{quote}
If scientific, technical, or other specialized knowledge will assist the trier of fact to understand the evidence or to determine a fact in issue, a witness qualified as an expert by knowledge, skill, experience, training, or education, may testify thereto in the form of an opinion or otherwise, if (1) the testimony is based upon sufficient facts or data, (2) the testimony is the product of reliable principles and methods, and (3) the witness has applied the principles and methods reliably to the facts of the case.
\end{quote}

\(^{117}\) \textit{Id.} at 593–94.
\(^{118}\) \textit{Id.} at 594.
\(^{120}\) Kumho Tire Co. v. Carmichael, 526 U.S. 137, 150 (1999).
\(^{121}\) \textit{Id.} at 153.
\(^{122}\) Cole, \textit{supra} note 87, at 462–63.
criterion, which echoed the Frye “general acceptance” criterion. Nonetheless, Professor Saks reasoned that under Daubert “[a] vote to admit fingerprints is a rejection of conventional science as the criterion for admission. A vote for science is a vote to exclude fingerprint expert opinions.”

The first challenge to latent print evidence was brought in federal court, where Daubert applied, in the 1999 case of United States v. Mitchell. Two latent fingerprints were offered against Mitchell, who was charged with armed robbery. Ten witnesses—seven for the government and three for the defense—spent five days in court debating whether latent fingerprint evidence was scientific. The Mitchell case division of experts foreshadowed the polarization between practitioners and scientists that would occur in future court challenges. Three of the government witnesses were latent print examiners without any advanced scientific background, while the defense witnesses were scientists and scholars who had conducted research on the validity of latent fingerprint evidence. At the Daubert hearing, the government contended 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, that the error rate of the technique could be meaningfully parsed into “methodological” and “human” categories, and that the methodological error rate was zero. Mitchell’s defense experts conceded “uniqueness” and argued that one could not infer the accuracy of the techniques from the uniqueness of its target of analysis, or from casework or trials, and that the error rate of fingerprint individualizations, although unknown, was certainly not zero. Ultimately, the court, although not ruling on the question of whether fingerprint evidence was scientific (finding the question irrelevant because a witness’s expertise could be based upon technical, rather than scientific, knowledge), accepted the prosecution’s position that fingerprint evidence is reliable.

Subsequent challenges to latent print evidence under Daubert have not been successful; nearly every state and federal court has held that latent print evidence is

---

123 Id. at 463.
124 Saks, supra note 64, at 1106.
126 COLE, supra note 13, at 284.
127 Id. at 284–85.
128 See Cole, supra note 87, at 502–04. Indeed, Cole demonstrates that “only with difficulty can the government produce any non-practitioner scientists who accept the claim that latent print individualization is valid. And yet, paradoxically, few courts have acknowledged this deficiency, and those that have have tended to deem it irrelevant to admissibility.” Id. at 540–41.
130 See id. at 239–41.
131 See id.
133 COLE, supra note 13, at 285.
admissible, that it is generally accepted by the scientific community, and that ACE-V meets the requirements of Daubert.\textsuperscript{134} Courts have opined that the reliability of latent fingerprint identification was established by its century of testing in the courts\textsuperscript{135} and that its error rate is essentially zero.\textsuperscript{136}

V. THE NAS REPORT’S FINDINGS ON LATENT PRINT EXAMINATION

It was against this backdrop of scientific and legal debate regarding latent fingerprint comparisons that the NAS Report was released. As a general matter, the report endorses the scientific method, concluding that every forensic discipline must be “founded on a reliable scientific methodology that gives it the capacity to accurately analyze evidence and report findings.”\textsuperscript{137} It rejects the idea that adversarial testing can serve as a substitute for the scientific method, instead finding that the process was “not suited to the task of finding ‘scientific truth.’”\textsuperscript{138}

More specifically, the report examines the three premises underlying latent fingerprint evidence: (1) uniqueness, (2) individualization, and (3) infallibility.

Although the NAS Report notes that there is some scientific evidence supporting the presumption of uniqueness, it rejects the claim that reliability of latent print individualization can be inferred from the uniqueness of friction ridge skin.\textsuperscript{139} It concludes that nuclear DNA testing is the only forensic method that can support a finding of individualization.\textsuperscript{140}

Notably, the report rejects the claim made by many in the latent fingerprint community regarding infallibility, or the “zero-error rate,” of a properly conducted test administered by a trained examiner. Instead, the report finds that because there is “limited information about the accuracy and reliability of friction ridge analyses, claims that these analyses have zero-error rates are not scientifically plausible.”\textsuperscript{141}

The report repeatedly asserts that latent print comparisons are dependent upon the quality of the latent print, concluding that “not all fingerprint evidence is equally good, because the true value of the evidence is determined by the quality of the latent fingerprint image.”\textsuperscript{142} Furthermore, the NAS Report expresses concerns about the effect cognitive bias can have on conclusions reached by forensic

\begin{footnotesize}
\begin{itemize}
\item \textsuperscript{135} United States v. Havvard, 117 F. Supp. 2d 848, 854 (S.D. Ind. 2000).
\item \textsuperscript{136} See id. at 854–55; see also United States v. Crisp, 324 F.3d 261, 268–69 (4th Cir. 2003).
\item \textsuperscript{137} NAS REPORT, supra note 2, at 9.
\item \textsuperscript{138} Id. at 12 (“The adversarial process relating to the admission and exclusion of scientific evidence is not suited to the task of finding ‘scientific truth.’”).
\item \textsuperscript{139} Id. at 43–44, 144.
\item \textsuperscript{140} Id. at 87.
\item \textsuperscript{141} Id. at 142.
\item \textsuperscript{142} Id. at 7–8, 86–87, 140.
\end{itemize}
\end{footnotesize}
examiners. It singled out the Mayfield case as an example of a situation in which fingerprint examiners were influenced in their evaluations by the knowledge that an FBI examiner had declared a match.

The NAS Report discusses data collection, the ACE-V method of interpretation, and issues surrounding the conclusions in the reporting of results. Regarding the ACE-V method, the report states:

ACE-V . . . is not specific enough to qualify as a validated method for this type of analysis. ACE-V does not guard against bias; is too broad to ensure repeatability and transparency; and does not guarantee that two analysts following it will obtain the same results. For these reasons, merely following the steps of ACE-V does not imply that one is proceeding in a scientific manner or producing reliable results.

Despite its critique and rejection of the principles of “individualization” and “infallibility,” the NAS Report acknowledges the value of fingerprints as a method of forensic identification and calls for greater research to implement improved standards and training. It urges the latent print community to provide sufficient documentation of the comparison of latent prints as well as the basis for conclusions to promote transparency in the identification process. It supports further research to study ridge flow and crease patterns of the hands and feet, as well as to assess the rarity and discriminating value of the patterns of ridge formations. Finally, it states that there is room for improvement in conducting research upon the multiple factors that affect the quality of latent print images, to prevent examiners from “explain[ing] a ‘difference’ as an ‘acceptable distortion’ in order to make an identification.”

The reaction to the NAS Report, reminiscent of a “déjà vu all over again” experience, has echoed the polarized debate that preceded the report. The International Association for Identification (IAI), the largest professional organization of latent print examiners in the world, issued a response to the NAS Report that states: “[U]nique anatomical features . . . have become the foundation upon which the individualization of a fingerprint to a single person becomes scientifically accepted and legally defensible.” SWGFAST also has disagreed

---

143 Id. at 122–25.
144 Id. at 123.
145 Id. at 142.
146 Id. at 142–44.
147 Id.
148 Id. at 144.
149 Id. at 145 (citing U.S. Dep’t of Justice, Office of the Inspector General, A Review of the FBI’s Handling of the Brandon Mayfield Case, 2006).
with the report’s findings, 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.” 151 Thus, the organizations continue to espouse the view, repudiated by the NAS Report, that claims of individualization are justified by assertions of uniqueness.

Other supporters of latent fingerprint individualization went back to the notion, also rejected by the NAS Report, that reliability of the technique could be inferred from adversarial testing in the courts:

Despite over 100 years of court acceptance worldwide, this most venerable of the forensic methods, which at one time was called “the gold standard” against which all forensic techniques ought to be compared, came in for some very serious criticism, at least when it came to comparing unknown crime scene impressions to known prints. 152

The authors go on to state: “The [NAS] report is affirmative of the technique and makes suggestions were (sic) additional research or the imposition of stricter standards—both of which have been advocated repeatedly by leaders in the profession—can strengthen forensic science in the United States.” 153

Ultimately, the question of whether the NAS Report will fulfill its promise as a legal “blockbuster” will be decided by the courts. In United States v. Rose, the federal trial court considered the findings of the NAS Report in its Daubert analysis and nonetheless concluded:

[F]ingerprint identification evidence based on the ACE-V methodology is generally accepted in the relevant scientific community, has a very low incidence of erroneous misidentifications, and is sufficiently reliable to be admissible under Fed Rule Ev. 702 generally and specifically in this case. 154

Shortly before the publication of this Article, a trial court in California denied a defendant’s motion to exclude fingerprint evidence under the state’s equivalent of the Frye test. 155 In State v. Greenwood, the court ruled that:

Fingerprint comparisons, which “jurors essentially can see for themselves” . . . do not qualify for scrutiny under Kelly. Rather, the Defendant ought to address his concerns using the standard framework

153 Id. § 10.08.
for expert opinion testimony. The [latent print examiner] will not be permitted to testify that her opinion is the result of an infallible scientific process, and the Defendant is free to vigorously cross-examine the LPE on the shortcomings of the ACE-V method raised in the 2009 National Academy of Science Report entitled *Strengthening Forensic Science in the United States: A Path Forward*. The fingerprint identification evidence is the quintessence of expert testimony, and the Defendant’s numerous critiques of the ACE-V process go to the weight of the People’s evidence, not its admissibility.\(^{156}\)

At the very least, the NAS Report can serve to educate the judge and jury about the troubling questions raised by the preeminent committee that was tasked by Congress to study the forensic disciplines, including friction ridge analysis. If a court is not willing to exclude evidence that a latent print found at the crime scene matches a suspect, the report can be used (as it was in *Greenwood*) to refute or preclude claims of infallibility made by latent print examiners.

**VI. SCIENTIFIC DEVELOPMENTS AS A BASIS FOR NEWLY DISCOVERED EVIDENCE CLAIMS**

The initial reaction of trial courts to the question of whether the NAS Report supports pretrial exclusion of latent fingerprint comparisons does not bode well for petitioners challenging convictions based upon fingerprint “matches.” Petitioners raising a claim post-conviction must overcome substantial hurdles to be successful. The next Part discusses obstacles petitioners face when raising claims of newly discovered evidence and highlights areas in which petitioners have successfully raised such claims based upon new developments in science. The Article then returns to the question of whether petitioners who claim to have been wrongly convicted on the basis of latent fingerprint evidence can obtain relief from the courts based upon the findings of the NAS Report.

Every state currently permits a motion for a new trial on the basis of newly discovered evidence.\(^{157}\) Newly discovered evidence of innocence “represents a ground for relief through the principal state post-conviction remedies in thirty-two states.”\(^{158}\) However, petitioners who raise newly discovered evidence claims in state post-conviction collateral proceedings face difficult procedural and substantive obstacles. First, many jurisdictions impose a relatively brief statute of

\(^{156}\) *Id.* (citation omitted).


\(^{158}\) *Id.* at 682 n.192 (discussing states that allow a free-standing innocence claim to be raised in habeas proceedings versus those that do not allow innocence claims to be raised and instead limit habeas proceedings to issues involving violation of constitutional rights or lack of jurisdiction).
limitations upon filing a motion for a new trial on the grounds of newly discovered evidence. 159 Second, the legal standard in most states for granting a new trial requires a petitioner to prove that the “newly discovered evidence is noncumulative, does not simply impeach the prosecution’s witnesses, and would probably yield a different result in a new trial.” 160 Third, if a trial judge rejects a new trial motion based on newly discovered evidence, “the standard of [appellate] review applied to that denial is extraordinarily deferential—[in most jurisdictions a petitioner] must prove the trial court abused its discretion in rendering its decision or failed to exercise that discretion.” 161

There are areas in which developments in the field of forensic science have not only altered the way in which cases are prosecuted, but also have offered avenues of relief for inmates erroneously convicted as a result of those findings. For example, post-conviction DNA testing using technology that was not available at the time of trial has been the basis for reversing a conviction and ordering a new trial. 162 However, DNA testing is not completely analogous to other methods of forensic analysis because the fundamental principles of the science supporting DNA testing have not been altered or discarded. However, recent developments in the scientific principles once used to support prosecutions based upon Comparative Bullet Lead Analysis (CBLA) and Shaken Baby Syndrome (SBS) are informative in demonstrating how new developments in forensic science have resulted in post-conviction relief for individuals convicted at trial on the basis of CBLA “matches” or evidence of SBS.

A. Comparative Bullet Lead Analysis (CBLA)

Bullet lead analysis, sometimes called compositional bullet lead comparison or comparative bullet lead analysis (CBLA), allows forensic investigators to identify the elemental composition and characteristics of bullets or bullet fragments found at a crime scene against unused bullets found in the possession of a suspect. 163 It was used for more than thirty-five years by expert witnesses,
primarily from the FBI, who testified that they were able to tell “where a bullet or bullet fragment was manufactured and even from which box it originated—all based on an analysis of the composition of the lead.”\(^{164}\) However, the FBI has recently rejected the basic tenets underlying CBLA, resulting in the reversal of convictions based solely on the now-discredited forensic science.

CBLA analysis is based on three premises:

The first premise . . . [is that] a bullet or bullet fragment is assumed to be compositionally representative of a molten source of lead. . . . The second premise is that the molten source from which the sample lead originated is compositionally uniform or homogeneous. . . . And the third, and most significant premise is the claim that each molten source of lead is unique in composition, establishing a chemical and metallurgical “fingerprint.”\(^{165}\)

Using these premises, the examiner would begin by taking three samples from each bullet or bullet fragment and conduct a quantitative and qualitative analysis of the concentrations of seven selected elements in the bullet lead alloy of both the crime-scene and the suspect’s bullets.\(^{166}\) The examiner would then apply a statistical analysis to determine whether the samples were “analytically indistinguishable.”\(^{167}\) If so, the examiner would provide expert testimony that the “crime scene bullet(s) came from the same manufacturer, molten source, batch, or box as the bullet(s) traceable to the suspect.”\(^{168}\)

CBLA expert witness evidence was routinely and unquestionably accepted by the courts for thirty-five years.\(^{169}\) The premises underlying the technique were not seriously challenged until early this decade, when William Tobin, the former chief metallurgist for the FBI laboratory, began questioning the technique in court testimony and scientific and legal articles.\(^{170}\) In 2002, Tobin and Wayne Duerfeldt published an article calling into question the use of CBLA in criminal trials.\(^{171}\) Tobin and Duerfeldt concluded that the metallurgical composition of bullets could either be fairly uniform throughout the melt, dramatically varied, or, most importantly, two different melts could be “compositionally indistinguishable.”\(^{172}\) They found that most bullet manufacturers obtained their lead from secondary

\(^{164}\) Id.
\(^{165}\) Id.
\(^{166}\) See id. at 29, 31.
\(^{167}\) See id. at 29.
\(^{169}\) Id. at 45.
\(^{171}\) Tobin & Duerfeldt, supra note 163, at 27.
\(^{172}\) See id. at 28.
refiners and that most secondary refiners obtained their scrap lead from old automobile batteries. Because automobile batteries have tight specifications due to “electrical conductivity, corrosion, processing, and other considerations,” Tobin and Duerfeldt found that bullets made from two different refiners could be “analytically indistinguishable” from one another. Their conclusion “that compositional uniqueness as an underlying premise or universal statement for the practice of CBLA is not scientifically valid.”

As a result of this research, the FBI asked the National Research Council (NRC) of the National Academy of Sciences to convene an independent committee of experts to evaluate the scientific basis of CBLA. The ensuing report—Forensic Analysis: Weighing Bullet Lead Analysis—concluded that an FBI examiner’s testimony “matching” a fragment of a bullet found at a crime scene to bullets found in the possession of a defendant was so overstated that such evidence was “seriously misleading under Federal Rule of Evidence 403.” Specifically, the report found that the methods of statistical analysis used during the second step of the CBLA analysis were not the best available. With regard to testimony concerning the statistical analysis, the report recommended that FBI expert witnesses limit their interpretations of findings to refer to “compositionally indistinguishable volumes of lead” instead of melts or boxes when discussing origin, and acknowledge uncertainties in the CBLA statistical analysis. Although the report thoroughly discussed the general limitations of CBLA, it ultimately concluded that CBLA was a “reasonably accurate way of determining whether two bullets could have come from the same compositionally indistinguishable volume of lead” and can “in appropriate cases provide additional evidence that ties a suspect to a crime.”

However, in 2005, the FBI announced that after extensive study and consideration, it would no longer conduct examinations of bullet lead. The Bureau’s decision to discontinue CBLA was based primarily on the “inability of scientists and manufacturers to definitively evaluate the significance of an association between bullets made in the course of a bullet lead examination.”

In 2007, the FBI took the unprecedented step of reviewing cases to determine

---

173 Id. at 27.  
174 Id. at 28–31.  
175 Id. at 28.  
178 Id. at 26–70.  
179 Id. at 112.  
180 Id. at 110.  
182 Id.
whether its CBLA testimony may have contributed to a wrongful conviction. In such cases, the FBI issued letters identifying its error and seeking to remedy erroneous convictions based upon CBLA.

The “newly discovered evidence” undermining the validity of CBLA “individualization” was used to overturn a defendant’s conviction during post-conviction proceedings in New Jersey. In State v. Behn, the New Jersey Court of Appeals reversed the conviction of Michael Behn after considering studies that proved the CBLA testimony offered against Behn was unsupported by science. At trial, an FBI examiner testified that “the lead fragments recovered from the decedent’s body and the defendant’s bullets came from the same source of lead, and both the fragments recovered from the decedent’s body and the defendant’s bullets came from the same box or boxes and were packaged on the same date by the manufacturer.” The court held that the newly discovered evidence, consisting of studies on composition bullet lead analysis that had not yet been developed and completed prior to the defendant’s murder trial, would have “effectively neutralized” the testimony of the state’s expert and was of such caliber that “it possessed, ‘to a probability—not a certainty,’ the capacity to change the jury’s verdict.” Thus, the court ruled that the defendant was entitled to new trial based on the newly discovered evidence.

**B. Shaken Baby Syndrome (SBS)**

Shaken Baby Syndrome (SBS) originated in the 1970s as a hypothesis to explain respiratory distress or death in a small group of infants from apparent head injury with an identifiable impact site and three diagnostic symptoms that came to be known as the SBS “triad.” Scholars have noted that the presence of the triad—retinal hemorrhaging, subdural hematoma, and cerebral edema—“was taken to mean that a baby had been shaken hard enough to produce what were conceptualized as whiplash forces.” The theory was that the disproportionately large heads and weak necks of infants allowed their heads to course back and forth during shaking, causing subdural hemorrhage through rupture of the bridging veins.

---

183 *See id.*
184 *Id.*
186 *Id.* at 335.
187 *Id.* at 345 (internal citation omitted).
188 *Id.* at 346.
between the brain and the dura. If the “triad” was present, it was understood by experts to be “pathognomic—or exclusively characteristic—of SBS.”

SBS prosecutions became prevalent in the last decade of the twentieth century: fifteen hundred incidents of alleged SBS are reported nationwide every year. In her seminal article on the topic, Professor Turkheimer notes that the science surrounding SBS has undergone a striking transformation since the mid-1990s. Many of those in the scientific community who were early proponents of SBS’s validity have come to perceive the diagnosis as illegitimate. New research conducted by neurosurgeons, biomechanical engineers, and pathologists have “eroded confidence in the existence of a pathognomonic relationship between shaking and the SBS triad.” Akin to some of the debates that are taking place in the field of latent print examinations, others in the medical field have responded to the new research with a passionate defense of SBS, making the debate highly polarized. And yet, Turkheimer concludes that there is enough common ground between the two factions that each would agree that the scientific underpinnings of SBS have evolved.

Turkheimer identifies three areas in which, from a perspective of “pure” science, the thinking on SBS has shifted, and where consensus has been achieved among supporters and opponents of SBS. She calls the first area “The Myth of Pathognomony,” explaining “[doctors] have conceded that the triad is not necessarily induced by shaking, and that a differential diagnosis must be considered.” The second area is called “Lucid Intervals.” In the past, doctors testified with certainty that there could be no “periods of lucidity” between the abuse and the loss of consciousness. This necessarily meant that the caretaker who was with the infant immediately prior to the loss of consciousness was identified and prosecuted as the person who perpetrated the abuse. Recent studies have shown, however, that children may in fact be lucid for more than seventy-two hours between the time of the fatal injury and its resulting death. These findings have dramatically altered the forensic landscape because the

---

192 Turkheimer, supra note 189, at 4.
194 Id.
195 Id. at 14.
196 Id. at 16.
197 Id.
198 See id. at 17–18.
199 Id. at 18.
200 Id.
201 Id.
202 Id.
203 Id.
identity of the perpetrator (if there is a crime) cannot be established by showing who was with the infant before he or she became unconscious. Turkheimer calls the third area of consensus “Removing the Shaking from the Syndrome.” Biomechanical studies have established that shaking alone cannot create the forces necessary to cause subdural hemorrhage without damaging the infant’s neck and cervical spinal cord or column. These types of injuries are not present in most infants with the “triad” of symptoms, and therefore they could not have been simply shaken to death. Thus, the possibility of “short-fall” incidents, although still controversial, must be considered as an explanation for the injuries, and other non-traumatic symptoms previously assumed to be pathognomonic of shaking must be considered and evaluated. Congenital malformations, metabolic disorder, hematological diseases, infectious diseases, and autoimmune conditions are types of differential diagnoses for symptoms that were previously exclusively associated with SBS.

The transformation in the medical understanding of SBS formed the basis for Audrey Edmunds’s request for a new trial after she was convicted for what was then determined to be an SBS death of an infant in her care. In State v. Edmonds, the Wisconsin Court of Appeals summarized the trial evidence as follows:

Audrey Edmunds was charged with first-degree reckless homicide following the death of seven month old Natalie on October 16, 1995, while Edmunds was caring for Natalie at Edmunds’s home. At trial, the State presented numerous medical expert witnesses who testified to a reasonable degree of medical certainty that the cause of Natalie’s death was violent shaking or violent shaking combined with impact that caused a fatal head injury. The State’s witnesses also testified that after being injured, Natalie would have had an immediate and obvious response and would not have appeared normal. Natalie’s mother, and the father of another child in Edmunds’s care who observed Natalie, testified that Natalie was acting normally when she was dropped off at Edmunds’s home on the morning of her death.

As Turkheimer explains,

In early 2007, the judge who presided over Audrey Edmunds’s trial over a decade earlier conducted a five-day evidentiary hearing upon her motion for a new trial based on newly discovered evidence. The defense

\[\text{Id. at 18–19.}\]
\[\text{Id. at 19.}\]
\[\text{Id. at 20.}\]
\[\text{Id. at 21–22.}\]
\[\text{Id. at 22.}\]
\[\text{State v. Edmunds, 746 N.W. 2d 590, 592 (Wis. Ct. App. 2008).}\]
\[\text{Id.}\]
experts testified that since the mid-1990s, “significant research has undermined the scientific foundations for SBS, creating substantial challenges to matters that were nearly universally accepted in the medical community at the time of Edmunds’s trial.”

After defense experts—who told the court that they would have sided with the prosecution at trial—testified that the evolving science had changed their opinions as to the likely cause of death, the trial judge denied the motion for a new trial. The judge acknowledged that “[s]tanding alone and unchallenged, the defense witnesses provide[d] a sufficient evidentiary basis to order a new trial based upon newly discovered medical evidence,” but found that “the newly discovered evidence . . . is significantly outweighed by the evidence presented by the prosecution.” However, an appellate court reversed this decision and concluded that there was a reasonable likelihood that a different result would be reached at a new trial, because of the “shift in mainstream medical opinion since the time of Edmunds’s trial.” Audrey Edmunds was released from prison after serving eleven years, having been convicted of a crime for which she had always maintained her innocence. The State ultimately elected not to retry her at the request of the deceased baby’s parents. She returned home to reunite with her own three daughters, who were 11, 13, and 16 when she was finally released.

VII. CONCLUSION

A petitioner seeking to challenge a conviction that was based upon a fingerprint match will face difficulties, even when armed with the findings of the NAS Report. In the areas described above, CBLA and SBS, there were “insiders” who began to challenge the science underlying the forensic conclusions reached by experts in the field. The community of latent print examiners has yet yielded a William Tobin, or multiple experts such as those in the Edmunds case who testified they had changed their opinion about the scientific validity of the conclusions they had once endorsed. The committee convened by Congress to study the forensic disciplines and issue the NAS Report was a highly credentialed and impressive assembly of scientists, academics, and judges. However, the

---

212 Turkheimer, supra note 189, at 48 (citing Brief of Defendant-Appellant at 11, Edmunds, 746 N.W. 2d 590 (No. 2007AP933)).
214 Id.
215 Edmunds, 746 N.W. 2d at 598–99.
217 See id.
committee appears to be viewed as a group of “outsiders” by the IAI and SWGFAST, who represent the “insider” community of latent fingerprint examiners. Both organizations have publicly rebutted the findings of the NAS Report. As a result, lines will continue to be drawn between practitioners and scientists. At this time, it is too early to determine which side the courts will ultimately align themselves with after the publication of the report.

As discussed above, petitioners raising post-conviction claims of newly discovered evidence face hurdles in addition to those confronted by a defendant during trial. The procedural requirements of a post-conviction newly discovered evidence claim universally require a petitioner to show that new evidence is “substantive” and not “merely impeaching.” If the findings of the NAS Report are viewed as being admissible merely for the purpose of cross-examination, a reviewing court will deny a petitioner’s claim for relief because cross-examination evidence is classic impeachment evidence. The challenge will be to educate courts that the NAS Report findings are substantive evidence challenging the underlying principles that have long been the claim of fingerprint identifications: “uniqueness,” “individualization,” and “infallibility.” Indeed, as the report indicates, one of the claims (infallibility) is not “scientifically plausible,” and the others have not been supported by reliable research.\(^219\) Finally, to succeed on a claim of newly discovered evidence, a petitioner must show that the evidence would “probably change the outcome of the trial.” This assessment is highly case-specific and will depend upon the nature of the expert testimony given and the importance of the fingerprint match in establishing the defendant’s guilt.

And so, petitioners challenging convictions based solely upon what they claim are erroneous fingerprint “matches” may find themselves where they were before the publication of the NAS Report. They can seek out defense experts to examine and critique the conclusions of the government’s expert witnesses, as did many of the individuals whose cases were discussed in Cole’s recent study of latent fingerprint misattributions.\(^220\) In the twenty-two cases he identified, ten were cases in which the individual was convicted on the basis of the fingerprint identification (the others were released prior to being charged or tried, or were acquitted).\(^221\) Three of the ten convictions were overturned after the FBI conducted a review of a state fingerprint laboratory, and another three were overturned based upon the re-examination by defense expert witnesses.\(^222\) However, a more promising avenue could be to seek experts outside the discipline to challenge the premises endorsed by fingerprint examiners at the petitioner’s trial.

Ultimately, the success of petitioners challenging convictions based upon what they claim are erroneous fingerprint “matches” may depend upon the new “gold standard” of forensic evidence—DNA typing. Certainly, cases like those of Stephan Cowans will bring into question the claimed infallibility of latent print evidence.

\(^{219}\) NAS REPORT, supra note 2, at 142–44.
\(^{220}\) See Cole, supra note 9, passim.
\(^{221}\) Id. at 1067–70 tbl.1.
\(^{222}\) Id.
individualization. The most promising way at present to obtain post-conviction relief may be to use emerging DNA technology to obtain profiles from epithelial cells or perspiration left on the latent print. These types of DNA exclusions will lend more unequivocal truth to the claim that “latent fingerprints are a corporeal signature of their maker, capable of conclusively identifying the individual who impressed them.”

223 See, e.g., Jennifer J. Raymond, Claude Roux and Simon J. Walsh, Fricton Ridge Skin-Interaction Between Fingerprint Detection and DNA/Biological Material, in WILEY ENCYCLOPEDIA OF FORENSIC SCIENCE [3] 1318 (A. Jaemison & Andre A. Moenssens eds. 2009); David E.O. Van Hoofstat et al., DNA Typing of Fingerprints and Skin Debris: Sensitivity of Capillary Electrophoresis in Forensic Applications Using Multiplex PCR, in PROCEEDINGS OF THE 2ND EUROPEAN SYMPOSIUM ON HUMAN IDENTIFICATION 131, 131–33 (Promega Corp. ed., 1998) (reporting on the successful use of DNA analyses on fingerprints and skin debris left on tools and clothing and bags); David E. O. Van Hoofstat, et al., DNA Typing of Fingerprints Using Capillary Electrophoresis: Effect of Dactyloscopic Powders, 20 ELECTROPHORESIS 2870, 2870–76 (1999) (examining the effect of different dactyloscopic powders on DNA typing performed on physical fingerprints left on glass and wooden plates and concluding that some powders may be used with DNA analyses so long as precautions are taken to avoid contamination).

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

Jane Campbell Moriarty*

1993

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

2009

With the exception of nuclear DNA analysis . . . no forensic method has
been rigorously shown to have the capacity to consistently, and with a
high degree of certainty . . . demonstrate a connection between evidence
and a specific individual or source.2

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

I. INTRODUCTION

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

---

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


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

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

4 Id. at 109 (quoting Peter J. Neufeld, The (Near) Irrelevance of Daubert to Criminal
Justice and Some Suggestions for Reform, 95 AM. J. PUB. HEALTH 107, 109 (2005)).
handwriting, bullet and bullet cartridge, hair, dental imprint, footprint, tire track, or even a lip print to its unique source (collectively, “individualization evidence”). Not only can they match it, they claim, they can do so often without any error rate.

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

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

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

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

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

II. THE NAS REPORT

In February 2009, the National Academy of Sciences issued its report on the status of forensic science, Strengthening Forensic Science in the United States: A Path Forward.6 It is a detailed discussion, with each chapter providing a
compendium of a separate subject relevant to forensic science. The NAS Report, highlighting the myriad shortcomings and failures of what we call forensic science, seems quite shocked at the current situation, remarking that “[t]he Law’s greatest dilemma in its heavy reliance on forensic evidence, however, concerns the question of whether—and to what extent—there is any science in any given ‘forensic science’ discipline.”

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

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

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

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

---

7 Id. at 87.
8 Id. at 109 (stating that the courts have been ineffective in addressing the issue of the validity of forensic science).
9 See id. at 1–33 (explaining the shortcomings of forensic science and providing a summary of sought-for changes in the system).
10 Id. at 4.
11 Id.
12 Id.
13 Id.
14 Id. at 5–6.
15 Id. at 6.
A wide variability across disciplines with respect to techniques, methodologies, reliability, types and numbers of potential errors, research, general acceptability, and published material, and

Problems relating to interpretation of forensic science—the questionable conclusion about individualization where a specimen is “matched” to a particular individual or source.

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

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

---

16 Id. at 6–7.
17 Id. at 7–8.
18 Id. at 12.
19 See id.
20 Id.
21 Id.
22 Id.
23 See id. at 87 (“The degree of science in a forensic science method may have an important bearing on the reliability of forensic evidence in criminal cases.”). For a searching empirical analysis explaining the relationship between forensic science errors and wrongful convictions, see Brandon L. Garrett & Peter J. Neufeld, *Invalid Forensic Science Testimony and Wrongful Convictions*, 95 VA. L. REV. 1 (2009).
24 The foundations of arson investigation, erroneously believed by many to rest on firm scientific footing, suffers from multiple shortcomings. The NAS Report discusses the “paucity of research” and states that “many of the rules of thumb that are typically assumed to indicate an accelerant was used . . . have shown not to be true.” NAS REPORT, supra note 2, at 173.
25 Id. at 136–64. See generally Jane Campbell Moriarty and Michael Saks, *Forensic Science: Grand Goals, Tragic Flaws, and Judicial Gatekeeping*, 44 JUDGES’ J. 16, 17 (2005) (explaining that “individualization specialties seek to associate an item of evidence
A conclusion of individualization implies that the evidence originated from that source, to the exclusion of all possible sources. The determination of uniqueness requires measurement of object attributes, data collected on the population frequency of variation in these attributes, testing of attribute independence, and calculation of the probability that different objects share a common set of observable attributes.26

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

Despite the need for such research to be done to validate methodology and underlying presumptions, the prosecution continues to argue such evidence is reliable, and, to date, many courts seem to be agreeing.51 When describing the lack of research or scientific scrutiny concerning individualization specialties, the report concludes that “although the precise error rates of these forensic tests are still unknown, comparison of their results with DNA testing in the same cases has revealed that some of these analyses, as currently performed, produce erroneous results.”32 These “erroneous results” caused wrongful convictions in some cases. Although people on both sides of the aisle may debate how important this lack of

---

26 Id.
27 Id. at 41 (stating DNA analysis, “with its well-defined precision and accuracy, has set the bar higher for other forensic science methodologies”).
28 Id. at 43–44.
29 Id. at 87.
30 Id. at 44.
31 See Part VI, infra, discussing post-NAS Report cases. Even since the report was issued, many federal prosecutors have continued to argue the evidence is completely reliable and the NAS Report should not be taken into consideration. See Harry T. Edwards, The National Academy of Sciences Report on Forensic Sciences: What it Means for the Bench and Bar, presentation at the Superior Court of the District of Columbia, May 6, 2010, at 4.
32 NAS REPORT, supra note 2, at 42.
research is, the inescapable conclusion is that without scientific proof of the foundations necessary to legitimately declare a match, much expert testimony is simply a hunch, supported by experiential “observations of countless samples.”

There are numerous reasons that judges admit such testimony so readily: they trust the FBI and other forensic scientists; they assume they would have heard on a more global level if forensic science were unreliable; they share an intuitive belief that forensic comparison is valid; they are comforted by other courts’ decisions admitting the evidence (what I term the “generally accepted by other courts” standard); they find the analysts’ experiential knowledge convincing; and they believe that the long history of use confers some field reliability to the work. Although many of these justifications might seem compelling on an intuitive, “hunch-like” level, further scrutiny reveals some serious cracks in the foundation of such opinions.

---


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

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

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

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

Proving the accuracy of Judge Posner’s comments, the Supreme Court has created, from virtual whole cloth, the “rules” governing admissibility of expert testimony, in the Daubert/Joiner/Kumho Tire trilogy (“trilogy”). To be admissible, the party seeking to introduce forensic evidence must be able to establish, to the satisfaction of the trial court, that the proposed expert evidence meets the Supreme Court-created standard of reliability, incorporated into the Federal Rules of Evidence in 2000. A substantial minority of courts use versions

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

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


41 See, e.g., FED. R. EVID. 702 (providing, in pertinent part, that an expert may testify if “(1) the testimony is based upon sufficient facts or data, (2) the testimony is the product
of the so-called “Frye general acceptance test,” while the federal courts and a majority of state jurisdictions use some variant of the Daubert reliability standard.\textsuperscript{42}

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

The trilogy governing the admission of expert testimony claims to envision a flexible standard, in which the trial court, as gatekeeper of the evidence, determines whether expert evidence meets a minimal standard of evidentiary reliability.\textsuperscript{44} Thus, in the case of expert evidence in the federal courts, the Supreme Court has created the rules by which the courts and litigants must abide\textsuperscript{45} and requires the inferior courts to serve as the arbiters of whether the evidence

\textsuperscript{42} The “Frye general acceptance test” requires proof that the novel scientific evidence is generally accepted in the field to which it belongs. The general acceptance test does not analyze the reliability of the proposed evidence; it asks whether novel scientific evidence has reached the tipping point at which it has become generally accepted by scientists in the field. See Frye v. United States, 293 F. 1013, 1014 (D.C. Cir. 1923). Many federal cases substantially incorporate some variant of the Frye test in their reliability determination, because it is one of the factors listed by the Daubert court. See 509 U.S. at 588–90. For a state-by-state breakdown of which states use the respective tests, see 2 Jane Campbell Moriarty, Psychological and Scientific Evidence in Criminal Trials app. 1A (2008) (providing a state-by-state analysis of admissibility standards) [hereinafter Moriarty, Criminal Trials].

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

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

\textsuperscript{45} See Daubert v. Merrell Dow Pharm., Inc., 43 F.3d 1311, 1316 (9th Cir. 1995) (On remand, Judge Kozinski, writing for the court, applied the standard created by the Supreme Court with the statement, “[m]indful of our position in the hierarchy of the federal judiciary, we take a deep breath and proceed with this heady task.”). Many states, through their supreme courts’ decisions, also have either adopted or rejected Daubert, in whole or in part. See Moriarty, Criminal Trials, supra note 42, at app. 1A. Thus, very clearly, courts both federal and state are engaged in rule-making in the area of expert testimony.
complies with those rules in this judicially-created gatekeeping role.\textsuperscript{46} Thus, the judiciary has created an entire universe where none previously existed and has appointed itself the master of such universe.

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

Nonetheless, after reading a multitude of cases involving forensic science evidence and empirically driven studies about what courts have done post-trilogy, it is not at all clear that judges “make sure everybody plays by the rules.” Indeed, research suggests that many judges do not require the prosecution to play by the same rules that other litigants in both civil and criminal cases do when the subject is forensic science evidence. These findings are not based solely upon my own impressions,\textsuperscript{47} but rather upon empirical data collected and analyzed by others\textsuperscript{48} and the conclusions of the NAS Report: “[T]he vast majority of the reported opinions in criminal cases indicate that trial judges rarely exclude or restrict expert testimony offered by prosecutors; most reported opinions also indicate that appellate courts routinely deny appeals contesting trial court decisions admitting forensic evidence against criminal defendants.”\textsuperscript{49}

It seems abundantly clear that many courts merely give lip service to pressuring the prosecution to meet its burden of proof in \textit{Daubert} hearings; much the way some bartenders mix an extra dry martini, as if waving the bottle of vermouth next to the glass will suffice. In other words, the court will find in favor of the prosecution if the prosecution meets its burden of proof. Thus, the “make sure everybody plays by the rules” standard is not met in many cases.

\textsuperscript{46} Daubert v. Merrell Dow Pharm., Inc., 509 U.S. 579, 589 n.7 (1993) (mentioning that Fed. R. Evid. 702 is the source of the judge’s gatekeeping duty).

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

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

\textsuperscript{49} NAS REPORT, supra note 2, at 11 (emphasis in original). Of course, the universe of reported decisions is far smaller than that of all decisions courts make, a point the NAS Report concedes. Id. Nonetheless, if the reported cases are at all representative, the discrepancy is noteworthy.
Deirdre Dwyer writes compellingly about the asymmetries in the application of Daubert between civil and criminal cases. She considers the normative expectations that either all evidence would be treated equally, or that evidence rules “in a criminal action [would be] geared more towards reducing the risk of an erroneous outcome than they are in a civil action.” After looking at data-driven studies, however, she concludes, “[i]t would seem that the expert evidence of civil plaintiffs, particularly in toxic tort cases, is subject to greater scrutiny than that of civil defendants, while the expert evidence of criminal prosecutors is subject to less scrutiny than that of criminal defendants, or than that of civil parties.” This seems to be an odd outcome, indeed, when civil cases involve only monetary damages, while criminal cases deal with stakes of much greater value.

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

---

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


52 Id. at 384–85.

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

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

55 The following cases provide a small sample of courts’ in-depth opinions on toxic tort expert evidence: Gen. Elec. Co. v. Joiner, 522 U.S. 136, 143–47 (1997) (analyzing admissibility of plaintiffs’ proposed expert testimony that exposure to PCBs promoted decedent’s lung cancer); McLain v. Metabolife Int’l, Inc., 401 F.3d 1233, 1239–52 (11th Cir. 2005) (reversing jury verdict for plaintiffs because plaintiffs’ expert evidence that ingestion of defendant product for weight loss was a likely cause of their respective strokes and heart attack was not reliable); Norris v. Baxter Healthcare Corp., 397 F.3d 878, 883–87 (10th Cir. 2005) (explaining why plaintiffs’ experts’ opinions that silicone breast implants caused autoimmune diseases was not reliably grounded on existing data); Moore v. Ashland Chem. Inc., 151 F.3d 269, 274–79 (5th Cir. 1998) (upholding trial court’s decision to exclude expert evidence on allegations that plaintiff’s exposure to toluene caused his respiratory disease); Daubert v. Merrell Dow Pharm., Inc., 43 F.3d 1311, 1316–22 (9th Cir. 1995) (on remand from the U.S. Supreme Court, analyzing whether plaintiffs’ claim that Bendectin exposure to fetuses in utero caused their limb malformation); Henricksen v. Conoco Phillips Co., 605 F. Supp. 2d 1142, 1152–79 (E.D. Wash. 2009) (analyzing expert testimony that plaintiff’s exposure to benzene caused the development of his myelogenous leukemia); Amorgianos v. Nat’l R.R. Passenger Corp., 137 F. Supp. 2d 147, 160–191, (E.D.N.Y. 2001), aff’d, 303 F.3d 256 (2d Cir. 2002) (explaining why experts’ opinions that
these cases, the district courts, affirmed by courts of appeal on an abuse of discretion standard, have found such testimony insufficiently reliable.\textsuperscript{56} Indeed, the NAS Report notes the different standards in civil and criminal cases, remarking that “courts have not . . . imposed standards ensuring the application of scientifically valid reasoning and reliable methodology in criminal cases involving \textit{Daubert} questions.”\textsuperscript{57} Furthermore, the report concludes that upon reviewing the reported decisions, “at least in criminal cases, forensic science evidence is not routinely scrutinized pursuant to the standard of reliability enunciated in \textit{Daubert}.”\textsuperscript{58}

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

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

\footnotesize
\textsuperscript{57} NAS REPORT, supra note 2, at 96.

\footnotesize
\textsuperscript{58} Id. at 106.

\footnotesize
\textsuperscript{59} See cases mentioned infra at notes 60–62.

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

\footnotesize
\textsuperscript{61} See, e.g., United States v. Glynn, 578 F. Supp. 2d 567, 574–75 (S.D.N.Y. 2008) (acknowledging that firearm toolmark comparison “lacks the rigor of science” and “suffers from greater uncertainty than many other kinds of forensic evidence,” yet allows testimony that a firearms match was “more likely than not”); United States v. Greene, 405 F. Supp. 2d 104, 124 (D. Mass. 2005) (allowing expert to testify about the points of similarity between a shell casing at the scene and one from defendant’s weapon but disallowing any testimony that the shell casings matched); United States v. Hines, 55 F. Supp. 2d 62, 70–71 (D. Mass. 1999) (limiting testimony of handwriting comparison witness to discussing points of
the reliability of those forms of evidence as well. Microscopic hair comparison, implicated for its role in wrongful convictions in many exonerations, has been admitted by both federal and state courts, with seemingly no concern for the potentially substantial rate of error such evidence presents.

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

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

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

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

---

68 Id. at 576.
69 Id.
70 Id. at 572.
71 United States v. Prime, 220 F. Supp. 2d 1203, 1210 (W.D. Wash. 2002), vacated and remanded on other grounds, 543 U.S. 1101 (2005), subsequently aff’d on this issue, 431 F.3d 1147 (9th Cir. 2005).
72 Prime, 431 F.3d at 1153.
73 Id. (noting that in one professor’s studies, experts arrived at the correct conclusions 87 percent of the time). As noted by the Llera Plaza court (and echoed by numerous scholars), the proficiency tests are not rigorous: “[T]he proficiency tests are less demanding than they should be. To the extent that this is the case, it would appear that the tests can be of little assistance in providing the test makers with a discriminating measure of the relative competence of the test takers.” 188 F. Supp. 2d at 565. For further information on proficiency tests of fingerprint examiners, see NAS REPORT, supra note 2, at 143–44.
lip print at the scene. Although there was absolutely no foundation for such evidence, the trial court, without the smallest amount of critical inquiry, found such proof met both Frye and reliability standards. The defense lawyer handling this case acted incompetently by not satisfactorily challenging the testimony, yet both the trial court and the court of appeals in the direct appeal were satisfied with little more than the ipse dixit testimony of the government’s witnesses.

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

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

---

75 See id. at 1258–59.
76 People v. Davis, No. 94 CF 76, 2006 WL 2641753 (Cir. Ct. Ill. Mar. 7, 2006) (holding that defendant was denied his constitutional right to effective counsel and overturning the defendant’s conviction).
77 See Davis, 710 N.E.2d at 1258–59.
78 On the other hand, this case may not be that much of an outlier because federal courts have permitted much expert testimony in criminal cases that appears to have dicey scientific grounding, including profile testimony about specific types of pedophiles. See, e.g., U.S. v. Hitt, 473 F.3d 146, 151, 158–59 (5th Cir. 2006) (upholding admission of expert testimony “to explain the behavior of those accused of sexual offenses”). Federal courts also have admitted allegedly scientific comparisons between photographs of items and the items themselves. See, e.g., U.S. v. McKreith, 140 F. App’x 112, 116 (11th Cir. 2005) (allowing expert testimony on the comparison of the defendant’s shirt that was seized in his apartment with the shirts depicted in bank surveillance images).
80 See Beecher-Monas, supra note 79, at 1378–84.
81 See id. at 1385–87.
Bite mark analysts often testify using phrases such as “consistent with” or “positively match.” Yet such statements have no grounding in science, because they are purely subjective conclusions unsupported by data. In addition, as with other forms of individualization forensic science evidence, statements about a match are neither scientifically supportable nor scientifically meaningful. Early research, however, appears to indicate that such statements are convincing to juries.

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

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

---

82 See, e.g., State v. Garrison, 585 P.2d 563, 566 (Ariz. 1978) (affirming trial court’s decision admitting expert testimony that the probability of two sets of teeth being identical was eight in one million, based on “articles written throughout the literature that do mention the possibility or the numerical values of finding two (sets of teeth) that match.”).
83 Beecher-Monas, supra note 79, at 1386.
85 See NAS REPORT, supra note 2, at 87 (“[N]o forensic method other than nuclear DNA analysis has been rigorously shown to have the capacity to consistently and with a high degree of certainty support conclusions about ‘individualization’ (more commonly known as ‘matching’ of an unknown item of evidence to a specific unknown source).”).
86 See NAS REPORT, supra note 2, at 21 (noting the need for standardized testimony).
87 See Dawn McQuiston-Surrett & Michael J. Saks, Communicating Opinion Evidence in the Forensic Identification Sciences: Accuracy and Impact, 59 HASTINGS L.J. 1159, 1188 (2008) (explaining studies conducted to determine how jurors analyze such terms as “match” and “consistent with” and finding that jurors overestimated the meaning of such terms); accord NAS REPORT, supra note 2, at 21 (“[U]se of such terms can and does have a profound effect on how the trier of fact . . . perceives and evaluates scientific evidence.”).
88 See NAS REPORT, supra note 2, at 176 (“[T]here is considerable dispute about the value and reliability of the collected data for interpretation.”).
89 See cases collected in Beecher-Monas, supra note 79, at 1372–74, and accompanying notes.
principles underlying fingerprint identification has, for decades, been misplaced.”

Thus, in a sweeping but unsupported statement, the court concluded that “the district court was well within its discretion in accepting at face value the consensus of the expert and judicial communities that the fingerprint identification technique is reliable.”

With regard to handwriting comparison, the court stated, “[t]he fact that handwriting comparison analysis has achieved widespread and lasting acceptance in the expert community gives us the assurance of reliability that Daubert requires.”

By not actually evaluating the testimony using Daubert’s suggested factors, the court evaded the problem that the evidence did not meet reliability requirements, a point noted by the dissent, which explained in detail how the government “utterly failed to meet its burden” of establishing reliability under Daubert.

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

Similarly, handwriting comparison fares even more poorly. Conclusions are entirely subjective, and the NAS Report recommends that its “scientific basis . . . needs to be strengthened.”

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

---

91 Id. at 268–69.
92 Id. at 269.
93 Id. at 271.
94 See Kumho Tire Co. v. Carmichael, 526 U.S. 137, 152 (1999) (“[A] trial court should consider the specific factors identified in Daubert where they are reasonable measures of the reliability of expert testimony.”).
95 Crisp, 324 F.3d at 273 (Michael, J., dissenting).
96 NAS REPORT, supra note 2, at 139 (noting that the threshold for making a source identification is “deliberately kept subjective”).
97 Id.
98 Id. at 143 n.31 and accompanying text.
99 Id. at 143–44.
100 Id. at 166.
101 NAS REPORT at 161.
simply unsupported by research, and serious questions are raised about its
evidentiary reliability.

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

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

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

102 Id. at 96–97.
103 Id. at 110.
104 Id.
105 Mnookin, supra note 35, at 141.
106 See cases cited supra note 55. Granted, part of the courts’ apparent comfort in the
civil cases may arise from the parties’ well-structured challenges and responses to the
expert testimony. Many criminal practitioners are not developing the challenges as well as
their civil counterparts. Nonetheless, serious, science-based challenges have been raised
against fingerprints, toolmarks, and handwriting comparison in many criminal cases, and
nearly universally courts have rejected them. See cases cited supra notes 60–64.
107 Daubert, 509 U.S. 579, 590 n.9 (describing the standard as “evidentiary
reliability—that is, trustworthiness”).
IV. “HISTORY MAY BE SERVITUDE”: THE ERRONEOUS JUDICIAL FOCUS ON THE LONG HISTORY OF USE OF INDIVIDUALIZATION EVIDENCE

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

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


110 Seigworth, supra note 109, at 204.

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

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

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

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

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

V. JUDICIAL DECISION MAKING AND THE PROBLEMS OF COGNITIVE BIAS

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

\(^{114}\) Hugh Bowden, Classical Athens and the Delphic Oracle, Divination and Democracy 19 (2005).

\(^{115}\) Id. at 29–30.

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

\(^{117}\) Bowden, supra note 114, at 19.


\(^{119}\) For more on this idea, see Chad M. Oldfather, Judges and Humans: Interdisciplinary Research and the Problems of Institutional Design, 36 Hofstra L. Rev. 125, 128 n.11 (2007) (collecting articles referencing the different ways that judges are subject to the same frailties as other humans). Jeff Rachlinski also reminded me that many
by heuristic decision making, friendships, beauty, the strength of a case, public opinion, fear of reversal, and the normal set of cognitive biases to which we all are subject: expectation bias, hindsight bias, confirmation bias, tunnel vision, and so forth.

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

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

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

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

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

121 Id. at 7.

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

123 Id. at 27.

124 Id. at 36.
could solve intuitively, the error rate was substantially higher than with a problem the judges had to actually reason their way through to resolve.\footnote{Id. at 10–13 (discussing the Cognitive Reflection Test Model from Shane Frederick, Cognitive Reflection and Decisionmaking, \textit{supra} note 122, at 26–28, and the judges’ performance on the test).}

Moreover, the problem of cognitive bias is most apparent, according to researchers, in the use of intuitive decision making, and it is where such problems as stereotyping, prejudice, and discrimination are likely to arise.\footnote{Thomas Gilovich and Dale Griffin, \textit{Introduction—Heuristics and Biases: Then and Now}, in \textit{JUDGMENT UNDER UNCERTAINTY}, \textit{supra} note 120, at 7.} Relying on heuristic shortcomings can lead to systematic biases.\footnote{See \textit{Benjamin R. Newell, et al., THE PSYCHOLOGY OF DECISION MAKING}, 71 (2007) (citing the oft-quoted Kahneman, D. \& Tversky, A., \textit{The Simulation Heuristic}, in \textit{Kahneman et al., supra} note 120; see also Burton, \textit{supra} note 36, at 149).}

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

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

Consider the diametric approaches courts employ when deciding admissibility of scientifically complex expert evidence (such as DNA comparison or toxic tort causation) versus the non-DNA individualization evidence such as fingerprints and handwriting. In the complex scientific evidence cases, courts appear to use a more rational processing system and engage in deliberate, analytic reasoning throughout the opinion to determine whether the evidence is reliable.\footnote{This does not mean that error cannot occur in the deliberative process or that judges always get the right answer. All it means, according to research psychology, is that the probability of error is reduced using this method.} A good example of this methodology is found in the Supreme Court’s opinion in \textit{General Electric v. Joiner}, in which the court engaged in a long methodological analysis of the quality of epidemiological evidence to determine whether plaintiff’s proof that PCB exposure had promoted his lung cancer was sufficiently reliable to be admitted.\footnote{Gen. Elec. Co. v. Joiner, 522 U.S. 136 (1997).} None of the writing appears intuitively-based.

By contrast, very few cases involving individualization evidence seem to be of the long, methodological analysis present in \textit{Joiner}. To the contrary, courts seem to rely on such heuristic devices as “long history of use” or “generally
accepted by other courts” to support their decisions, rather than engaging the evidence and subjecting it to a rational, science-based analysis.\textsuperscript{130}

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

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

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


\textsuperscript{133} Id. at 286.

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


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

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

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

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

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

---

136  Seigworth, supra note 109, at 2027.
137  In fact, in comparisons of clinical judgments and actuarial judgments (controlled studies), clinical judgments fare more poorly. One scholar notes that “[f]ailure to accept a large and consistent body of scientific evidence over unvalidated personal observation may be described as a normal human failing or, in the case of professionals who identify themselves as scientific, plainly irrational.” Robyn M. Dawes et al., Clinical Versus Actuarial Judgment, in KAHNEMAN ET AL., supra note 120, at 716, 727.
138  See Melendez-Dias v. Massachusetts, 129 S. Ct. 2527, 2537 n.6 (2009) (noting the report’s conclusion that the forensic science system is badly flawed and refuting the suggestion that forensic examiners are “uniquely reliable”).
analysts’ honesty, proficiency, and methodology—the features that are commonly the focus in the cross-examination of experts.\(^\text{141}\)

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

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

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

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

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

. . .

> The point here is simple: *When scientific methodologies once considered sacrosanct are modified or discredited, the judicial system must accommodate the changed scientific landscape.*\(^\text{146}\)

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

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

\(^{142}\) *Montalvo-Rangel*, 2010 WL 1484708, at *3.

\(^{143}\) *Rose*, 672 F. Supp. 2d at 725.

\(^{144}\) *Id.*

\(^{145}\) *Id.*

United States v. Taylor, a district court decision from New Mexico, provides a detailed explanation of the shortcomings of firearm toolmark comparison, discussing the lack of standards by which a match is declared, the subjective basis for the conclusion of a match, and the lack of standards for even distinguishing between class, subclass, and individual characteristics.¹⁴⁷ “[T]he . . . theory is circular. An examiner may make an identification when there is sufficient agreement, and sufficient agreement is defined as enough agreement for an identification.”¹⁴⁸ The court also quotes the damning language from the NAS Report that, even with better training and new techniques, “the decision of the toolmark examiner remains a subjective decision based on unarticulated standards and no statistical foundation for estimation of error rates.”¹⁴⁹ Nonetheless, the significance of those failures¹⁵⁰ is swept aside as orthogonal to the underlying validity. The court, citing pre-NAS Report cases, finds the practice of cartridge comparison sufficiently reliable to be admitted and permits the expert to give his opinion “within a reasonable degree of certainty in the firearms examination field.”¹⁵¹

United States v. Willock¹⁵² is a deeper, more thoughtful analysis. The court cites the conclusion expressed in the NRC Ballistic Imaging Report: “The validity of the fundamental assumptions of uniqueness and reproducibility of firearms-related toolmarks has not yet been fully demonstrated,” and “additional general research on the uniqueness and reproducibility of firearms-related toolmarks would have to be done if the basic premises of firearms identification are to be put on a more solid scientific footing.”¹⁵³ Yet, in its analysis, the court finds the pre-NAS Report decisions compelling, noting that “the furthest” any court to date has gone is to exclude testimony where examiners’ results were not confirmed or documented; or “to restrict the degree of certainty to which the examiners could express their identifications.”¹⁵⁴ The courts seem to note—without disapproval—

¹⁴⁸  Id. at 1177.
¹⁴⁹  Id. at 1178 (citing the NAS REPORT, supra note 2, at 5–20).
¹⁵⁰  These are not the only shortcomings the court notes. While proficiency tests have been conducted, none of them was done as a blind test, which raises doubts about the value of the tests. Id. at 1176.
¹⁵¹  Id. at 1180.
¹⁵³  Id. at *15 (quoting NATIONAL RESEARCH COUNCIL, BALLISTIC IMAGING 3 (2008)).
that one 2007 federal case allowed the expert to state a conclusion with “100% degree of certainty.” 155

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

While both of the courts recognize that toolmark comparison “does not have sufficient rigor to be evaluated as science,” 159 they categorize this testimony as either technological or other specialized testimony 160 and then proceed to admit the testimony with some limitations. What the courts continue to miss, however, is that the declaration of a match (a conclusion of individualization), is, by its very nature, a scientific finding:

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

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

---

156 See id. at *19.
157 Id. at *20.
158 Id. at *24.
161 NAS REPORT, *supra* note 2, at 43–44.
proving the triumph of belief perseverance in the face of contradictory (and compelling) information.

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

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

VII. CONCLUSION

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

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

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

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

163 Tetlock, supra note 132, at 290–91 (suggesting that pre-exposure accountability information may reduce the primacy effect).
appeared to in the language of *Melendez-Diaz*, the inferior federal courts likely will realize that they will be held accountable for their decisions on forensic science and will begin to evaluate the testimony in a more critical, thoughtful fashion.

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

INTRODUCTION

Good law depends on good science. The February 18, 2009, National Academy of Sciences report, *Strengthening Forensic Science in the United States: A Path Forward* (“NAS Report”), reveals that, for the most part, forensic science is bad science. The NAS Report also suggests that when confronted with forensic science, most courts make bad law.

The NAS Report is a wake-up call for courts, forensic scientists, law professors, and lawyers. For almost two decades, we have hoped (and more recently despaired) that *Daubert v. Merrell Dow Pharmaceuticals, Inc.* could revolutionize the courts by ensuring that judges only open their courtroom gates to demonstrably valid science. The NAS Report provides new and detailed evidence that the so-called “Daubert Revolution” has failed to transform the practice of science-based law or (law-based) forensic science—especially in the criminal courts.

Bad science leads to bad law because legal decisions based on many forms of forensic evidence are suspect. Less obviously, bad law leads to bad science because courts grant an apparent (if unwarranted) imprimatur of legitimacy by relying on forensic evidence of indeterminate or inadequate validity. The NAS Report illuminates this symbiosis and enhances our understanding of how underfunded, standardless, subjective forensic methods and lax judicial review of forensic evidence redound to both science and law.
The NAS Report is not entirely bleak. Retrospectively, the drafters recognized that some forensic fields have advanced and that a few (e.g., nuclear DNA analysis) are demonstrably valid. Prospectively, the drafters purport to chart a path forward (presumably) to a better future. This path leads to the creation of a National Institute of Forensic Science (“NIFS”), which will develop and enforce standards, help fund and promote independent research, and engage in other projects designed to enhance scientific validity within the various forensic fields. However, this path forward may be more circuitous than even a careful read of the report might reveal. Although the time and expense required to create a new federal agency suggest that the NIFS may be a hard sell, to those familiar with both progress and problems within the forensic fields, the NAS Report contained few shocking revelations. However, the report itself has subsequently provoked some surprises. One of the most potentially significant and unexpected developments was that on June 25, 2009, the Supreme Court relied on the NAS Report in Melendez-Diaz v. Massachusetts.

Melendez-Diaz is the most recent decision, since Crawford v. Washington, defining the scope of the Confrontation Clause. According to the Melendez-Diaz plurality, a defendant’s confrontation rights are violated when prosecutors introduce forensic lab reports without making the forensic analyst available for cross-examination. Melendez-Diaz was principally based on the plurality’s conclusion that the lab reports at issue were “testimonial statements” (under Crawford) and Justice Thomas’s fifth-vote concurrence limiting the Court’s holding to his view that extrajudicial statements only implicate the Confrontation Clause.

---

5 NAS REPORT, supra note 1, at 128–33.
6 Id. at 14–33.
7 Id. at 19–22.
8 129 S. Ct. 2527, 2536–38, 2555 (2009). Four days after Melendez-Diaz was decided, on June 29, 2009, the Supreme Court granted cert in Briscoe v. Virginia, 129 S. Ct. 2858 (2009) (mem). Because Briscoe involved the constitutionality of a state statute that provided criminal defendants with the right to subpoena the prosecution’s forensic analyst in lieu of confrontation, many commentators thought the question had been resolved by Melendez-Diaz. In fact, the unusual sequencing led to speculation about whether Melendez-Diaz would be overruled or limited and about the potentially pivotal role of Justice Sotomayor (who had been questioned about Melendez-Diaz during her confirmation hearings). The Court held oral argument on January 11, 2010, and then sixteen days later issued a unanimous, one-sentence per curiam opinion vacating the judgment of the Supreme Court of Virginia and remanding the case for further proceedings not inconsistent with Melendez-Diaz. Briscoe v. Virginia, 130 S. Ct. 1316 (2010) (per curiam).
9 541 U.S. 36 (2004). In Crawford, the Supreme Court held that the admission of testimonial hearsay invokes the Confrontation Clause. Id. at 61–62, 68–69.
10 See Melendez-Diaz, 129 S. Ct. at 2531–32.
11 Id. (characterizing the lab reports at issue in Melendez-Diaz as affidavits and noting that the Court’s decision in Crawford expressly included affidavits in the category of “testimonial statements” invoking the Confrontation Clause).
Clause if they have been adequately formalized. However, this narrowing of the Melendez-Diaz holding does not derogate from the significance of Justice Scalia’s rather remarkable opinion for four members of the Court. According to the plurality, because “[s]erious deficiencies have been found in the forensic evidence used in criminal trials,” defendants need confrontation “to weed out not only the fraudulent analyst, but the incompetent one as well.” The plurality’s reliance on the NAS Report to support these conclusions was the first time that the Court used the Confrontation Clause to resolve serious problems within the forensic fields. In effect, Melendez-Diaz has constitutionalized traditional evidentiary concerns. This new development raises a variety of new and interesting legal questions.

Melendez-Diaz is also notable because the Court explicitly recognized the “serious deficiencies” that continue to plague forensic evidence. As the plurality noted, the NAS Report attributed these ongoing problems to the “wide variability across forensic science disciplines with regard to techniques, methodologies, reliability, types and numbers of potential errors, research, general acceptability, and published material.” More specifically, Justice Scalia relied on the NAS Report to conclude that confrontation of forensic analysts must be guaranteed because “an analyst’s lack of proper training or deficiency in judgment may be disclosed in cross-examination.” Given the plurality’s focus on the problems that arise when courts routinely rely on fraudulent or incompetent experts, it is more than passing strange that not one member of the Court mentioned Daubert or the Federal Rules of Evidence.

This omission ignores the fact that for the past seventeen years, federal courts and more than half of state courts have relied on Daubert and the evidentiary rules to screen juries from seriously deficient forensic (and other expert) evidence proffered by any party in the criminal and civil courts. The best explanation for

---

12 Id. at 2543 (noting that, in his opinion, the Confrontation Clause is implicated in connection with extrajudicial statements only if they are contained in “formalized testimonial materials, such as affidavits, depositions, prior testimony, or confessions”) (internal quotations and citations omitted).
13 Id. at 2537.
14 Id. at 2538 (quoting NAS REPORT, supra note 1, at 6–7).
15 Id. at 2536–37.
16 See, e.g., Hodges v. Mack Truck, Inc., 474 F.3d 188, 195 (5th Cir. 2006) (stating judges have wide latitude in determining whether an expert and his testimony are reliable); Wills v. Amerada Hess Corp., 379 F.3d 32 (2d Cir. 2004) (holding that a district court did not abuse its discretion by refusing to admit a forensic toxicologist’s testimony in part because the expert’s theory had not been tested or subjected to peer review); Truck Ins. Exchange v. MagneTek, Inc., 360 F.3d 1206 (10th Cir. 2004) (allowing for the exclusion of an insurer’s expert physics testimony regarding the cause of a fire in the insured’s restaurant in-part because the testimony was not supported by scientific testing); U.S. v. Harris, 192 F.3d 580, 588 (6th Cir. 1999) (stating the Sixth Circuit always applies the Daubert test to evidence entered under Rule 702); Chemipal Ltd. v. Slim-Fast Nutritional Foods Intern., Inc., 350 F. Supp. 2d 582 (D. Del. 2004) (refusing to admit damages expert testimony due to concerns about methodology and overall reliability); U.S. v. Youngberg,
the Supreme Court’s recent failure to discuss evidentiary challenges to forensic science may be that *Melendez-Diaz* and *Daubert* reflect two different approaches to the same problem. The *Daubert* Court assumed that cross-examination was inadequate to the task of exposing and neutralizing specious expertise and that judges would be more accurate and consistent arbiters of scientific validity. After almost two decades of continued reliance on numerous forms of specious forensic evidence as documented in the NAS Report, *Melendez-Diaz* now suggests that confrontation can compensate for judges’ failure to screen *seriously deficient* expert evidence. Apparently, the *Melendez-Diaz* plurality believed that confrontation will succeed where *Daubert* has failed because “the analyst who provides false results may, under oath in open court, reconsider his false testimony” and because “[c]onfrontation is designed to weed out not only the fraudulent analyst, but the incompetent one as well.” Thus, at least for prosecution-sponsored expert evidence, *Melendez-Diaz* embraces the idea rejected by *Daubert*—that cross-examination will ensure that law does not rely on bad science.

There are good reasons to doubt that cross-examination is the right tool for this laudable purpose. In specific cases, defense counsel may prefer not to provide prosecution experts with the opportunity to flaunt their expertise. For example, in Professors Garrett and Neufeld’s recent study of forensic science evidence proffered by prosecutors in 137 cases where defendants were subsequently exonerated using DNA, they found that “[d]efense counsel rarely made any objections to the invalid forensic science testimony in these trials and rarely effectively cross-examined forensic analysts who provided invalid science testimony.” As a more general matter, *Melendez-Diaz* reinforces the idea (rejected in *Daubert*) that “[c]ross-examination is not merely accorded historic or structural importance in the adversary process; [but] it is also regarded as a panacea, a cure-all.” However, as Professor Jules Epstein recently opined, in

---

17 See *Daubert v. Merrell Dow Pharm., Inc.*, 509 U.S. 579, 592 (1993) (“Faced with a proffer of expert scientific testimony, then, the trial judge must determine at the outset . . . whether the expert is proposing to testify to . . . scientific knowledge . . . ”). However, the *Daubert* Court did acknowledge that after the gatekeeper judge had evaluated the proffered evidence, litigants might use “traditional and appropriate means of attacking shaky but admissible evidence,” which could include “[v]igorous cross-examination, presentation of contrary evidence, and careful instruction on the burden of proof.” *Id.* at 596.


19 *Id.* at 2537.

20 *See id.* at 2536–38.


many cases cross-examination of experts “actually impedes accurate fact-finding because leading questions are not always an appropriate or sufficient tool for truth finding . . . [and because] it lacks utility when confronting the honest-but-mistaken witness.”23 Regardless of whether Melendez-Diaz has the effect of neutralizing some of the specious forensic evidence that sneaks in under the Daubert gate, the new extension of the Confrontation Clause to forensic analysts has implications for both the practice of criminal law and the interpretation of constitutional doctrine.

In practice, although Melendez-Diaz is just over a year old, the case has already begun to confound the lower courts. A narrow reading of Melendez-Diaz would bar admission only of out-of-court statements by unavailable, non-testifying, prosecution-sponsored forensic experts when these statements are deemed “testimonial” because they are embodied within an affidavit or a similarly formalized document.24 However, Melendez-Diaz has proved almost impossible for the lower courts to decipher or apply. Over the past few months, state courts across the country have struggled to determine which lab reports (and other state records) are “testimonial statements” mandating confrontation.25 In fact, the post-Melendez-Diaz cases are so disparate and bizarre that they include decisions based (in whole or part) on the following factors: (1) whether a lab analyst subjectively anticipated that his autopsy report would be used in court; (2) whether a state requires that forensic reports be certified or accompanied by some form of attestation; (3) whether prosecutors can avoid confrontation by having testifying analysts describe lab reports prepared by non-testifying analysts (if they do not introduce the report in evidence); (4) whether the test and report were contemporaneous; (5) whether Melendez-Diaz is understood to guarantee confrontation of testimonial statements relating to experts’ methods or conclusions; (6) whether expert reports were created as part of a standard lab protocol without any effort to incriminate the defendant; and (7) efforts to reconcile Melendez-Diaz with Federal Rule of Evidence 703 or its state corollaries (which have long allowed experts to testify to opinions based on inadmissible evidence including out-of-court statements by non-testifying witnesses).26

Ambiguity is not the only problem with the post-Melendez-Diaz cases. If state legislators and administrative agencies seek to deprive future criminal defendants of confrontation opportunities, they can circumvent Melendez-Diaz by simply removing certification and attestation requirements from state records. This easy

24 See id. at 2531–32, 2542.
26 See infra Part VI.A (describing the effect of the Melendez-Diaz decision on challenges to forensic evidence and experts).
end run around the Sixth Amendment would have the paradoxical effect of making less reliable state records more readily admissible.

The doctrinal implications of Melendez-Diaz are equally complex. Melendez-Diaz seems to blend originalist and historical concerns about the Confrontation Clause with contemporary data about problems within the forensic community. Thus, it provokes, but does not resolve, questions that implicate numerous assumptions about the nature and purpose of confrontation, how the scope of the clause should be defined, and how expert evidence can and should be challenged. Melendez-Diaz itself offers little guidance. Although the Melendez-Diaz plurality acknowledged that “there are other ways—and in some cases better ways—to challenge or verify the results of a forensic test,” in Justice Scalia’s view, any other way would not be a sufficient alternative, because “the Constitution guarantees one way: confrontation.” Of course, this outcome was predetermined when the plurality deemed the certified lab report a “testimonial statement.”

This Article focuses on the NAS Report and the relationship between the concerns embodied within the report and the rapidly evolving confrontation doctrine. My thesis is that, taken together, these developments suggest that the “Daubert Revolution” has failed, at least in the criminal courts. Accordingly, the “path forward” charted by the NAS Report and the Melendez-Diaz Court does not lead to a Daubert-style solution involving better pretrial judicial screening or more conscientious application of evidentiary rules and standards. Instead, change should be sought in new directions, including coordinated efforts to standardize and improve the forensic fields and newly “constitutionalized” opportunities for criminal defendants to use cross-examination to expose specious and fraudulent forensic evidence.

This Article examines the practical and doctrinal implications of this new approach. Part I describes the origins of the NAS Report, including the creation of the committee and the congressional charge. Part II explores the NAS Report recommendations for the forensic fields and the recommendations for future criminal courts. Part III predicts the likely impact of these recommendations on both the forensic community and the courts. Part IV describes the Melendez-Diaz case and places the decision in its appropriate confrontation context. Part V explains how and why the Melendez-Diaz plurality’s reliance on the NAS Report reflects an effort to constitutionalize concerns about forensic evidence and experts. Finally, Part VI anticipates the future of Melendez-Diaz by exploring its impact on the lower courts and the likely implications of these new developments for the future of forensic science and law.

27 See Melendez-Diaz, 129 S. Ct. at 2536.
28 Id.
29 Id. at 2532.
I. THE NATIONAL ACADEMY OF SCIENCES REPORT

A. The Creation of the Forensic Science Committee

The NAS Report is based on a study authorized by Congress in November 2005. The NAS Report was based on a study authorized by Congress in November 2005. Four years ago, Congress charged the NAS with the creation of a new independent Forensic Science Committee (“FSC”) that would “assess the present and future resource needs of the forensic science community . . . [in order to] make recommendations for maximizing the use of forensic technologies and techniques to solve crimes, investigate deaths, and protect the public.” The FSC was co-chaired by Judge Harry T. Edwards of the United States Court of Appeals for the D.C. Circuit and Dr. Constantine Gatsonsis, a professor of biostatistics at Brown University. It was composed of forensic science practitioners, a variety of other scientists, and members of the legal community. Over the past four years, the committee heard testimony from many members of the forensic science community, including analysts from the FBI, the United States Secret Service, the National Institute of Justice, and other forensic science professional associations and advocacy groups. The committee also gathered evidence from judges, lawyers, and legal scholars.

The creation of the FSC reflected the most recent serious national effort to identify problems of validity and consistency across the range of forensic science fields and to shape the future of forensic evidence in our courts. Although the NAS Report may have the broadest scope and the longest view, this was not the first federally funded investigation into the forensic science community. For example, a decade earlier, the National Institute of Justice published a report entitled Forensic Sciences: Review of Status and Needs. Although the scope and detail of the earlier report were not as extensive as the NAS Report, the earlier report raised similar concerns about the need for more funding, better research, and greater standardization and coordination among local, state, and federal crime laboratories. The NAS has also generated previous reports on a variety of specific forensic questions including DNA analysis and ballistics identification.

---

30 NAS REPORT, supra note 1, at 1.
31 Id. at 1–2 (citing S. REP. NO. 109-88, at 46 (2005)).
32 See id. at v, xx.
33 See id. at v–ix, xx.
34 See id. at xi–xii, xix–xx, 1–2.
35 See id. at v–ix, xx, 1–2.
37 See id. at 3–4.
B. The Congressional Charge

The congressional charge to the FSC echoed concerns raised within and outside the forensic fields. In sum Congress instructed the FSC to identify the resource needs of state and local crime labs, make recommendations to maximize the use of existing forensic technologies, make recommendations that will increase the number of qualified forensic scientists, and disseminate practice guidelines for collecting and analyzing forensic evidence.40 Thus, the NAS Report was designed to provide current data on the general quality of forensic science and to generate specific recommendations that could be used to strengthen the forensic science community. The overarching goal was to improve the forensic science community’s ability to contribute to a fair and effective criminal justice system.41

In response to specific concerns about the forensic science community, the FSC engaged in an ambitious effort to understand the spectrum of forensic fields and their operation at the federal, state, and local levels.42 This focus reflected the fact that much of the impetus for the NAS Report had come from individuals within the forensic science community (some of whom were likely motivated by the significant lack of funding, especially among state and local crime labs).43 Over the past four years, the FSC explored a range of forensic fields including analysis of fingerprints, shoe prints, tire tracks, tool marks, firearms, hair, fiber, handwriting, paint, explosives, fire debris, bite marks, and bloodstains.44 The scope of this inquiry required the FSC to gather information from a wide range of forensic science laboratories and service providers, including crime labs operated by the FBI, the Department of Homeland Security, the Bureau of Alcohol, Tobacco, Firearms and Explosives, the Department of Defense, the National Bioforensic Analysis Center, the National Counterproliferation Center, state and

40 See NAS REPORT, supra note 1, at 1–2 (citing S. REP. NO. 109-88, at 46 (2005)).
41 See id. at 4–5.
42 See id. at 5–6, 55–56.
43 See National Research Council’s Publication “Strengthening Forensic Science in the United States: A Path Forward”: Hearing Before the H. Subcomm. on Crime, Terrorism, and Homeland Security of the H. Comm. on the Judiciary, 111th Cong. 16, 19–20 (2009) [hereinafter Hearing Before the H. Subcomm. on Crime, Terrorism, and Homeland Security] (statement of Peter M. Marone, Director, Virginia Department of Forensic Science) (describing how the NAS Report was motivated by concerns that came in part from state and local crime labs and the state and local medical examiner communities, which have not received adequate financial support).
44 NAS REPORT, supra note 1, at 3. For the FSC’s description of and analysis of various forensic fields, see id. at 127–182.
local crime labs, forensic science funding organizations, and forensic science professional associations.45

C. Forensic Science in the Criminal Courts

In response to specific concerns about the admission and use of forensic science evidence in court, the FSC also examined both the extant and potential interplay between forensic science and law.46 Although the FSC analysis of forensic science in litigation was limited to the federal criminal courts, this is an essential component of the NAS Report because, with the notable exception of DNA analysis, courts are the only real consumers of forensic evidence.

As part of the FSC exploration of the admission and use of forensic evidence in the criminal courts, the committee assessed the effectiveness of the Federal Rules of Evidence and the operation of Daubert challenges.47 The NAS Report recommendations are discussed in detail below.48 However, in essence the NAS Report concluded that there are vast systemic problems with the use of forensic science in criminal litigation because courts “continue to rely on forensic evidence without fully understanding and addressing the limitations of different forensic science disciplines.”49 Moreover, according to the NAS Report, seventeen years after Daubert required that judges pre-screen challenged expertise to determine whether it is “scientific knowledge,”50 there has been little improvement in the quality of forensic evidence.51 In the view of the FSC, this is because courts “have not with any consistency or clarity imposed standards ensuring the application of scientifically valid reasoning and reliable methodology in criminal cases involving Daubert questions.”52

II. THE NAS REPORT: ANALYSIS AND RECOMMENDATIONS

A. Recommendations for the Forensic Science Community

The FSC explicitly acknowledged that the forensic science community suffers from systemic structural problems.53 In fact, these problems are so pervasive that any accurate assessment of the forensic science community is impeded by its fragmented nature, which “makes it difficult to gather data on the entire universe

45 See id. at 4, 57–77.
46 See id. at 85–88.
47 See id. at 90–98.
48 See infra Part II.
49 NAS REPORT, supra note 1, at 85.
51 NAS REPORT, supra note 1, at 106.
52 Id. at 11.
53 Id. at 77–80.
of forensic service entities and activities."54 As the NAS Report revealed, the forensic science community is subdivided among a vast spectrum of distinct fields.55 It is further fragmented by the fact that forensic techniques are practiced in autonomous laboratories that operate at the local, state, and federal levels.56 Despite these impediments to a comprehensive understanding of the range of forensic fields, the FSC found that “the large amount of information provided to the committee by people engaged in the forensic science enterprise and by experts who have studied how well that enterprise functions all points to a system that lacks coordination and that is underresourced in many ways."57 Moreover, because most of the forensic scientific community lacks adequate funding and consistent professional standards, existing problems are further compounded by the fact that many forensic analyses are performed by “practitioners with different levels of education and training and different professional cultures and standards for performance."58

The FSC also recognized that most forensic fields suffer from endemic localized problems engendered by the nature of their common objective.59 Forensic fields almost invariably engage in a process of “individualization.” The goal of individualization is to match evidence found at a crime scene or on a law enforcement database to a specific suspect.60 Although the process may utilize a range of technologies, a vital component of most forensic individualization is the subjective analysis of the human interpreter.61 Recently, much has been written about both covert and overt bias among forensic scientists,62 including at least one essay complaining that this social science evidence was given short shrift by the FSC.63 To be fair, the NAS Report attempted to address a wide range of specific lab analyst problems including human observer bias, lack of consistent education and training requirements, inconsistent terminology, and the lack of uniform lab accreditation and employee certification standards.64

54 Id. at 77.
55 See id. at 78.
56 See id. at 77–78.
57 Id. at 77 (emphasis added).
58 Id. at 78.
59 See id. at 87.
60 See id.
61 See id.
63 D. Michael Risinger, The NAS/NRC Report on Forensic Science: A Glass Nineteenth Full (This Is About the Other Tenth), 50 JURIMETRICS J. 21 (2009).
64 See NAS REPORT, supra note 1, at 3, 21.
In light of the spectrum of forensic disciplines and the range of concerns, the NAS Report does not contain any one-size-fits-all forensic science community recommendations. To the extent there are any general findings, they are embodied within the FSC’s conclusion that “forensic laboratories are underresourced and understaffed”65 and “the knowledge base that underpins [forensic] analysis and interpretation of evidence—are not as strong as they could be”66 in part because “[t]he forensic science system . . . has only thin ties to an academic research base that could support the forensic science disciplines and fill knowledge gaps.”67 More general conclusions are also contained within the finding that “[w]ith the exception of nuclear DNA analysis . . . no forensic method has been rigorously shown to have the capacity to consistently, and with a high degree of certainty, demonstrate a connection between evidence and a specific or individual source.”68

The NAS Report includes thirteen specific recommendations responsive to the committee’s congressional charge.69 These begin with the threshold recommendation that Congress create and fund an independent federal National Institute of Forensic Science (“NIFS”).70 The remaining twelve recommendations principally elaborate on how a new NIFS could transform the forensic science community.71 These recommendations are described in some detail in the report,72 but they can be summarized as follows: (1) establish standard terminology to be used within the forensic fields to report on and testify about forensic science investigations;73 (2) fund peer-reviewed research to demonstrate the validity of forensic methods and develop and establish quantifiable measures of the validity of forensic analyses (including quantifiable measures of uncertainty in conclusions);74 (3) maximize the independence of forensic laboratories and professionals from law enforcement and prosecutors’ offices;75 (4) “encourage research programs on human observer bias and sources of human error in forensic examinations”;76 (5) work with the National Institute of Standards and Technology to advance standards that would control “measurement, validation, reliability, information sharing, and proficiency testing in forensic science and to establish protocols for forensic examinations”;77 (6) ensure mandatory laboratory accreditation and mandatory

---

65 Id. at 14.
66 Id.
67 Id. at 15.
68 Id. at 7.
69 Id. at 19–33.
70 Id. at 19.
71 Id. at 22–33.
72 Id. at 19–33.
73 Id. at 22.
74 Id. at 22–23.
75 Id. at 24.
76 Id.
77 Id. at 24–25.
individual certification of forensic science professionals; 78 (7) “establish routine quality assurance and quality control procedures”; 79 (8) “establish a national code of ethics for all forensic science disciplines”; 80 (9) improve and develop graduate education programs in multidisciplinary fields essential to the practice of forensic science; 81 (10) improve the quality of medico-legal death investigations; 82 (11) “launch a new broad-based effort to achieve nationwide fingerprint data interoperability”; 83 and (12) work with the Centers for Disease Control and Prevention, the FBI, forensic scientists, and crime scene investigators to ensure the capacity to manage and analyze evidence from future events that affect homeland security. 84 As discussed below, some insight into the viability of these recommendations might be gleaned from recent congressional hearings on the NAS Report.

B. Recommendations for the Criminal Courts

The FSC examined the interplay between forensic science and law by exploring the admission of forensic science evidence in criminal litigation. 85 According to the NAS Report, “[t]he law’s greatest dilemma in its heavy reliance on forensic evidence, however, concerns the question of whether—and to what extent—there is science in any given ‘forensic science’ discipline.” 86

1. The Failure of Daubert

The question of whether there is science in any scientific evidence has preoccupied evidence scholars (at least) since 1993, when the Supreme Court was galvanized into action in Daubert v. Merrell Dow. 87 I have written elsewhere about the myopic post-Daubert approach to the broad range of science and law questions, 88 but the NAS Report further supports the conclusion that a post-Daubert emphasis on “rules governing the admissibility of forensic evidence [and] the applicable standards governing appellate review of trial court decisions . . . is

78 Id. at 25.
79 Id. at 26.
80 Id.
81 Id. at 27–28.
82 Id. at 29–30.
83 Id. at 31.
84 Id. at 33.
85 See id. at 85–110.
86 Id. at 9, 87.
88 See, e.g., Joëlle Anne Moreno, Eyes Wide Shut: Hidden Problems and Future Consequences of the Fact-Based Validity Standard, 34 SETON HALL L. REV. 89, 90–91 (2003) (explaining that, under Daubert, judges may not defer to the scientific community but must instead decide for themselves whether evidence is scientifically reliable).
inadequate to the task of curing the documented ills of the forensic science disciplines.” As the NAS Report recognized, almost two decades after Daubert, trial courts “continue to rely on forensic evidence without fully understanding and addressing the limitations of different forensic science disciplines.” These systemic problems are compounded by appellate courts that “have not with any consistency or clarity imposed standards ensuring the application of scientifically valid reasoning and reliable methodology in criminal cases involving Daubert questions.”

2. Impediments to Doctrinal Analysis

After acknowledging the failures of the existing legal regime to effectively root out specious forensic science, the NAS Report proposed that a better approach would start with

- two very important questions that should underlie the law’s admission of and reliance upon forensic evidence in criminal trials: (1) the extent to which a particular forensic discipline is founded on a reliable scientific methodology that gives it the capacity to accurately analyze evidence and report findings and (2) the extent to which practitioners in a particular forensic discipline rely on human interpretation that could be tainted by error, the threat of bias, of the absence of sound operational procedures and robust performance standards.

However, discerning the validity of the range of forensic science methodologies and their reliance on subjective human interpretation is a difficult, complex, and time-consuming task. It is also a task that many members of both the forensic and legal communities (because of their traditional interdependence) may not be anxious to undertake.

Moreover, even if we assume that during the post-Daubert period the operation of the admissibility rules and standards has generally improved, these developments are extremely difficult to identify or measure. When expert evidence is challenged, trial courts frequently issue judgments on the admission of evidence without published opinions; especially in criminal cases and even in the federal courts, evidentiary rulings are only infrequently subject to appellate review. Thus, as the NAS Report noted, “reported opinions do not offer in any way a complete sample of federal trial court dispositions of Daubert-type questions in

89 NAS REPORT, supra note 1, at 85.
90 Id. at 53, 85.
91 Id. at 96.
92 Id. at 87.
93 See id. at 11–12, 110 (noting the “highly deferential nature of the appellate review afforded trial courts’ Daubert rulings”).
This may explain why there have been few efforts to quantify Daubert’s real impact on the criminal and civil courts.

3. Two Models of Valid Forensic Science

Perhaps in an effort to illustrate that scientists and courts sometimes get it right, the NAS Report also discussed two forensic fields where both the science and law are generally sound. The report described the fact that, like most forensic evidence, nuclear DNA evidence is universally accepted for individualization purposes in U.S. courts. However, unlike most forensic evidence, judges and jurors should rely on DNA evidence because this evidence is the product of the only forensic method that “has been rigorously shown to have the capacity to consistently, and with a high degree of certainty, demonstrate a connection between evidence and a specific individual or source.” The second forensic field identified with approval by the FSC, drug identification, poses fewer problems because the analysts’ goal is not individualization. Forensic drug identification generally relies on widely accepted principles and technologies of chemical analysis, and the analyst’s goal is limited to determining the chemical composition of the recovered substance. Thus, forensic substance identification evidence is also routinely and appropriately admitted in the criminal courts. Judicial decisions admitting nuclear DNA analyses and drug identification evidence reflect examples of accurate determinations by the courts that certain forms of forensic evidence are based on valid forensic science methodologies.

---

94 Id. at 11, 97.
95 Id. at 99–102.
96 Id. at 99–100.
97 Id. at 7. Nuclear DNA evidence is touted throughout the NAS Report as the proverbial “gold standard” of forensic evidence. However, it is worth noting that an August 18, 2009 article in The New York Times described how Israeli scientists have proved that it is possible to fabricate DNA evidence. Andrew Pollack, Scientists Show That It’s Possible to Create Fake DNA Evidence, N.Y. Times, Aug. 18, 2009, at D3. Dr. Dan Frumkin and his team report that they were able to fabricate both blood and saliva samples that contained non-donor DNA. Id. Perhaps even more troubling, they claim that, without access to biological source material, they could build a sample of DNA to match a DNA database profile. Id. However, this is very new research and, as the article indicates, Dr. Frumkin is a founder of Nucleix, a for-profit Israeli company that has developed a test designed to distinguish fake DNA that Nucleix plans to market to forensics laboratories throughout the world. Id.
98 See NAS REPORT, supra note 1, at 134.
99 See id. at 101–02.
100 See id. at 7, 102.
4. Avoiding the Vicious Circle

The NAS Report recognized that the legal system creates bad incentives for the forensic community and vice-versa.\(^{101}\) Despite the evidentiary rules and \textit{Daubert}, judges consistently fail to prevent “forensic science methodolog[ies] [from being] condoned by the courts before the techniques have been properly studied and their accuracy verified.”\(^{102}\) For example, even when defendants raise \textit{Daubert} objections, forensic evidence is routinely admitted without serious judicial scrutiny, and this evidence includes “‘even the most vulnerable forensic sciences—hair microscopy, bite marks, and handwriting.’”\(^{103}\) These problems are further compounded when judges admit this evidence “[by] citing earlier decisions rather than facts established at a hearing.”\(^{104}\) In most forensic fields, there is “‘no evident reason why [rigorous, systematic] research would be infeasible’”;\(^{105}\) but it is simply not being done. In fact, many judges “appear to be loath to insist on such [rigorous, systematic] research as a condition of admitting forensic science evidence in criminal cases, perhaps because to do so would likely ‘demand more by way of validation [than] disciplines can presently offer.’”\(^{106}\) These problems will likely multiply in the future as the courts continue to rely on an ever-expanding range of forensic evidence.

5. The Future of Forensic Evidence in the Criminal Courts

Although the bulk of the NAS Report recommendations focused on improving the practice of forensic science, not the practice of law, law and forensic science are increasingly interdependent.

Forensic science is crucial to the criminal justice system from start to finish. During an investigation, forensic science evidence is a vital exculpatory tool, often excluding potential suspects and narrowing the focus of investigations for the police. Forensic evidence may provide important clues to places, objects or people that can lead police to an arrest before another crime has been committed by a particular individual, thus harnessing the power of crime prevention. In a post-mortem context, forensic examinations are imperative for suspicious

\(^{101}\) \textit{See id.} at 16–17 (asserting that the forensic science field needs strong governance to encourage jurisdictions to adopt best practices and discourage bad practices).

\(^{102}\) \textit{See id.} at 109 (emphasis added).

\(^{103}\) \textit{Id.} at 107 (quoting P.J. Neufeld, \textit{The (Near) Irrelevance of \textit{Daubert} to Criminal Justice and Some Suggestions for Reform}, 95 \textit{Am. J. Pub. Health} S107, S109, S110 (2005)).

\(^{104}\) \textit{Id.}


\(^{106}\) \textit{Id.}
deaths and are vital to determining a cause of death. . . . After an arrest, forensic evidence often expedites dispositions of cases and, frequently, when confronted with the results of forensic analyses, defendants choose to accept a plea rather than assume the risk of going to trial. At trial, forensic evidence and the expert testimony proffered by forensic scientists can be key to securing a conviction or appropriate sentence.  

As discussed above, the NAS Report proposes long-term, systemic, transformative recommendations for the forensic community designed to ensure that future “forensic science experts will be better able to analyze evidence and coherently report their findings in court.” The report was released in early 2009, in response to significant concerns from within and outside the forensic fields. Although the NAS Report recommended a path forward to a future of science-based forensic evidence, implementing these recommendations will require substantial administrative coordination and a significant allocation of funds. Over the past few months, Congress has begun to consider the costs and benefits of the NAS Report recommendations.

III. PREDICTING THE IMPACT OF THE NAS REPORT

It is far too early to accurately assess the impact of the NAS Report on the forensic community. However, in the wake of its release, a number of congressional committees have held public hearings to consider how the federal government should respond to the report’s recommendation. These hearings may provide some insight into the likely congressional response to the NAS Report. They may also reveal nascent congressional reaction to the proposal to create a new NIFS, which is a predicate to most of the other recommendations. As discussed in more detail below, the initial congressional response revealed little beyond an attempt to understand the details of the NAS Report. More recent hearings have also included testimony supporting responsive federal action. However, more recent hearings have also begun to incorporate more skeptical

testimony suggesting that the NAS Report concerns are exaggerated, arguing that
the creation of an NIFS is unnecessary or inappropriate, and revealing sensitivity
to any Congressional response that might be viewed as undermining the legitimacy
of the criminal justice system.111 Because the NAS Report recommendations are
structured so that most of the changes would be implemented through a new
federally funded oversight agency,112 a threshold question for Congress is whether
federal funds will be allocated to create an NIFS.

A. Predicting the Impact on the Forensic Science Community:
The Initial Congressional Response

1. Senate Judiciary Committee

Immediately after the NAS Report was released, the Senate Judiciary
Committee met to consider its findings.113 At this hearing, the committee’s
chairman, Senator Patrick Leahy, opined that the problems identified in the report
were very serious because they “go to the heart of our criminal justice system.”114
Apparently, many of the NAS Report conclusions were unfamiliar to the
committee members and to the public. For example, Senator Leahy noted that the
NAS Report dispelled prevalent misconceptions regarding forensic science
because it illuminated the fact that forensic methods bear little resemblance to
“television shows like ‘CSI,’ [because] forensic scientists too rarely get to review
crime scene evidence in sleek, ultra-modern, state-of-the-art laboratories.”115

2. House Committee on Science and Technology

That same month, the House Committee on Science and Technology held
hearings on the NAS Report.116 The committee heard testimony from Professor
Peter Neufeld, who is the co-director of the Innocence Project.117 Professor
Neufeld reminded the committee that that the criminal justice system relies heavily
on a wide range of forensic evidence.118 According to the Bureau of Justice
Statistics 2005 Census, non-DNA forensic lab analysis requests constitute 97
percent of the workload in publicly funded forensic crime laboratories.119

Professor Neufeld emphasized the link between bad forensic evidence and
wrongful convictions. According to Professor Neufeld,

111 See discussion Part III.B.
112 NAS REPORT, supra note 1, at 19–20.
113 Hearing Before the S. Comm. on the Judiciary, supra note 109, at 1.
114 Id.
115 Id.
116 Hearing Before the Subcomm. on Technology and Innovation, supra note 109.
117 Id. at 32–39.
118 Id. at 34.
119 Id.
[U]nvalidated and improper forensics contributed to approximately 50% of wrongful convictions overturned by DNA testing. . . . [And] we have had wrongful convictions based on unvalidated or misapplied serological analysis, microscopic hair comparisons, bite mark comparisons, shoe print comparisons, fingerprint comparisons, forensic geology (soil comparison), fiber comparison, voice comparison, and fingernail comparison . . . .

Professor Neufeld specifically endorsed the NAS Report conclusion that the problems of jury exposure to and reliance upon bad forensic evidence cannot improve without better national standardization, coordination, and centralization. According to Professor Neufeld, a new NIFS would create opportunities to “conduct research into the scientific validity and reliability of forensic disciplines and set standards for their use in the courtroom.” Thus, the NIFS would ensure a future where we “don’t have 50 states operating under 50 [different] definitions of ‘science.’” In Professor Neufeld’s view, these developments are essential because “forensic science in America needs one standard of science so we can have one standard for justice.”

B. Recent Congressional Response

In mid-May 2009, the House Judiciary Committee Subcommittee on Crime, Terrorism, and Homeland Security held its own hearings on the NAS Report. More recently, on September 9, 2009, the Senate Judiciary Committee held a second round of hearings. These more recent hearing explored the NAS Report in greater detail and revealed nascent skepticism about its findings and recommendations.

1. House Judiciary Committee’s Subcommittee on Crime, Terrorism, and Homeland Security Responds

(a) Daubert and the Federal Rules of Evidence Are Inadequate

According to the chairman, Representative Robert Scott, the NAS Report confirmed that the current post-Daubert legal regime of judicial gatekeeping has not worked. In his view, “[W]e have decided that the judge would be inadequate
as a gatekeeper to decide what kind of scientific evidence comes in and comes out.”

(b) Distinguishing Among the Forensic Fields

The committee heard testimony from Pete Marone, the Director of the Commonwealth of Virginia’s Department of Forensic Science. According to Mr. Marone, the NAS Report confirmed that

[the] disciplines based on biological or chemical analysis, such as toxicology, drug analysis, and some trace evidence sub-disciplines such as explosives, fire debris, polymers to include paint and fiber analysis, are generally well-validated and should not be included in the same category as the experience-based disciplines, such as fingerprints, firearms and toolmarks, and other pattern-recognition types of analysis.127

Mr. Marone focused on the need for more and better independent research within the forensic fields.128 According to Mr. Marone, “We need studies, for instance, that look at large populations of fingerprints and toolmarks so as to quantify how many sources might share similar features . . . [i]n addition to investigating the limits of the techniques themselves[, and] research is also needed on the issues of context effect and examiner bias.”129 Thus, most of Mr. Marone’s testimony, based on his professional experience directing Virginia’s state crime lab system, was consistent with the conclusions and recommendations contained within the NAS Report.

(c) Caution and Skepticism

The committee also heard testimony from Kenneth Melson, acting director of the Bureau of Alcohol, Tobacco, Firearms and Explosives.130 Mr. Melson was more cautious than Mr. Marone. His testimony clarified that the NAS Report “does not, and was never intended to, comprehensively assess the forensic sciences themselves . . . [or] undermine the use of forensic science generally—or any specific discipline—in the courtroom.”131 According to Mr. Melson, “the report highlights the lack of research and other scientific validation methods within several disciplines.”132 However, Mr. Melson’s principal concern seemed to be the

126 Id. at 62.
127 Id. at 21.
128 Id.
129 Id.
130 Id. at 5–15.
131 Id. at 9.
132 Id. at 13.
risk that the NAS Report’s conclusions had been, and would continue to be, overstated. Mr. Melson testified that “the report does not take the position that any of the forensic disciplines is scientifically invalid. . . . [Yet it] has been taken by the public and the defense bar as labeling forensics not ‘real’ science.”

Mr. Melson also appeared to oppose the creation of a National Institute of Forensic Science with federal oversight authority. In his view, an NIFS is not necessary, because solutions are already being generated from within the forensic community. According to Mr. Melson, these include the following programs: (1) the National Institute of Justice and the National Institute on Standards and Technology joint Expert Working Group on Human Factors in Latent Print Analysis; (2) nine FBI-sponsored Scientific Working Groups composed of state and federal experts in nine different forensic fields; and (3) National Institute of Justice efforts to facilitate and encourage forensic lab accreditation and analyst certification. Given the time and expense associated with the creation of a new centralized federal agency, Mr. Melson’s alternative approach of allowing the forensic community to fix itself via a variety of different initiatives and programs may have congressional appeal.

2. Senate Judiciary Committee

During the second Senate Judiciary Committee hearing on the NAS Report, held on September 9, 2009, the committee heard testimony from Barry Matson, the Deputy Director of the Alabama District Attorneys Association. Mr. Matson disagreed with the recommendation that the federal government create an NIFS for a variety of reasons. Like Mr. Melson, Mr. Matson stated that an NIFS is unnecessary. However, Mr. Matson went further, suggesting that the creation of an NIFS would reflect a blatant effort to politicize science. In Mr. Matson’s view, “[f]orensic sciences is [sic] the search for truth and if you’re going to have an agency with a new director appointed every four years and different ideologies

133 See id. at 13–15.
134 Id. at 13.
135 Id. at 12–13.
136 Id. at 11–12.
138 See id. (“[T]here are institutions available that are already meeting many of the challenges mentioned in the NAS report.”); Mary Orndorff, Congress Looks at Court Evidence Standards: Sessions Disputes Need for New Agency, 122 THE BIRMINGHAM NEWS 4 (Sept. 10, 2009).
139 Orndorff, supra note 138.
140 Id.
coming in and new national bureaucracies, it’s not what we need.” Of course, Mr. Matson’s testimony ignored the wide range of federal agencies, from the EPA to the FDA, that routinely evaluate scientific information. Mr. Matson’s testimony received support from his home state senator, Jeff Sessions. Senator Sessions, a former prosecutor, acknowledged that the NAS Report could create uncertainty about evidence long-relied upon by police, prosecutors, judges, and juries. For example, without addressing any of the specific concerns raised in the NAS Report (which included a section on fingerprints and a discussion of the Brandon Mayfield Madrid-train bombing debacle), Senator Sessions said simply, “I don’t accept the idea that they seem to suggest that fingerprints are not proven technology.”

Finally, Senator Leahy noticed the relationship between the NAS Report and the Supreme Court’s recent Melendez-Diaz v. Massachusetts decision. According to Senator Leahy, after Melendez-Diaz, cross-examination of prosecutors’ expert witnesses should play a much more important role in routing out the type of specious forensic science evidence identified in the NAS Report because government-sponsored experts will now be forced to explain and defend their conclusions. Senator Leahy suggested that Melendez-Diaz and the NAS Report together will enhance the quality of forensic evidence proffered by prosecutors in the nation’s criminal courts. According to Senator Leahy, the Melendez-Diaz decision “stems from a recognition that forensic findings may not always be as reliable as we would hope or they might appear.” As the congressional hearings continue, additional information regarding the political and economic viability of the NAS Report recommendations will be revealed.

141 Id.
142 Id.
143 See id.
144 NAS REPORT, supra note 1, at 46.
145 Orndorff, supra note 138.
146 129 S. Ct. 2527 (2009).
147 National Academy of Sciences Report: Strengthening Forensic Science in the United States: A Path Forward: Hearing Before the Senate Judiciary Comm., 111th Cong. (2009) (statement of Sen. Patrick Leahy, Chairman, Sen. Judiciary Comm.) (noting that the Supreme Court held in Melendez-Diaz “that forensic examiners must present evidence in court and be subject to cross examination, rather than simply submitting reports of their findings. This Supreme Court holding stems from a recognition that forensic findings may not always be as reliable as we would hope, or they might appear”).
148 Id. (“The report issued by the National Academy of Sciences earlier this year is detailed and far-reaching, and can provide a foundation for building broad consensus for change.”).
149 Id.
C. Predicting the Impact of the NAS Report on the Criminal Courts

Judges, prosecutors, defense counsel and legal commentators have just begun to speculate about the impact of the NAS Report on the courts. One of the first, Professor Edward J. Imwinkelried, has reflected upon the fact that in the past,

[the] NAS’ issuance of reports has sometimes persuaded courts to change their stance on the admissibility of specific types of scientific evidence. . . . When a scientific organization as large and highly respected as the NAS raises questions about the reliability of an expert technique, that development arguably proves the existence of a major controversy that is the antithesis of the general acceptance required by [United States v.] Frye.150

Professor Imwinkelried may be correct that the NAS Report could have a more powerful effect in states still governed by Frye, because it may be used to effectively demonstrate a lack of “general acceptance.” However, there is reason to believe that the NAS Report could have a similar impact in the federal courts and Daubert jurisdictions because (even under Daubert) judges continue to rely on “general acceptance” as an important admissibility criterion.151 Some commentators have also speculated that the NAS Report may prove useful to lawyers. Professor Imwinkelried has opined that the report’s specific findings may generate better opportunities for lawyers to challenge individualization testimony based on matching fingerprints, toolmarks, firearms, hairs/fibers, handwriting, or bitemarks.152 Professor Jules Epstein has suggested that in Daubert jurisdictions, litigants might use the NAS Report to ask courts to revisit earlier decisions admitting evidence.153 If these evidence scholars are


151 See Sophia Gatowski et al., Asking the Gatekeepers: A National Survey of Judges on Judging Expert Evidence in a Post-Daubert World, 25 LAW & HUM. BEHAV. 433, 447 (2001) (“[T]he vast majority of judges . . . , regardless of operating admissibility standard, indicated that general acceptance was a useful criterion for determining the merits of the proffered scientific evidence . . . .”); see also id. at 437 (citing Daubert, 509 U.S. at 594) (noting that “[t]he [Daubert] Court recognized . . . the ‘general acceptance’ of the proposed testimony of the scientific community” as one of the “guidelines” for consideration).

152 See Imwinkelried, supra note 150, at 11–12 (noting that opponents of expert testimony could make a Daubert objection based on the NAS Report).

153 Jules Epstein, The NAS Report: An Evidence Professor’s Perspective, IT’S EVIDENT, July 20, 2009, at 1 (“Because the Daubert inquiry is fundamentally one of
correct that future lawyers (and especially future defense counsel) will begin to rely on the specific conclusions of the NAS Report to challenge “individualization”/source attribution evidence involving ballistics or fingerprint evidence, they are probably also correct that the data contained in the report could make these challenges more persuasive.

Finally, some commentators have speculated that the NAS Report might spark systemic change in states that still rely on Frye. These predictions assume that the NAS Report will be viewed as revealing such serious problems within the forensic community that it would provoke Frye jurisdictions to convert to a Daubert standard. This, in turn, could result in greater scrutiny of the “traditional methodologies that Frye jurisdictions routinely admit as generally accepted.”

These are all optimistic predictions. Clearly, the NAS Report has the capacity to illuminate specific problems for judges, lawyers, and jurors who must decide whether to admit and how to use different forms of forensic evidence. However, the more general impact of the NAS Report will be determined by the eventual congressional response and the overall reaction from the courts. So far, the most important response to the NAS Report from the courts came from the Supreme Court at the very end of the 2008–09 term.

IV. THE NAS REPORT IN THE SUPREME COURT: MELENDEZ-DIAZ V. MASSACHUSETTS

On June 25, 2009, the Supreme Court decided Melendez-Diaz v. Massachusetts. Melendez-Diaz is the most recent decision defining the scope of the Confrontation Clause. As discussed in more detail below, the Melendez-Diaz plurality concluded that a state statute designed to admit certified forensic reports was an impermissible end run around the Confrontation Clause. Rather evidentiary reliability, there is no bar to asking a court to revisit prior decisions of admissibility.”).

See Imwinkleried, supra note 150, at 12 (“One question is whether the [NAS] report should prompt additional Frye jurisdictions to rethink their standard for admitting scientific testimony.”).

See id. at 11 (noting that the NAS Report “could play a major role in shaping the future treatment of expert evidence”).

See id. (“Many of the techniques discussed in the NAS report are traditional methodologies that Frye jurisdictions routinely admit as generally accepted. The argument could be made that the NAS report is potent evidence that the Frye test is ineffective in separating the wheat from the chaff.”).


See infra Part IV.B.

Melendez-Diaz, 129 S. Ct. at 2532 (quoting Crawford v. Washington, 541 U.S. 36, 54 (2004)) (“In short, under our decision in Crawford the analysts’ affidavits were testimonial statements, and the analysts were ‘witnesses’ for purposes of the Sixth
surprisingly for a committed originalist, Justice Scalia found significant support for his confrontation analysis in the just-released NAS Report conclusion that “[t]he forensic science system . . . has serious problems.”\footnote{Melendez-Diaz, 129 S. Ct. at 2537 (quoting NAS Report, supra note 1, at xx) (emphasis omitted).} This linkage enabled the Court to craft new constitutional solutions to old evidentiary problems.

*Melendez-Diaz* is not a simple case, and any analysis of its implications for the future of forensic science in the criminal courts is complicated by its constitutional context. Legal history scholars far better suited to the task have explored the history of the Confrontation Clause at length in a wide range of books and articles. However, to understand *Melendez-Diaz* it is important to start with the fact that, although the Sixth Amendment was incorporated to the states almost fifty years ago in *Pointer v. Texas*,\footnote{Pointer v. Texas, 380 U.S. 400, 403 (1965) (“We hold today that the Sixth Amendment’s right of an accused to confront the witnesses against him is likewise a fundamental right and is made obligatory on the States by the Fourteenth Amendment.”).} the Supreme Court’s confrontation doctrine was revitalized and transformed just five years ago in *Crawford v. Washington*.\footnote{Crawford v. Washington, 541 U.S. 36, 54 (2004).}

### A. Understanding Melendez-Diaz in Context: Crawford v. Washington

#### 1. Rejecting Ohio v. Roberts

In 2005, *Crawford* significantly expanded the criminal defendant’s Sixth Amendment right “to be confronted with the witnesses against him; [and] to have compulsory process for obtaining witnesses in his favor . . . .”\footnote{U.S. CONST. amend. VI.} For the past quarter-century, conflicts between the Confrontation Clause and the Federal Rules of Evidence had been resolved under the rule of *Ohio v. Roberts*.\footnote{Ohio v. Roberts, 448 U.S. 56 (1980).} In *Ohio v. Roberts* the Supreme Court held that the Confrontation Clause restricts otherwise admissible hearsay in two ways. First, it requires that the prosecution produce the “witness against” the defendant or demonstrate her unavailability.\footnote{Id at 65.} Second, if the witness is unavailable, “the [Confrontation] Clause countenances only hearsay marked with such trustworthiness that ‘there is no material departure from the reason of the general rule.’”\footnote{Id. (quoting Snyder v. Massachusetts, 291 U.S. 97, 107 (1934)).} Prosecutors could satisfy the “indicia of reliability”\footnote{Roberts, 448 U.S. at 65–66.} requirement either (1) by convincing the court that the proffered out-
of-court statement fit within a “firmly rooted hearsay exception” or (2) by demonstrating that the statement had “particularized guarantees of trustworthiness.” Thus, under *Ohio v. Roberts*, confrontation analysis was inextricably intertwined with the Rules of Evidence, and a defendant’s confrontation rights could expand and contract based on judicial determinations of evidentiary reliability.

The *Crawford* Court took a dim view of *Ohio v. Roberts*. The Sixth Amendment guarantees the criminal defendant the right of confrontation; yet *Ohio v. Roberts* enabled judges to continue to substitute their own ad hoc determinations of reliability for the “crucible of cross-examination.” This sidestepping of the Confrontation Clause bore the full brunt of Justice Scalia’s estimable ire. Writing for the *Crawford* majority in a case that involved admission of a witness statement to the police that had been deemed admissible as a statement against interest by the trial court, he opined that the *Ohio v. Roberts* practice of “[admitting] statements deemed reliable by a judge is fundamentally at odds with the right of confrontation” and that “[d]ispensing with confrontation because testimony is obviously reliable is akin to dispensing with jury trial because a defendant is obviously guilty.” Although much has been written elsewhere about *Crawford*, one aspect of Justice Scalia’s *Crawford* analysis is especially important to an accurate understanding of the Court’s decision five years later in *Melendez-Diaz*.

---

168 *Id.* at 66.
169 *Id.*
170 *Crawford v. Washington*, 541 U.S. 36, 54–59 (2004). The *Crawford* Court further stated, “[T]he [Ohio v. Roberts] framework is so unpredictable that it fails to provide meaningful protection from even core confrontation violations.” *Id.* at 63.
171 *Id.* at 61–63.
172 *Id.* at 63, 65–66.
173 *Id.* at 61.
174 *Id.* at 62. The Court further noted, “The unpardonable vice of the Roberts test, however, is not its unpredictability, but its demonstrated capacity to admit core testimonial statements that the Confrontation Clause plainly meant to exclude.” *Id.* at 63.
2. Redefining the Confrontation Inquiry: The Testimonial Statement

The Crawford Court relied on the text and history of the Sixth Amendment to create a more vigorous confrontation standard. Justice Scalia began with a textualist analysis that initially appeared designed to address the question of who might be a “witness[] against” the accused. The majority cited to the 1828 edition of Webster’s American Dictionary of the English Language, which defined a “‘witness[’] against the accused” as one who “bear[s] testimony” and defined “testimony” as “a solemn declaration or affirmation made for the purpose of establishing or proving some fact.”

Justice Scalia also examined the history of the Confrontation Clause. Following a lengthy disquisition of the treason trial of Sir Walter Raleigh along with other historical materials, the Court concluded that “the principal evil at which the Confrontation Clause was directed was the civil-law mode of criminal procedure, and particularly its use of ex parte examinations as evidence against the accused.” According to the majority, the text and history together led to the inference that “the Framers would not have allowed the admission of testimonial statements of a witness who did not appear at trial unless he was unavailable to testify, and the defendant had had a prior opportunity for cross-examination.”

Thus, Crawford replaced the Ohio v. Roberts reliability focus with an entirely different inquiry. The Crawford Court concluded that “[t]he constitutional text, like the history underlying the common-law right of confrontation, thus reflects an especially acute concern with a specific type of out-of-court statement.” Future courts attempting to understand the scope of a criminal defendant’s confrontation right would now need to start by distinguishing testimonial statements from non-testimonial statements.

Despite the importance of this new task, the Crawford Court declined the opportunity to define testimonial statements. Instead, Justice Scalia noted that, “[w]hatever else the term covers, it applies at a minimum to prior testimony at a preliminary hearing, before a grand jury, or at a former trial; and to police interrogations.” Crawford also contained dicta that (rather confusingly)

---

177 Id. at 42–44.
178 Id. at 51.
179 Id.
180 Id.
181 Id. at 42–49.
182 Id. at 50.
183 Id. at 53–54 (emphasis added).
184 Id. at 51.
185 Id. at 68 (“We leave for another day any effort to spell out a comprehensive definition of ‘testimonial.’”).
186 Id.
identified, but did not endorse, various possible alternative definitions. The narrowest definition would limit the right to confrontation to statements “contained in formalized testimonial materials, such as affidavits, depositions, prior testimony, or confessions.” Both before and after Crawford, this definition would be a consistent favorite of Justice Thomas, and five years later he would rely on it to provide his critical fifth-vote concurrence in Melendez-Diaz. The two other alternatives construed testimonial statements more broadly. Under the second definition, testimonial statements might also include the functional equivalent of ex parte in-court testimony such as “affidavits, custodial examinations, prior testimony that the defendant was unable to cross-examine, or similar pretrial statements that declarants would reasonably expect to be used prosecutorially.” Finally, under the third possible definition, judges might inquire into whether statements “were made under circumstances which would lead an objective witness reasonably to believe that the statement would be available for use at a later trial.” This final definition provoked Professor Ronald J. Allen to predict that Crawford would create a “spectacle of deciding what is testimonial by the oxymoronic standard of what, objectively speaking, the primary purpose of a government/citizen interaction might be” and that “the Crawford regime will be subject to just as much, if not more, ambiguity as what it replaced.”

B. The Melendez-Diaz v. Massachusetts Decision

1. The Facts

The Melendez-Diaz case arose out of a Boston police investigation based on an informant’s tip that Kmart employee Thomas Wright was engaged in suspicious activity. According to the tipster, Mr. Wright regularly received phone calls at work that were followed immediately by the arrival of a blue sedan. Mr. Wright would enter the sedan and then return to work a short time later. The Boston police set up surveillance outside the Kmart, and following a search of Mr. Wright, found four clear plastic bags of white powder resembling cocaine. Luis
Melendez-Diaz was one of two suspects arrested in the blue sedan. The three men were taken into custody. After they arrived at the police station, the officers discovered that an additional nineteen bags of white powder had been hidden in the back seat of the patrol car.

Melendez-Diaz was charged with distributing and trafficking in cocaine. At his trial, the Commonwealth submitted three “certificates of analysis.” These certificates reported the amount of white powder seized from the defendant and detailed how the powder “[had] been ‘examined with the following results: The substance was found to contain: Cocaine.’” As required by state law, the three certificates had been sworn to before a notary public. The Massachusetts statutory design was quite clear. Certificates containing sworn statements describing the results of laboratory substance analyses provided prima facie evidence of the composition, quality, and weight of the tested substance. Thus, the Commonwealth could, but need not, provide live trial testimony from the lab analyst. At trial, the defendant objected to admission of the certificates as a violation of his confrontation rights as construed by the Supreme Court in Crawford. The defendant’s request was denied by the trial court, and the decision to admit the certificates was affirmed by the Massachusetts Appellate Court. The United States Supreme Court granted certiorari on March 17, 2008. In the wake of Crawford, there was substantial disagreement among the states regarding whether forensic lab reports were testimonial statements. Professor Jennifer Mnookin has explored the post-Crawford cases on expert evidence and confrontation. Her research, which was published three years after Crawford, revealed that courts across the country were using a range of criteria and reaching

---

198 Id.
199 Id.
200 Id.
201 Id.
202 Id. at 2531.
203 Id.
204 Id. (citing MASS. ANN. LAWS ch. 111, § 13 (LexisNexis 2009)).
205 Melendez-Diaz, 129 S. Ct. at 2532 (“[U]nder Massachusetts law the sole purpose of the affidavits was to provide ‘prima facie evidence of the composition, quality, and the net weight’ of the analyzed substance[.]” (quoting MASS. ANN. LAWS ch. 111, § 13)).
206 See id. at 2531.
207 Id.
209 See Jennifer L. Mnookin, Expert Evidence and the Confrontation Clause After Crawford v. Washington, 15 J.L. & POL’Y 791, 795–96 (2007) (“This article focuses on one domain within the post-Crawford universe that has received rather little academic scrutiny: the intersection of expert evidence with the Confrontation Clause.” (citation omitted)).
inconsistent results. Presumably, the Supreme Court granted cert in *Melendez-Diaz* to resolve these ongoing problems. However, as the post-*Melendez-Diaz* cases reveal, *Melendez-Diaz* has not resulted in a more consistent or systematic approach and may actually have added to the confusion.

2. The Post-Crawford Question: Are Forensic Lab Reports Testimonial Statements?

When *Melendez-Diaz* reached the Court in 2009, Justice Scalia wrote for an eclectic plurality that included Justices Stevens, Souter, and Ginsburg. Although the *Melendez-Diaz* plurality characterized the case as an easy question that required nothing more than a “straightforward application of our holding in *Crawford*,” the simplicity of the question is belied by both Justice Thomas’s razor thin concurrence and Justice Kennedy’s broad and vigorous dissent.

(a) Justice Scalia’s Plurality Opinion

In the first few pages of the plurality opinion, Justice Scalia concluded that the certified lab certificates at issue were testimonial statements because: (1) they were affidavits, (2) they were “made under circumstances which would lead an objective witness reasonably to believe that the statement would be available for use at a later trial,” and (3) because “[w]e can safely assume that the analysts were aware of the affidavits’ evidentiary purpose.” Ultimately the first rationale is the most important because it provides the only point of agreement with Justice Thomas, who added the fifth vote. Thus, the holding of *Melendez-Diaz* is rooted in a two-part analysis that starts from the assumption that certified lab reports are akin to affidavits and proceeds to the conclusion that, after *Crawford*, the Confrontation Clause precludes prosecutors from admitting affidavits and other similarly formalized statements because they are testimonial statements.

The *Melendez-Diaz* Court overstates the *Crawford* holding. As discussed above, the only statements that the *Crawford* Court defined as testimonial were “prior testimony at a preliminary hearing, before a grand jury, or at a former trial; and police interrogations,” a list that clearly does not include affidavits. Although Justices Scalia and Thomas were correct that the *Crawford* Court twice

---

210 Cf. id. at 796 (“[A] great many lower court opinions have wrestled with the potential Confrontation Clause implications of expert evidence . . . . Most of these courts have endeavored to find ways around *Crawford*’s dictates . . . .”).

211 *Melendez-Diaz*, 129 S. Ct. at 2530.

212 See id. at 2533.

213 See id. at 2531 (citing *Crawford v. Washington*, 541 U.S. 36, 51 (2004)).

214 *Melendez-Diaz*, 129 S. Ct. at 2531 (quoting *Crawford*, 541 U.S. at 52).

215 *Melendez-Diaz*, 129 S. Ct. at 2532.

216 Id. at 2542.

217 *Crawford v. Washington*, 541 U.S. 36, 68 (2004); see supra Part IV.A.
mentioned “affidavits”; these references were embodied within a discussion of various possible alternative definitions of testimonial statements not specifically embraced by the Court. Thus, the Melendez-Diaz Court did not simply apply Crawford, as the plurality claimed, but expanded the definition of testimonial statements to include documents akin to affidavits.

(b) Justice Thomas’s Concurring Opinion

Justice Thomas concurred that certified lab reports are testimonial statements, but on narrower grounds. In a very short concurring opinion that highlighted the consistency of his own confrontation analysis, Justice Thomas quoted his 1992 concurrence in White v. Illinois to support his position that “the Confrontation Clause is implicated by extrajudicial statements only insofar as they are contained in formalized testimonial materials, such as affidavits, depositions, prior testimony, or confessions.” According to Justice Thomas, the certified lab reports “at issue in this case ‘are quite plainly affidavits’ . . . . As such, they ‘fall within the core class of statements’ governed by the Confrontation Clause.”

(c) Justice Kennedy’s Dissenting Opinion

Justice Kennedy (joined by Chief Justice Roberts and Justices Breyer and Alito) drafted a lengthy and vigorous dissent. According to the dissenters, Melendez-Diaz “sweeps away an accepted rule governing the admission of scientific evidence . . . [that] has been established for at least 90 years . . . based on two recent opinions that say nothing about forensic analysts . . . .” Echoing concerns raised by the Commonwealth, Justice Kennedy articulated a range of practice problems that will inevitably inure to the detriment of the criminal justice system. The dissenters were deeply troubled by the new obstacles that state and federal prosecutors would now confront whenever they seek to introduce forensic evidence. According to Justice Kennedy, by mandating that every criminal defendant has a right to cross-examination of forensic lab analysts, the Court “threatens to disrupt forensic investigations across the country and to put prosecutions nationwide at risk of dismissal based on erratic, all-too-frequent

218 Melendez-Diaz, 129 S. Ct. at 2532 (“There is little doubt that the documents at issue in this case fall within the ‘core class of testimonial statements’ thus described. Our description of that category mentions affidavits twice.”).
219 See id. at 2543 (Thomas, J., concurring).
220 502 U.S. 346, 365 (Thomas, J., concurring).
221 Melendez-Diaz, 129 S. Ct. at 2543 (citing White, 502 U.S. at 365 (Thomas, J., concurring)).
222 Id.
223 Id. (Kennedy, J., dissenting).
224 Id. at 2543–51.
225 Id. at 2543–46, 2556.
instances when a particular laboratory technician, now invested by the Court’s new constitutional designation as the analyst, simply does not or cannot appear.” 226 The dissenters even predicted that the effect of Melendez-Diaz will be that, “in many cases . . . the prosecution cannot meet its burden of proof, and the guilty defendant [will] go[] free on a technicality that, because it results in an acquittal, cannot be reviewed on appeal.” 227 Ultimately, in Justice Kennedy’s view, “[g]uilty defendants will go free, on the most technical grounds, as a direct result of [this] decision, [which] add[s] nothing to the truth-finding process.” 228

Finally, the dissenters attacked the plurality on originalist and historical grounds. According to Justice Kennedy, “[a]ll of the problems with [this] decision . . . would be of no moment if the Constitution did, in fact, require the Court to rule as it does today. But the Constitution does not.” 229 Thus, the dissenters’ most “immediate systemic concern is that the Court makes no attempt to acknowledge the real differences between laboratory analysts who perform scientific tests and other, more conventional witnesses—‘witnesses’ being the word the Framers used in the Confrontation Clause.” 230 The plurality’s jumbling together of expert and non-expert witnesses for confrontation purposes is, in the dissenters’ view, attributable to

[t]he Court’s fundamental mistake[ which] is to read the Confrontation Clause as referring to a kind of out-of-court statement—namely a testimonial statement—that must be excluded from evidence. The Clause does not refer to kinds of statements. Nor does the Clause contain the word ‘testimonial.’ The text, instead, refers to kinds of persons, namely, to ‘witnesses against’ the defendant. 231

Although little has been written about Melendez-Diaz so far, one astute commentator has summarized this originalism debate as follows:

Justices Scalia, Thomas, and Kennedy strive to determine the original meaning of the Confrontation Clause, more specifically the word “witnesses,” but arrive at differing conclusions. Scalia’s version of originalism in Melendez-Diaz is bolder than the others. In his determination to get it right and avoid confusion, however, he downplays contrary historical evidence, serious practical concerns, and the amount of existing authority his rule will overrun. Thomas’s variety of originalism sticks closer to the historical record . . . . Kennedy’s brand of

226 Id. at 2549.
227 Id. at 2550.
228 Id.
229 Id.
230 Id. at 2543.
231 Id. at 2550.
originalism is humbler . . . . Kennedy does not want to throw originalism overboard, but he does not want to go overboard with originalism either.232

Finally, the dissenters assert that not only does the plurality misread the text, but “[n]o historical evidence supports the Court’s conclusion that the Confrontation Clause was understood to extend beyond conventional witnesses to include analysts who conduct scientific tests far removed from the crime and the defendant. Indeed, what little evidence there is contradicts this interpretation.”233

V. MELENDEZ-DIAZ AND THE NAS REPORT: CONSTITUTIONALIZING EVIDENTIARY CONCERNS

Melendez-Diaz redefined the scope of the Confrontation Clause to accommodate a new problem provoked by Crawford—whether out-of-court forensic expert statements proffered by the prosecution raise confrontation concerns.234 The Melendez-Diaz plurality held that a defendant’s confrontation rights are violated when prosecutors introduce forensic lab reports without making the analyst available for cross-examination.235 It is too early to tell whether Melendez-Diaz will be understood as limited to the type of explicitly formalized statements envisioned by Justice Thomas. However Melendez-Diaz is understood and applied by future courts, the decision reflects a significant development in our overall understanding of problems of forensic validity in the criminal courts and a new constitutionalized solution designed to address the “[s]erious deficiencies [that] have been found in the forensic evidence used in criminal trials.”236

The NAS Report figured prominently in the Melendez-Diaz plurality opinion. Justice Scalia described the report as “a recent study conducted under the auspices of the National Academy of Sciences”237 that effectively revealed that “[t]he majority of [laboratories producing forensic evidence] are administered by law enforcement agencies, such as police departments, where the laboratory administrator reports to the head of the agency.”238 According to the plurality, the NAS Report concluded that administrative ties create evidentiary problems “[b]ecause forensic scientists often are driven in their work by a need to answer a particular question related to the issues of a particular case, [and] they sometimes

233 Melendez Diaz, 129 S. Ct at 2552 (Kennedy, J., dissenting).
234 See id. at 2530 (majority opinion).
235 Id. at 2536.
236 Id. at 2537.
237 Id. at 2536.
238 Id. (quoting NATIONAL RESEARCH COUNCIL OF THE NATIONAL ACADEMIES, STRENGTHENING FORENSIC SCIENCE IN THE UNITED STATES: A PATH FORWARD 6-1 (Prepublication Copy Feb. 2009) [hereinafter PREPUBLICATION NAS REPORT].
face pressure to sacrifice appropriate methodology for the sake of expediency.\(^{239}\) Justice Scalia also referred to the NAS Report finding that “[t]he forensic science system, encompassing both research and practice, has serious problems that can only be addressed by a national commitment to overhaul the current structure that supports the forensic science community in this country.”\(^{240}\) In the plurality’s view, this new data has established that it is now far from “evident that what respondent calls ‘neutral scientific testing’ is as neutral or as reliable as respondent suggests.”\(^{241}\)

The plurality also used the NAS Report to illustrate how cross-examination of forensic experts could effectively expose specious and fraudulent evidence.\(^{242}\) The right to confront and cross-examine would enable defendants to uncover and discredit the forensic analyst who had “sacrifice[d] appropriate methodology for the sake of expediency.”\(^{243}\) The plurality confidently predicted that as defendants begin to avail themselves of their new constitutionally guaranteed confrontation rights, “the analyst who provides false results may, under oath in open court, reconsider his false testimony.”\(^{244}\) In Justice Scalia’s view, successful case-specific confrontation may even begin to resolve some of the systemic problems within the forensic science fields identified in the NAS Report because “the prospect of confrontation will deter fraudulent analysis in the first place.”\(^{245}\) Ever mindful of the “serious deficiencies” of much forensic evidence, the Melendez-Diaz plurality provided criminal defendants with the new tool of guaranteed cross-examination “to weed out not only the fraudulent analyst, but the incompetent one as well.”\(^{246}\)

The Melendez-Diaz plurality’s lengthy disquisition of the NAS Report was especially notable because it was neither necessary nor apt—given the facts of the case. It was not necessary if, as the plurality acknowledged, Melendez-Diaz required nothing more than a straightforward application of Crawford.\(^{247}\) If the analysis was that simple, the plurality’s lengthy and detailed detour into the sorry state of the forensic sciences and elaborate speculations about the revelatory effect of future cross-examinations were entirely unnecessary. Moreover, this detour was inapt because the facts of Melendez-Diaz suggest that the defendant would have gained little or nothing had he been afforded an opportunity to cross examine the Commonwealth’s expert at trial. The forensic analyst who prepared this lab report simply ran a chemical analysis of the white powder that had been recovered by the police from the defendant and, based solely on this chemical analysis, determined

\(^{239}\) Id. (quoting PREPUBLICATION NAS REPORT, supra note 238, at S-17).

\(^{240}\) Id. at 2537 (quoting PREPUBLICATION NAS REPORT, supra note 238, at P-1).

\(^{241}\) Id. at 2536.

\(^{242}\) See id. at 2536–37.

\(^{243}\) See id. (quoting PREPUBLICATION NAS REPORT, supra note 238, at S-17).

\(^{244}\) Id. at 2537 (citing Coy v. Iowa, 487 U.S. 1012, 1019 (1988)).

\(^{245}\) See id.

\(^{246}\) See id.

\(^{247}\) See id. at 2542–43.
that the evidence was cocaine.\textsuperscript{248} As the NAS Report repeatedly acknowledged, chemical substance analysis and nuclear DNA testing are the two most consistently valid forensic fields.\textsuperscript{249}

Although the \textit{Melendez-Diaz} plurality cites extensively from the NAS Report, in a footnote, Justice Scalia explained that “[w]e discuss the report only to refute the suggestion that this category of evidence is uniquely reliable and that cross-examination of the analysts would be an empty formalism.”\textsuperscript{250} However, the text of the plurality’s opinion belies this disavowal. First, if the Sixth Amendment clearly mandates confrontation, it cannot matter whether cross-examination is likely to be effective or fruitless; these practice considerations are irrelevant. Second, if the plurality is really only concerned with the rare case where experts claim that their conclusions are “uniquely reliable,”\textsuperscript{251} there would be no need to engage in such a lengthy discussion of the pervasive research and practice problems currently plaguing all of the forensic science fields. Finally, Justice Scalia’s odd and unlikely claim that the need for confrontation would be just as great “if all analysts always possessed the scientific acumen of Mme. Curie and the veracity of Mother Theresa,” further undermines the plurality’s effort to disavow the significance of the NAS Report.\textsuperscript{252}

\section*{VI. The Post-\textit{Melendez-Diaz} Cases: Future Challenges to Forensic Evidence and Experts}

\subsection*{A. Briscoe v. Virginia}

The petition for writ of certiorari to the Supreme Court in \textit{Briscoe v. Virginia}, was filed in May 2008.\textsuperscript{253} \textit{Melendez-Diaz} was decided on June 25, 2009,\textsuperscript{254} and four days later, the Court granted cert in \textit{Briscoe}. The Virginia statute at issue in \textit{Briscoe} provided that certificates of analysis that report the results of crime lab analyses are “admissible in evidence as evidence of the facts therein stated and the results of the analysis or examination referred to therein.”\textsuperscript{255} Under the challenged state statute, the defendant had “the right to call the person performing such analysis or examination or involved in the chain of custody as a witness therein, and examine him in the same manner as if he had been called as an adverse witness. Such witness shall be summoned and appear at the cost of the

\begin{footnotesize}
\begin{enumerate}
\item \textit{Id.} at 2530–31.
\item \textit{See NAS REPORT, supra} note 1, at 7, 40–41, 47, 87, 100–02, 128, 130, 135.
\item \textit{Melendez-Diaz}, 129 S. Ct. at 2537 n.6 (emphasis added).
\item \textit{Id.}
\item \textit{Id.}
\item \textit{Melendez-Diaz}, 129 S. Ct. at 2527.
\item \textit{VA. CODE ANN.} §19.2-187 (2009).
\end{enumerate}
\end{footnotesize}
However, if the defendant failed to request the subpoena, she lost the right to confront and cross-examine the analyst, and the lab report was admitted.257

The question for the Briscoe Court, had been defined by Professor Richard D. Friedman in his petition to the Court as follows: “[i]f a state allows a prosecutor to introduce a certificate of a forensic laboratory analysis, without presenting the testimony of the analyst who prepared the certificate, does the state avoid violating the Confrontation Clause of the Sixth Amendment by providing that the accused has a right to call the analyst as his own witness?”258 The question presented by Briscoe seemed to have been resolved by Melendez-Diaz. In fact, Justice Scalia had explicitly rejected the Commonwealth’s substantially identical claim in Melendez-Diaz that there had been “no Confrontation Clause violation in this case because petitioner had the ability to subpoena the analysts.”259 According to the Melendez-Diaz plurality, the opportunity for the defendant to subpoena the crime lab analyst “whether pursuant to state law or the Compulsory Process Clause—is no substitute for the right of confrontation . . . [because u]nlike the Confrontation Clause, those provisions are of no use to the defendant when the witness is unavailable or simply refuses to appear.”260 However, the Court’s decision granting certiorari in Briscoe raised the possibility that a state statute guaranteeing a defendant the right to subpoena the prosecution’s forensic analyst provided a constitutionally acceptable substitute for confrontation. The Court held arguments in Briscoe on January 11, 2010,261 and approximately two weeks later issued a unanimous, one-sentence per curiam opinion vacating the judgment of the Supreme Court of Virginia and remanding the case for “further proceedings not inconsistent with the opinion in Melendez-Diaz . . . .”262

Over the past year, as the Supreme Court was deciding how to proceed in Briscoe, state courts across the country have been struggling to decipher and apply Melendez-Diaz in a range of criminal cases involving prosecution-proffered forensic lab reports and other state records.263 Although the Melendez-Diaz plurality may have intended to clarify some of the post-Crawford confusion on the nature of testimonial statements, a review of some of the new post-Melendez-Diaz cases reveals that the confrontation doctrine developing in our state criminal courts is utterly inconsistent. In fact, these new cases are so disparate and bizarre that decisions interpreting the scope of the Confrontation Clause have been based on factors that include:

256 Id. § 19.2-187.1.
257 Petition for Writ of Certiorari, supra note 253, at 3–4.
258 Id. (question presented).
259 Melendez-Diaz, 129 S. Ct. at 2540.
260 Id.
261 Transcript of Oral Argument at 1, Briscoe v. Virginia, 130 S. Ct 1316 (No. 07-11191).
262 Briscoe, 130 S. Ct. at 1316.
263 Petition for Writ of Certiorari at 5–10, Briscoe, 130 S. Ct. 1316 (No. 07-11191).
(1) whether a lab analyst subjectively anticipated that his autopsy report would be used in court;\textsuperscript{264}

(2) whether a state requires that forensic reports be certified or accompanied by some other form of attestation;\textsuperscript{265}

(3) whether prosecutors can evade confrontation by asking testifying analysts to describe lab reports prepared by non-testifying analysts (if they do not seek to introduce the non-testifying analyst’s report in evidence);\textsuperscript{266}

(4) whether the test and report were contemporaneous;\textsuperscript{267}

\textsuperscript{264} See People v. Dungo, 98 Cal. Rptr. 3d 702, 710–11 (Cal. Ct. App. 2009) (concluding that an autopsy was a testimonial statement because “the report was prepared during the midst of a homicide, a circumstance of which he was no doubt aware given that a homicide detective . . . was present at the autopsy”).

\textsuperscript{265} See, e.g., id. at 711 (finding that, “[a]s with the certificates at issue in Melendez-Diaz, the autopsy report constitutes a ‘solemn declaration or affirmation made for the purpose of establishing or proving some fact’”); Tabaka v. District of Columbia, 976 A.2d 173, 175–76 (D.C. 2009) (finding a record of the District of Columbia Motor Vehicle Department that had been certified was a testimonial statement under Melendez-Diaz); Grant v. Commonwealth, 682 S.E.2d 84, 89 (Va. Ct. App. 2009) (concluding that certificates attesting to the results of breath tests that, prior to Melendez-Diaz, were frequently admitted in DWI cases were now testimonial statements because “the attestation clause included in the certificate is testimonial in nature and its admission, over the objection of [the defendant], constitutes a violation of the Confrontation Clause”).

\textsuperscript{266} In Dungo, Dr. George Bolduc performed the autopsy and wrote the autopsy report. 98 Cal. Rptr. 3d at 704. Dr. Robert Lawrence, who was not present at the autopsy, testified at trial based on Dr. Bolduc’s report. \textit{Id.} The Dungo Court held that, despite the fact that the prosecutor never sought to admit the autopsy report prepared by Dr. Bolduc, if Dr. Lawrence’s opinions were based on the autopsy report prepared by Dr. Bolduc, the Confrontation Clause required that the defendant have an opportunity to cross-examine Dr. Bolduc. \textit{Id.} at 705. The court reached a different conclusion in People v. Navarro, which involved facts very similar to Melendez-Diaz. People v. Navarro, No. B211266, 2009 WL 2992543, at *3 (Cal. Ct. App. Sep. 21, 2009). The evidence at issue was the forensic analysis of a substance deemed by the analyst to be methamphetamine. \textit{Id.} at *1. In Navarro, the state did not seek to admit the lab report, but relied on the forensic lab analyst’s trial testimony, which was based on a lab test performed by a non-testifying analyst. \textit{Id.} Although the Navarro court concluded that admission of the analyst’s testimony was harmless error, the court opined on the scope of the Confrontation Clause post-Melendez-Diaz, noting that Justice Thomas had limited the scope of the right to formal testimonial materials and that four dissenting justices described the vast difference between witnesses against the accused and “laboratory analysts [who] are not ‘witnesses against’ the defendant as those words have been understood at the framing.” \textit{Id.} at *3 n.4.

\textsuperscript{267} In People v. Gutierrez, the court made a distinction between contemporaneous and near-contemporaneous lab reports. 99 Cal. Rptr. 3d 369, 376–77 (Cal. Ct. App. 2009). According to the Gutierrez court, if reports are prepared “at the time the tests and examinations were conducted, not ‘almost a week after the test were performed,’” the lab reports are non-testimonial. \textit{Id.} at 377.
(5) whether Melendez-Diaz is understood to guarantee confrontation of testimonial statements relating to experts’ methods or conclusions; and
(6) whether expert reports were created as part of a standard lab protocol without any effort to incriminate the defendant; and
(7) efforts to reconcile Melendez-Diaz with Federal Rule of Evidence 703 (or a state corollary), which has long allowed experts to testify based on inadmissible evidence including out-of-court statements by non-testifying witnesses.

268 In Hamilton v. Texas, the prosecutor’s expert, Garon Foster, was a forensic scientist supervisor from the county criminal investigations lab. 300 S.W.3d 14, 19 (Tex. App. 2009). Foster testified to opinions that were based on DNA tests performed by Erica Graham (an analyst who worked in his lab) that revealed that the defendant could not be excluded as the donor of spermatozoa found on the victim. Id. The Hamilton court drew a bright line between the methods and conclusions of scientific inquiry by holding that Foster did not violate the Confrontation Clause when he testified about “the procedures and protocols employed by Graham to produce the DNA profiles Foster used to reach his opinion” if he refrained from mentioning Foster’s conclusions. Id. at 21. Thus, the only portion of Foster’s testimony that raised confrontation concerns was the description of Graham’s findings. Id. at 22.

269 In People v. Navarro, the court found that the Sixth Amendment was not violated because “the report was generated as part of a ‘standardized scientific protocol’ and was made as part of the scope of employment, not as an effort to incriminate the defendant.” Navarro, 2009 WL 2992543, at *2.

270 Federal Rule of Evidence 703 states:

The facts or data in the particular case upon which an expert bases an opinion or inference may be those perceived by or made known to the expert at or before the hearing. If of a type reasonably relied upon by experts in the particular field in forming opinions or inferences upon the subject, the facts or data need not be admissible in evidence in order for the opinion or inference to be admitted. Facts or data that are otherwise inadmissible shall not be disclosed to the jury by the proponent of the opinion or inference unless the court determines that their probative value in assisting the jury to evaluate the expert’s opinion substantially outweighs their prejudicial effect.

FED. R. EVID. 703; see also People v. Rutterschmidt, 98 Cal. Rptr. 3d 390, 412 (Cal. Ct. App. 2009) (“It is well established in this state that expert testimony may ‘be premised on material that is not admitted in evidence so long as it is material of a type that is reasonably relied upon by experts in a particular field in forming their opinions.’”) (quoting People v. Gardeley, 927 P.2d 713, 721 (Cal. 1996)); People v. Johnson, 915 N.E.2d 845, 850 (Ill. App. Ct. 2009) (finding that DNA analyst could testify based on a DNA test that she did not perform without violating defendant’s Confrontation Clause rights because this testimony was consistent with well-established rules “that an expert may testify about the findings and conclusions of a non-testifying expert that he used in forming his opinions”) (quoting People v. Williams, 895 N.E.2d 961, 969 (Ill. App. Ct. 2008)).
These new cases demonstrate the prescience of Professor Allen’s prediction that “the Crawford regime will be subject to just as much, if not more, ambiguity as what it replaces.”271

However, ambiguity is not the only problem with the post-Melendez-Diaz cases. If state legislators and administrative agencies want to deprive future criminal defendants of confrontation opportunities, they can circumvent Melendez-Diaz by simply removing certification and attestation requirements from state records. This would effectively transform these documents from testimonial statements to non-testimonial statements because they would no longer be akin to affidavits. This easy end run around the Sixth Amendment would have the paradoxical effect of making state records both less reliable and more readily admissible.

If this confusion among the state courts persists, and if states find new legislative or administrative routes around Melendez-Diaz, the Supreme Court may reconsider its recent decision not to provide additional clarification. As courts continue to wrestle with Melendez-Diaz, the following questions relating to the appropriate interplay between confrontation and evidentiary challenges to expert evidence are likely to arise. First, is the opportunity to confront experts satisfied if the defendant has an opportunity to cross-examine the prosecutor’s expert during a Daubert hearing? Second, will the Melendez-Diaz emphasis on cross-examination as an effective tool for exposing specious or fraudulent expertise affect how judges understand and apply the evidence rules and/or Daubert? Third, how will courts reconcile the fact that Melendez-Diaz (but not Daubert or any federal or state evidentiary rules) creates new opportunities for criminal defendants to test the validity of forensic evidence unavailable to all other parties? Fourth, does Melendez-Diaz affect the admission of evidence under Federal Rule of Evidence 703 and its state corollaries, which have long allowed experts to explain the bases for their opinions? Fifth, if Melendez-Diaz has already been extended to include testimonial statements by experts in non-forensic fields, what constitutional criteria define the scope and limits of testimonial statements contained within federal, state, or local records? These are just a few of the many questions likely to arise in our courts in the very near future.

CONCLUSION

On February 18, 2009, the NAS Report revealed systemic problems across the range of forensic fields and proposed specific recommendations for the forensic communities and the courts. According to the NAS Report, the Federal Rules of Evidence and Daubert have failed to prevent our criminal courts from “continu[ing] to rely on forensic evidence without fully understanding and addressing the limitations of different forensic science disciplines.”272 Four months

271 Allen, supra note 192, at 14.
272 NAS REPORT, supra note 1, at 53.
later, the Supreme Court expanded the criminal defendant’s right to confrontation, citing the “[s]erious deficiencies [that] have been found in the forensic evidence used in criminal trials.” Together, the NAS Report and Melendez-Diaz raise important questions about how future courts should test the validity of proffered forensic evidence. For decades, and especially since Daubert was decided in 1993, these have traditionally been viewed as evidence questions and addressed with evidentiary rules and standards. However, the NAS Report and Melendez-Diaz offer new and different solutions. The NAS Report recommendations begin with an ambitious plan to centralize and coordinate the fields and to improve funding and support for legitimate forensic research. These extrajudicial approaches require a substantial national commitment of time and money and the creation of new programs designed to improve oversight, increase standardization, and enhance interdisciplinary coordination. Melendez-Diaz opts for a more immediate constitutional solution designed to neutralize the impact of some forensic evidence by guaranteeing defendants the right to cross-examine prosecutors’ experts. This recent expansion of the Confrontation Clause purports to provide defendants with the power to expose the forensic analyst who “sacrifice[d] appropriate methodology for the sake of expediency” and “the analyst who provides false results [and] may, under oath in open court, reconsider his false testimony.”

Together, the NAS Report and Melendez-Diaz chart a new path forward that does not end at the Daubert destination of more accurate judicial screening. Instead, this path begins with constitutionally guaranteed confrontation of prosecutors’ forensic experts in court and ends with the hope that by implementing some or all of the NAS Report extrajudicial recommendations (maybe) there will finally be some science in the forensic sciences.

---

274 NAS REPORT, supra note 1, at 77–83.
275 Melendez-Diaz, 129 S. Ct. at 2550.
276 NAS REPORT, supra note 1, at 24.
277 Melendez-Diaz, 129 S. Ct. at 2537.
INTRODUCTION

The report of the National Academy of Sciences on the current state of forensic science in the United States is a remarkable document. In comprehensive detail, it exposes a number of troubling and sometimes scandalous problems with forensic science evidence. Some of these problems are economic: for example, many crime labs are severely underfunded and face large backlogs of cases. Some of the problems are political: many forensic experts operate as an arm of law enforcement or the prosecution in a manner that creates potential biases and prejudices, rather than as independent scientists. Some of the problems are legal: despite evidentiary rules that are meant to weed out expert testimony that has not been shown to be reliable, courts do not appear to be requiring that all forensic expertise meet this standard. Finally, many of the problems are with the forensic science itself: with the exception of DNA evidence, there is a lack of peer-reviewed studies or other credible evidence establishing the scientific validity of most forensic science; nor is there much evidence demonstrating exactly how reliable many forensic techniques are in practice. And there is widespread divergence regarding the types of training, certification, and controls on the methodologies and protocols of those who process and testify about forensic-science evidence. According to the NAS Report, the current state of forensic

© 2010 Michael S. Pardo, Associate Professor, University of Alabama School of Law. My thanks to Daniel Medwed for inviting me to participate in this symposium and to Dean Ken Randall and the Alabama Law School Foundation for generous research support.


2 Id. at 1-2, 2-3 to -5.

3 See id. at S-6 n.8, 2-19.

4 Id. at 3-17. Moreover, there are additional problems with the manner in which forensic science is presented in court. These problems include the conclusory manner in which testimony is often presented and the certainty with which conclusions are given. See, e.g., id. at 3-15 to -16.

5 Id. at S-5 to -6 (“[T]here is a notable dearth of peer-reviewed, published studies establishing the scientific bases and validity of many forensic methods.”).

6 See, e.g., id. at 3-16. These are distinct problems. One problem has to do with the quality of the evidence; the other has to do with our knowledge of the quality of the evidence. Even evidence of not great quality may be useful in litigation settings so long as it is probative and the fact-finder understands how much weight to give it. The primary problem with many forensic-science techniques appears not necessarily to be that they are unreliable; rather, the problem is that their reliability is uncertain.

7 Id. at S-4 to -6.
science poses a threat to effective, efficient, and accurate law-enforcement investigation (including national security) as well as to innocent defendants who may be wrongfully convicted on the basis of such evidence.\(^8\)

In addition to documenting these problems, the NAS Report calls for widespread reform. One common theme running through the report is that legal doctrine—primarily, the law of evidence—cannot fix the problems diagnosed in the report and that reform must focus on the forensic science practices themselves and the juridical evidence they produce.\(^9\) These reforms include creating an independent agency to oversee forensic science (the “National Institute of Forensic Science”); establishing best practices; establishing standardized terminology, measurements, and reporting procedures; developing accreditation procedures and ethical codes; increasing funding for education and training; and conducting research into the scientific foundations of the various forensic techniques and research into the potential biases and prejudices of practitioners, among others.\(^10\)

Other participants in this symposium are better qualified than I am to assess the feasibility as well as the likely successes and failures of these recommendations on their own terms, and I will leave this analysis to them. My focus will be on the relationships between current theoretical accounts of evidence (and the proof process more generally) and current forensic science evidence as described by the NAS Report. My hope is that the interaction between high-level evidence theory and the intensely practical issues raised by the NAS Report will help to illuminate both.

Aside from any theoretical insights this interaction may provide, its practical significance should not be underestimated. Although the NAS Report is surely right that legal doctrine cannot solve all of the problems identified, legal doctrine and the courts applying it need not sit idly by, replicating the status quo, waiting for forensic science reforms to take hold. The justification of doctrinal changes or applications by courts will depend upon a well-justified conceptual foundation of the legal proof process.\(^11\) Providing this conceptual foundation is the domain of

---

\(^8\) *Id.* at S-3 (“Further advances in the forensic science disciplines will serve three important purposes. First, further improvements will assist law enforcement officials . . . to identify perpetrators with higher reliability. Second, further improvements . . . should reduce the occurrence of wrongful convictions . . . . Third, any improvements . . . will undoubtedly enhance the Nation’s ability to address the needs of homeland security.”). The NAS Report also documents a number of problems with, and suggests reforms for, medical examiner and coroner offices throughout the country. *See id.* at 9-1 to -21.

\(^9\) *Id.* at 1-14 (“In short, judicial review, by itself, is not the answer.”).

\(^10\) *Id.* at S-14 to -24.

\(^11\) This justification will also depend on accurate empirical information. Any theoretical or conceptual project in this area must, of course, be informed by the best available empirical evidence of the proof process. On the relationship between theoretical and empirical investigations in the law of evidence, see generally Ronald J. Allen & Brian Leiter, *Naturalized Epistemology and the Law of Evidence*, 87 Va. L. Rev. 1491, 1503 (2001) (arguing that “[e]xisting attempts to make theoretical sense of the evidentiary process have inadequately attended to their empirical adequacy and, in addition, have fairly systematically run afoul of the two constraints of epistemology”).
Evidence theory. Reflecting on forensic science in light of evidence theory will thus help to illustrate potential options for doctrinal changes and potential applications of current doctrine, as well as illustrate potential ways in which these changes and applications may be justified in light of the goals and functions of the proof process. Or so I will attempt to demonstrate.

Rather than focus on the details of any one type of forensic science, I will focus on a general problem underlying many forensic techniques: evidence purporting to link a defendant to a particular crime scene when there is uncertainty regarding the reliability of technique to establish the link or uncertainty regarding the validity of the science underlying the technique. Given the way this problem is framed, I will focus most of my analysis on the use of forensic evidence by the prosecution as purportedly inculpatory evidence, but, aside from a few differences I will point out along the way, many of the considerations will also apply to defense evidence. My primary conclusion is that, in light of evidence theory, many of the problems with forensic science described in the NAS Report ought to be dealt with at the sufficiency-of-the-evidence stage rather than at the admissibility stage. Doing so will require courts to develop a more robust sufficiency jurisprudence for criminal cases. I will briefly sketch how such a development may proceed in light of the theoretical issues discussed.

Part I provides a brief outline of evidence theory. Part II discusses forensic science in light of the theoretical issues discussed in Part I, focusing on the general problem of forensic evidence of unknown probative value purporting to link a defendant to a crime. Part III argues that a sufficiency response may be better justified than an admissibility response and sketches how such a response may proceed in a theoretically justified manner.

I. **A BRIEF OUTLINE OF EVIDENCE THEORY**

Theoretical accounts in the law of evidence are primarily epistemological in nature. They are “epistemological” in the broad sense that they aim to either justify or to reform evidentiary rules or practices in light of their tendencies to produce true (factually accurate) outcomes or produce false (factually erroneous) outcomes. These accounts are also “epistemological” in the more narrow sense that the truth-conducive or truth-thwarting tendencies of evidentiary rules are typically evaluated based on the likely effects they will have on the rational evaluation of evidence by juries and judges. In short, legal fact finders are epistemic agents whose inferences and conclusions about evidence may be more or less justified and which the law may or may not endorse based upon the epistemic warrant of the inferences and conclusions. Although epistemic considerations largely drive evidence policy, the choices among evidentiary arrangements (with varying

---

epistemic consequences) must ultimately be defended in terms of political morality. Not all evidence theory is concerned with these questions about truth, errors, rationality, and epistemic justification, but a great deal of it is (at least implicitly), and it is the part on which I shall focus and to which I shall refer as “evidence theory.” Although evidentiary arrangements may at times be justified on other grounds, these considerations will typically provide necessary, if not sufficient, conditions for just legal judgments.

The law of evidence and the theory underlying it serve these epistemic functions by regulating the proof process at both the macro level and the micro level. The macro level involves evaluations of whether evidence as a whole proves—or is at least sufficient for a reasonable fact finder to conclude that it proves—the elements of a claim or an affirmative defense. The law operates at this level by assigning the burden of proof to one party and adopting a decision rule (such as “preponderance of the evidence,” “clear and convincing,” or “beyond a reasonable doubt”). The burden of proof and the decision rules function to distribute the potential errors between the parties in a justified manner. For example, the preponderance-of-the-evidence rule applicable in most civil cases is meant to distribute the risk of error roughly evenly among the parties, while the beyond-a-reasonable-doubt rule in criminal cases is meant to skew the risk of error against the prosecution and in favor of criminal defendants.

---

13 See Michael S. Pardo, The Political Morality of Evidence Law, 5 INT’L COMMENTARY ON EVIDENCE 1, 16–30 (2007), available at http://www.bepress.com/ice/vol5/iss2/art1 (reviewing ALEX STEIN, FOUNDATIONS OF EVIDENCE LAW (2005)) (arguing that “morally justified evidence law . . . lies . . . in deeper epistemic waters”). Evidentiary arrangements that are not the best from an epistemic perspective may nevertheless be better justified in terms of political morality. For example, the beyond-a-reasonable-doubt standard may create more overall errors than the preponderance standard, but it is justified in criminal cases based on the relative costs of false acquittals and false convictions. In general, however, the practices better justified epistemically will also be the ones better justified in terms of political morality. This is so because truth is a necessary condition for the application of justice. See WILLIAM TWINING, RETHINKING EVIDENCE: EXPLORATORY ESSAYS 76 (2d ed. 2006) (“Establishing the truth . . . is a necessary condition for achieving justice in adjudication; incorrect results are one form of injustice.”).

14 See TWINING, supra note 13.

15 See, e.g., Grogan v. Garner, 498 U.S. 279, 286 (1991) (noting “the preponderance-of-the-evidence standard results in a roughly equal allocation of the risk of error between litigants”); Herman & Maclean v. Huddleston, 459 U.S. 375, 390 (1983) (quoting Addington v. Texas, 441 U.S. 418, 423 (1979)) (“A preponderance-of-the-evidence standard allows both parties to ‘share the risk of error in roughly equal fashion.’”); In re Winship, 397 U.S. 358, 371 (1970) (Harlan, J., concurring) (“In a civil suit between two private parties for money damages, for example, we view it as no more serious in general for there to be an erroneous verdict in the defendant’s favor than for there to be an erroneous verdict in the plaintiff’s favor.”).

16 See, e.g., Addington v. Texas, 441 U.S. 418, 423–24 (1979) (“In the administration of criminal justice, our society imposes almost the entire risk of error upon itself. This is
One possibility at the macro level is for the law of evidence not to regulate it. In other words, this would allow fact finders in each case to decide for themselves when the evidence warrants a conclusion that a disputed fact has been proven. This would allow them, in effect, to distribute the risk of error among the parties as they see fit. The Supreme Court has rejected this approach as a general matter and has declared in particular that the decision rule in criminal cases of “beyond a reasonable doubt” (BARD) is constitutionally required. Although the decision rules potentially help to solve one problem—distributing the risk of error in a socially desirable manner—they may achieve this goal while still leaving much to be desired epistemically. Consider two extreme examples: the risk of error under the preponderance rule could be approximated in civil cases by deciding cases with a coin flip, and if BARD is interpreted to incorporate something like Blackstone’s ratio (i.e., ten false acquittals for every false conviction), then this risk of error could be approximated with an eleven-sided die or any other random procedure that made it ten times as likely the defendant would win.

The upshot of the examples is that the number and types of accurate outcomes matter, too. The decision rules distribute the risk of error based on a rational assessment of the evidence because a rational assessment of the evidence will produce more accurate outcomes. What is troubling about the examples is that they prevent the parties who ought to win from producing evidence showing they ought to win, thereby reducing the risk of an erroneous and adverse outcome. By producing favorable evidence, each side may reduce the risk of an adverse judgment, and fact finders have more information on which to decide given the residual uncertainty. In short, the epistemic focus on the proof process is not just on the decision rules but on the evidentiary base upon which decisions are made in light of these rules.

accomplished by requiring under the Due Process Clause that the state prove the guilt of an accused beyond a reasonable doubt.

17 See Santosky v. Kramer, 455 U.S. 745, 757 (1982) (noting that the Supreme Court “has never approved case-by-case determination of the proper standard of proof for a given proceeding” and finding that “[s]ince the litigants and the fact-finder must know at the outset of a given proceeding how the risk of error will be allocated, the standard of proof necessarily must be calibrated in advance”).

18 See, e.g., In re Winship, 397 U.S. at 364 (noting that “the reasonable-doubt standard is indispensable” to due process).

19 4 WILLIAM BLACKSTONE, COMMENTARIES *352 (“[I]t is better that ten guilty persons escape, than that one innocent suffer.”).


21 The evidence provided by the party who does not deserve to win, given the underlying events, may be misleading or truth-thwarting evidence.
This takes us, finally, to the micro level of the law of evidence. One possibility at this level is to allow each party—who presumably best know their respective cases—to present whatever evidence they wish. One constraint on such a free-proof system, however, is the requirement that evidence must be relevant to the disputed issues. It is hard to argue against such a requirement given that logically irrelevant evidence ought, by definition, to have no bearing on a rational evaluation of whether a fact has been proven. Moreover, some parties under a free-proof system may have an incentive to introduce irrelevant evidence to confuse the issues and/or stall the proceedings.

At the micro level, from an epistemic perspective, the more relevant evidence the better.\textsuperscript{22} In general, evidence is relevant if it makes a disputed fact appear more or less likely.\textsuperscript{23} The more of it presented to a fact finder, the better the evidentiary base on which legal decisions will be made. Thus, as a matter of evidence policy, the exclusion of relevant evidence requires a separate justification. These justifications fall into four categories: (1) economic, (2) other policy goals, (3) jury control, (4) party control. First, relevant evidence may have only minimal probative value in proving the disputed fact, or may be cumulative of other evidence, or may otherwise not be worth the costs of receiving it.\textsuperscript{24} Second, relevant evidence may sometimes be excluded to serve other non-epistemic purposes.\textsuperscript{25} The two other categories are meant to provide epistemic rationales for the exclusion of relevant evidence. Third, relevant evidence may be excluded if fact finders will “overvalue” the evidence to such an extent that it will detract from rather than aid a rational evaluation of the evidence as a whole, or will otherwise distract them from this function.\textsuperscript{26} Finally, relevant evidence may be excluded if doing so will induce parties to present better evidence instead.\textsuperscript{27}

The foregoing is relatively uncontroversial. Less understood, and perhaps more controversial, is how the macro-level and micro-level issues interact. Two related issues will highlight the complexity. First, in addition to attempting to improve the evidentiary base on which decisions are made, micro-level rules may also shift the risk of error between the parties (perhaps frustrating or amplifying

\textsuperscript{22} See TIMOTHY WILLIAMSON, KNOWLEDGE AND ITS LIMITS 189 (2000) (discussing this epistemic principle); Alvin I. Goldman, \textit{Quasi-Objective Bayesianism and Legal Evidence}, 42 JURIMETRICS J. 237, 253–60 (2002). This is not to deny that it may be quite difficult to determine whether particular evidence is relevant or not in the context of a particular case.

\textsuperscript{23} FED. R. EVID. 401.

\textsuperscript{24} FED. R. EVID. 402–03.

\textsuperscript{25} FED. R. EVID. 407–11.

\textsuperscript{26} FED. R. EVID. 403.

\textsuperscript{27} See FED. R. EVID. 1001–1008 (requiring original evidence unless other conditions of trustworthiness are satisfied); see also Dale A. Nance, \textit{Naturalized Epistemology and a Critique of Evidence Theory}, 87 VA. L. REV. 1551, 1555–56 (2001) (discussing the role of exclusionary rules in incentivizing parties to present better evidence in court).
the risks imposed by the macro-level proof rules). For example, a rule that excludes evidence that would otherwise typically be presented by one side (e.g., the prosecution) may shift more risk onto that side and away from the other side (e.g., defendants), while a rule routinely admitting that evidence may shift the risk of error in the other direction. Second, given this interaction, an asymmetrical application of micro-level rules may be justified by the fact that, as an empirical matter, a decision rule is failing to achieve its desired effect in distributing errors among the parties. For example, if BARD is known to be producing too many false convictions (as compared to false acquittals), then an asymmetrical micro-level rule (either of admission or exclusion) that shifts more risk away from defendants and onto the prosecution may be justified on that ground—and vice versa if the reverse were true. A priori there is no reason to prefer one evidentiary arrangement over another in terms of where it distributes the risk of error (micro and macro), but evidence theory must attend to the ways these levels interact.

This sketch of the basic structure of the evidentiary proof process still leaves open a number of conceptual issues. Most significantly, we still need some conception of what makes evidence relevant, how to measure probative value, and when evidence as a whole is sufficient to satisfy the decision rules. Within evidence theory, two competing conceptions speak to these issues. The first is a probabilistic conception. Under this conception, relevance and probative value can be evaluated based on the likelihood of a disputed fact of consequence given the particular item of evidence, typically measured as a cardinal probability between zero and one. Under this conception, decision rules are typically


30 See id.; see also Richard O. Lempert, Modeling Relevance, 75 Mich. L. Rev. 1021, 1025–26 (1977) (explaining how the Bayes’ Theorem and regret matrices may be used to analyze evidentiary rules). Under this conception, the probative value of evidence may be expressed as a “likelihood ratio,” that is, the likelihood of receiving the evidence given that the disputed fact for which it is offered is true compared with the likelihood of receiving the evidence given that the disputed fact is false. Id.
assigned cardinal probabilities (e.g., 0.5 for preponderance and 0.91 for BARD), and evidence is sufficient to satisfy the decision rule when the probative value of the evidence as a whole surpasses the decision rule.  

The second conception is explanatory. 33 Under this conception, relevance and probative value may be evaluated based on whether and how well a disputed fact, if true, would explain particular items of evidence. 34 Likewise, under this conception, decision rules may be explicated in terms of how well each side’s theory explains the evidence. 35 For example, consistent with the error-distribution goals of the decision rules, the preponderance rule is satisfied when the best available explanation of the evidence and events under dispute supports the party with the burden of proof. 36 Likewise, the BARD rule is satisfied when the model of probative value, see Ronald J. Allen & Michael S. Pardo, The Problematic Value of Mathematical Models of Evidence, 36 J. LEGAL STUD. 107, 111–14 (2007).


34 This conception relies on the notion of “inference to the best explanation” (or the inferential process of abduction), a notion best known in the philosophy of science. See Gilbert H. Harman, The Inference to the Best Explanation, 74 PHIL. REV. 88, 88–91 (1965); see generally Peter Lipton, Inference to the Best Explanation 1 (2d ed. 2004) (providing an explanation of the model of Inference to the Best Explanation); Paul R. Thagard, Evaluating Explanations in Law, Science, and Everyday Life, 15 CURRENT DIRECTIONS PSYCHOL. SCI. 141, 141–43 (2006) (discussing a “theory of explanatory coherence” to explain how people consider competing explanations). A number of general criteria determine whether one explanation is better than another (e.g., consistency, simplicity, coherence with background beliefs, consilience, and the absence of ad hoc premises, and so on), but this determination will depend on the context and details of particular cases as well as on the goals of decision makers. For further discussion, see Pardo & Allen, Juridical Proof and the Best Explanation, supra note 33, at 229–33.

35 See Pardo, supra note 33, at 1102–05; Pardo & Allen, Juridical Proof and the Best Explanation, supra note 33, at 233–42 (explaining how explanatory proof works at trial).

36 Assuming that the better of two explanations is more likely to be true, then this rule will serve the error-allocation and -distribution functions underlying the preponderance rule.
prosecution offers a plausible explanation of guilt and there is no plausible explanation consistent with innocence.\(^{37}\)

It should be noted that these conceptions, and the conceptual issues in evidence theory more generally, are related to—but distinct from—the empirical questions of how jurors actually draw inferences and decide cases.\(^{38}\) The best model regarding these empirical questions is the “story model,” where jurors impose a narrative structure on the evidence, organize the evidence into coherent versions of events through their background knowledge about analogous situations, generalize about the world in the general, and make assumptions about gaps in the evidence.\(^{39}\)

Finally, however, it should be noted that this empirical information by itself cannot answer the conceptual questions posed by evidence theory. For example, the story model cannot tell us when evidence is sufficient to satisfy a decision rule, when a jury verdict is unreasonable, or how to measure probative value.\(^{40}\) Evidence theory also places normative constraints on the empirical process: for example, a process of selecting the best story among those offered by each side in a criminal case would allocate an unjustified risk of error onto defendants.\(^{41}\)

II. FORENSIC SCIENCE AND EVIDENCE THEORY

The NAS Report documents several problems with the current state of forensic science. For purposes of current criminal litigation, the most serious better than a probabilistic rule. See Pardo & Allen, *Juridical Proof and the Best Explanation*, supra note 33, at 235, 261–62.

37 Assuming the quality of an explanation is a good indicator of its likely truth, then this rule will serve the error-allocation and -distribution functions underlying the BARD rule. See Pardo, supra note 33, at 1105; Ronald J. Allen & Michael S. Pardo, *Probability, Explanation, and Inference: A Reply*, 11 Int’l J. Evid. & Proof 307, 316–17 (2007).

38 Theoretical accounts must be empirically informed. See Allen & Leiter, supra note 11, at 1503–37.


40 For example, if the probative value of evidence were just what any individual jury thought it was, then this would render Federal Rule of Evidence 403 largely unintelligible (how could evidence be misleading or confusing or unfairly prejudicial if the only relevant criterion for probative value is what jurors think it is?). Moreover, sufficiency-of-the-evidence reviews would not make sense because it would be up to juries to determine for themselves what is sufficient. In other words, summary judgment, judgments as a matter of law in civil cases, and sufficiency challenges in criminal cases, which all turn on what “reasonable” juries could conclude, would not be possible. See Pardo, supra note 33, at 1097–99.

41 For further discussion of this point, see id. at 1102–05.
general issue appears to be the following: “With the exception of nuclear DNA analysis . . . no forensic method has been rigorously shown to have the capacity to consistently, and with a high a degree of certainty, demonstrate a connection between evidence and a specific individual or source.”

We may call this the “identity” problem, and it underlies many of the uses of forensic science evidence in criminal proceedings: fingerprints, hair, bite marks, ballistics, handwriting, shoe prints, and tire tracks, among others. We may state the identity problem generally and schematically in the following way: we have some evidence (E) and some forensic science technique (FS) such that the combination of E and FS renders it more or less likely that the defendant is the source of E (and thus more or less likely that the defendant is guilty). The reason the identity problem is a problem is because in most cases we simply do not know how much more or less likely the combination of E and FS makes it that the defendant is the source.

The NAS Report responds to this problem by proposing recommendations that aim at improving the precision and reliability of the techniques as well as discovering more information about the precision and reliability of the techniques. The report states that improving these techniques will improve the ability of law enforcement to identify true perpetrators of crime and “should reduce the occurrence of wrongful convictions.” The report also notes that the probative value and hence the admissibility of forensic science evidence depends on the precision and reliability of these techniques, suggesting, consistent with Daubert and Federal Rule of Evidence 702, that evidence must cross some “threshold of evidentiary value” for it to be admitted into court.

Faced with the world described by the NAS Report, courts in criminal cases have two different avenues for potential doctrinal responses to the coming challenges to forensic science. One avenue, and the one most discussed in the report, is at the micro level. The second avenue is at the macro level.

At the micro level, one response would be to exclude the evidence unless and until the precision and reliability of the techniques have been demonstrated to a sufficient degree. This could be done under current doctrine under a more rigorous (and perhaps more faithful) application of Daubert and Federal Rule of Evidence 702, or perhaps under Federal Rule of Evidence 403. This response, however,

---

42 NAS REPORT, supra note 1, at S-5.
43 Id. at S-10 to -24.
44 Id. at S-3.
45 Id. at ch. 3.
46 Id. at 5-37.
48 Daubert v. Merrell Dow Pharmaceuticals, Inc., 509 U.S. 579, 589–95 (1993) (imposing “reliability” standard on the admissibility of expert testimony, and providing criteria for assessing “reliability” such as whether a technique can be tested, has been
Evidence Theory and the NAS Report

raises a number of potential problems in light of evidence theory. First, let us begin with relevance. Only if a forensic technique is no better than a coin flip is it irrelevant. Assuming this is not the case (at least with regard to most of these techniques), then the evidence should presumptively be admitted unless there is some other reason to exclude it. The two possible epistemic reasons for doing so are because the fact finder will significantly overvalue it (jury control) or because it will induce the proponent to introduce better evidence (party control).

Jury-control considerations may be particularly salient in this context given the aura of reliability (and perhaps infallibility) surrounding science in our culture. But, as documented in the NAS Report, much of the problem here may have more to do with how forensic experts testify and to what they testify, rather than the fact that this type of evidence is introduced. Moreover, although far from conclusive, empirical evidence appears to support the competence of jurors in evaluating complex evidence. Thus, there are reasons to think (at least until there is good empirical evidence to the contrary) that jurors, once apprised of the many known and unknown limitations of the techniques, can give it proper weight in the subjected to peer-reviewed study, has known error rates, has standards controlling its operations, and is accepted in the relevant scientific community; see also Kuhmo Tire Co., Ltd. v. Carmichael, 526 U.S. 137, 141 (1999) (holding that Daubert also applies to non-scientific expert testimony); Gen. Elec. Co. v. Joiner, 522 U.S. 136, 146 (1997) (concluding that courts can assess expert conclusions as well as techniques under Daubert); FED. R. EVID. 702 (requiring that expert testimony must be based on sufficient facts or data along with reliable principles and methods, applied reliably to the facts of the case). Despite previously admitting forensic science, jurisdictions that employ the “general acceptance” standard, may also attempt to screen such evidence by concluding that it is no longer generally accepted in the scientific community, by redefining the relevant scientific community, or by concluding that it would not be helpful to the fact-finder. See Frye v. United States, 293 F. 1013, 1014 (D.C. Cir. 1923).

If a technique is no better than random, then it should be excluded on relevance grounds. Many forensic techniques purport to provide relevant evidence based on the fact that they rely on similarities in patterns between evidence found at crime scenes and evidence related to defendants. See Michael J. Saks & Jonathan J. Koehler, The Coming Paradigm Shift in Forensic Identification Science, 309 SCIENCE 892, 892 (2005).

Even if minimally relevant, the evidence may also be excluded based on the excessive costs of producing it. FED. R. EVID. 403.

See NAS REPORT, supra note 1, at 1-11. The NAS Report’s discussion of the education of potential experts also provides a nice model for the kinds of information an expert should be able to explain to judges and juries: “the [expert] should learn what to measure, the associated population statistics (if appropriate), biases and errors to avoid, other threats to the validity of the evidence, how to calculate the probability that a conclusion is valid, and how to document and report the analysis.” Id. at 8-1.

For a survey of the literature, see VIDMAR & HANS, supra note 39, at 177–80.

This knowledge may come from either explaining the technique and why its results are likely to be valid or through evidence indicating that the technique has been tested empirically and has been sufficiently validated (even if the expert cannot explain why it is valid). For an illuminating discussion of this distinction, see Jennifer L. Mnookin, Of Black Boxes, Instruments, and Experts: Testing the Validity of Forensic Science, EPISODEME 343
context of individual cases. And from an epistemic perspective, other things being equal, the more relevant evidence—generally understood and roughly properly weighed—the better. Likewise, the party-control rationale does not necessarily warrant exclusion. The exclusion of much current forensic science evidence may induce the development of more precise and reliable techniques. But, and the NAS Report documents, there are plenty of reasons and incentives to continue with this development regardless of whether current evidence is excluded. And exclusion may produce great epistemic costs by excluding probative evidence (assuming it could be accurately presented and evaluated).

Nor is it clear that systematically excluding non-DNA forensic science would reduce erroneous judgments as a whole or types of errors (including false convictions). We do know that in a large percentage of known wrongful convictions the prosecution introduced forensic science evidence against the defendant. But we do not know the extent to which this evidence caused these verdicts. More importantly, it is not clear what the systemic consequences might be should this evidence be routinely excluded. A number of perverse possibilities exist. Perhaps prosecutors may try to convict the same number defendants with worse evidence—worse epistemically but better persuasively—perhaps leading to more false convictions. Or perhaps prosecutors will select and prosecute a different class of defendants based on a greater perceived likelihood of conviction but, again, based on worse evidence, perhaps also increasing false convictions. And, of course, it may also lead to a great deal of more false acquittals, placing dangerous criminals back on the street. Assuming appropriate constraints can be placed on how the evidence is presented and that the evidence is generally understood by fact finders, the epistemic considerations underlying evidence exclusion are important. As a general matter, these limitations may be more apparent to juries and judges when expert testimony serves an educational function, rather than one of pure deference. See Ronald J. Allen & Joseph S. Miller, The Common Law Theory of Experts: Deference or Education? 87 NW. U. L. REV. 1131, 1133–41 (1993).

These include, for example, more accurate and efficient law-enforcement investigations and national-security interests. For example, according to a study of 200 wrongful convictions, some type of forensic-science evidence was introduced by the prosecution in 113 of the cases. See Brandon L. Garrett, Judging Innocence, 108 COLUM. L. REV. 55, 81 (2008).

For example, we do not know how many juries would have convicted anyway based on the remaining evidence. And, even if the evidence did play a causal role, we do not know the answer to the counter-factual question of what evidence the prosecution might have introduced instead if the forensic evidence were excluded.

As a general matter, the exclusion of even weak scientific evidence may lead to adverse epistemic consequences if non-scientific evidence relied on in its place is worse. See Schauer, supra note 30, at 15–25.

This possibility arises from the fact that exclusion may further shift the risk of error onto the prosecution. And, given the utilities of true verdicts (acquittals and convictions), these costs cannot be ignored by simply focusing on the ratio of errors. See supra note 20 and accompanying text.

See supra notes 51, 53 and accompanying text.
theory appear to counsel against the systematic exclusion of the evidence. This, of course, is not to deny the very real dangers presented by placing this problematic evidence before juries or judges. Daubert itself mentions “traditional and appropriate means” to deal with the dangers of “shaky but admissible evidence”—“[v]igorous cross-examination, presentation of contrary evidence, and careful instructions on the burden of proof.” The adversarial process itself, however, as the NAS Report notes, is not enough. But it is also not the only other doctrinal option. More detailed macro-level doctrine is also possible. The idea of “careful instructions on the burden of proof” points in the direction of such macro-level considerations but does not go far enough. Better jury instructions on BARD in general or connecting the forensic science with BARD in particular would be desirable.

Also, more importantly, the development of a more robust sufficiency-of-the-evidence jurisprudence would help considerably to cabin the dangers posed by current forensic science evidence. In cases in which otherwise weak evidence is coupled with weak or otherwise problematic forensic science evidence, courts should dismiss cases based on insufficient evidence or overturn convictions. This option would also, for reasons explored more fully in the next Part, respond to the dangers with forensic science in a manner more consistent with the epistemic goals of the proof process.

III. FORENSIC SCIENCE, BARD, AND SUFFICIENCY OF THE EVIDENCE

As a matter of doctrine, the relevant standards regarding burdens of proof, decision rules, and sufficiency of evidence are already in place and of constitutional significance. The prosecution must prove each element of the offense beyond a reasonable doubt, and due process requires that every conviction must be supported by sufficient evidence such that “after viewing the evidence in the light most favorable to the prosecution, any rational trier of fact could have found the essential elements of the crime beyond a reasonable doubt.” In theory, the combination of these standards provides a solid doctrinal foundation to respond to the dangers created by “shaky but otherwise admissible” forensic science. Moreover, this foundation provides a potential response that better accords with the epistemic goals elucidated by evidence theory.

---

62 See Richard D. Friedman, Squeezing Daubert out of the Picture, 33 SETON HALL L. REV. 1047, 1048 (2003) (arguing that sufficiency standards may provide a more appropriate response to problems with expert testimony than admissibility standards).
These goals include reducing errors (or types of errors) and allocating the risk of error among the parties. Assuming that forensic science evidence is relevant and properly understood and evaluated, macro-level rules will better align decision-making with these goals than a micro-level response. At the macro level, in any given case, three situations are possible: (1) the non-forensic-science evidence is sufficient to support a conviction; (2) the forensic science evidence in combination with the other evidence is insufficient; or (3) the forensic evidence pushes the totality of the prosecution evidence across the sufficiency line. In each situation, the forensic evidence improves the evidentiary base upon which decisions are made. In the second situation, even in the case of a conviction, problematic forensic evidence will be neutralized by overturning the verdict.

In the first situation, a conviction was already warranted anyway, and so problematic forensic evidence did no harm—indeed, it improved the evidentiary base, perhaps solidifying an already epistemically warranted judgment. The third situation is the most important and controversial. But if the macro-level rules are well designed in allocating the risk of error, then admission of the evidence is appropriate because it will allow for more accurate applications of the macro-level rules. In other words, adding in this additional relevant evidence will provide more information about on which side of the sufficiency line the case belongs. By contrast, a purely micro-level response will not further these goals. In each of the above three situations, it will reduce the evidentiary base on which decisions are made. Moreover, in situation two, it will not necessarily do anything to prevent convictions based on the other evidence. And in situation three, the reduced evidentiary base will provide less accurate information about on which side of the sufficiency line the case belongs—frustrating the goals of the macro-level rules.

65 See supra notes 51, 53.
66 Note that in this situation, the forensic evidence may still have played a causal role in the outcome—but if the other evidence is sufficient, it is not clear how this can be a problem epistemically. And if the other evidence by itself is not sufficient, then it is one of the other two situations described above.
67 Reducing the evidentiary base in this way might be justified if the decision rule were not doing enough to distribute the risk of error. See supra notes 29–30 and accompanying text. Under such an argument, increasing the risk of total errors would be justified because doing so would better allocate the risk of error among the parties. This, however, does not appear to be an argument currently being made about forensic-science evidence—for this rationale would apply regardless of the quality of the forensic evidence, and regardless of whether jurors understood its correct probative value, while the arguments for exclusion of forensic evidence typically depend on the poor quality of the evidence or the tendency of jurors to overvalue it.
68 Of course, we can employ both macro- and micro-level rules in tandem. See, e.g., Friedman, supra note 62, at 1047–48 (arguing for macro-level rules for expert testimony along with asymmetric micro-level rules). At least in theory, in situations like situation three, micro-level rules may not be necessary if the macro-level rules are operating appropriately.
69 See supra notes 67–68. Situation one is complicated because, one on hand, excluding the forensic evidence will weaken the evidentiary base upon which the decision
Unfortunately, in practice, macro-level doctrine has largely failed to serve these epistemic goals. Although, the BARD instructions play some role in jury decision making,\(^{70}\) they are vague, poorly understood, and epistemically problematic in their focus on subjective belief states rather than on objective qualities of the evidence.\(^{71}\) The sufficiency standard is also without much bite. In Brandon Garrett’s study of 200 wrongfully convicted defendants, for example, 60 raised sufficiency challenges, and only one was successful.\(^{72}\) As a general matter, Judge Jon Newman has called for courts to take sufficiency review more seriously, noting that courts overturn convictions on sufficiency grounds “very rarely,” and, even then, “they almost never do so by applying, in explicit terms, the ‘reasonable doubt’ standard.”\(^{73}\) He explains that “on those rare occasions when a federal appellate court accepts a claim that a case should not have gone to a jury, it typically says simply that the evidence is ‘insufficient’ . . . that there is no evidence to support a necessary element.”\(^{74}\) Indeed, the failures of the doctrine with regard to these macro-level issues may explain much of the concerns regarding admissibility in the criminal context. Admissibility becomes more critical when admissibility decisions are also, in effect, macro-level sufficiency determinations.

The forensic science issues present an important context for courts (and evidence scholars) to develop and improve this macro-level doctrine. This kind of macro-level doctrine is possible, as a brief reflection on the civil context demonstrates. In addition to providing rigorous review on the admissibility of expert testimony, courts in civil cases have developed quite sophisticated sufficiency-review doctrine: summary judgment and judgments as a matter of law.\(^{75}\) These reviews also depend on what a reasonable or rational fact finder can conclude based on the evidence,\(^{76}\) and the primary criticism in this context is not

---


\(^{72}\) Garrett, supra note 55, at 112.


\(^{74}\) Id.; see Cuellar v. United States, 128 S. Ct. 1994, 2005 (2008) (“[T]he Government failed to introduce any evidence that the reason drug smugglers move money to Mexico is to conceal or disguise a listed attribute of the funds.”).

\(^{75}\) FED. R. CIV. P. 56, 50.

\(^{76}\) See Reeves v. Sanderson Plumbing Prods., Inc., 530 U.S. 133, 149 (2000) (“Under Rule 50, a court should render judgment as a matter of law when ‘a party has been fully heard on an issue and there is no legally sufficient evidentiary basis for a reasonable jury to

---
that the doctrine is feckless. To the contrary, it is that courts are too eager to grant motions for summary judgment or judgment as a matter of law.\textsuperscript{77} My point is not to suggest that these areas provide a model for the criminal context to emulate;\textsuperscript{78} rather, the point is merely that courts are capable of taking this issue seriously on a grand doctrinal scale.\textsuperscript{79} Exactly how such review should proceed depends on the soundness of its theoretical foundation. The following paragraphs sketch how this might proceed in the forensic science context.

Under the explanatory conception of legal proof,\textsuperscript{80} the macro-level decision rules (as well as the micro-level issues regarding relevance and probative value) may be articulated based on the explanatory relationships between the evidence and the facts at issue. How ought this play out in the forensic science context? Under the explanatory conception of BARD, a fact is proven when it is part of a plausible explanation of guilt and there is no plausible explanation consistent with innocence.\textsuperscript{81} How forensic science fits into this framework will depend, as the NAS Report notes, on the context and details of particular cases.\textsuperscript{82} Generally, however, (1) a prosecution case built on weak forensic science will fail to be plausible when there is an absence of (or not much) other evidence rendering the prosecution’s explanation plausible, and (2) weak forensic science offered by the prosecution will not by itself render an otherwise plausible defense explanation implausible.\textsuperscript{83}


\textsuperscript{78} For criticism of this standard and its applications in the civil context, see \textit{Suja A. Thomas, The Fallacy of Dispositive Procedure}, 50 B.C. L. REV. 759, 769–78 (2009).

\textsuperscript{79} Furthermore, courts are capable of taking the issue seriously in light of the details of particular cases. For an example of a court doing so—and doing so consistent with the explanatory analysis presented above—see \textit{United States v. Navedo, 443 F. Supp. 2d 431, 434 (W.D.N.Y. 2006)} (“[T]here was insufficient evidence at trial to prove beyond a reasonable doubt that defendant \textit{knowingly} possessed the hidden drugs. In fact, the evidence equally supports an inference that defendant \textit{did not know} about the hidden drugs.”).

\textsuperscript{80} Although I proceed under the explanatory conception, a similar path could be taken under a probabilistic conception. Under this approach, one would need to specify a probabilistic conception of BARD, jury instructions that convey this conception accurately and usefully, and a probabilistic standard for determining when the evidence is sufficient (that is, for when jury conclusions are reasonable).

\textsuperscript{81} See \textit{supra} note 37 and accompanying text.

\textsuperscript{82} \textit{NAS REPORT, supra} note 1, at 1-6.

\textsuperscript{83} Similarly, a defense case built solely on weak forensic evidence may not provide a plausible explanation consistent with innocence. None of these points is meant to suggest that the ability of an expert to \textit{explain} the technique will necessarily be better evidence than evidence that the technique has been tested empirically and has a known error rate. For an
Jury instructions to this effect may better align decision making with the epistemic goals of the proof process. More importantly, courts ought to dismiss a case or overturn a conviction when either of two conditions obtain: (1) a prosecution case relies on weak forensic science and otherwise weak evidence (such that it fails to render the prosecution explanation plausible), or (2) the defense offers a plausible explanation of the evidence that is challenged only by otherwise shaky forensic science. Further generalities may emerge among common-law adjudication along these lines, but this general framework allows for doctrinal development in an epistemically justified manner in light of evidence theory. It allows for proof to proceed in a way that may reduce errors, while maintaining a justified allocation of the risk of errors among the prosecution and defendants, along with also providing a conceptual foundation for courts to place rational constraints on uses of forensic science evidence.

CONCLUSION

The NAS Report recommends a wholesale overhaul of the field of forensic science in the United States. Most of the proposed recommendations for reform focus on activities that take place outside of the courtroom, reforms aimed at improving the quality of the evidence that is used for law enforcement and criminal litigation. In many ways, however, the report invites courts to respond to, and to perhaps improve upon, the ways in which such evidence is admitted and presented currently and in the future. It also invites evidence scholars to develop solid theoretical foundations for possible avenues of response and reform. Toward these ends, I have outlined two general avenues for doctrinal response to problems with forensic science evidence—arguing that macro-level responses in terms of decision rules and sufficiency determinations may provide a more justified response than micro-level admissibility determinations—and I have sketched how such doctrinal reform might proceed in terms of explanatory criteria.
STABLE ISOTOPES AND COURTS

James R. Ehleringer* and Scott M. Matheson, Jr.**

TABLE OF CONTENTS

INTRODUCTION ................................................................. 386
I. BACKGROUND ON STABLE ISOTOPES AND FORENSIC APPLICATIONS 388
II. ADMISSIBILITY STANDARDS AND STABLE ISOTOPE RATIO ANALYSIS 398
   A. The Federal Rules and Daubert ........................................... 398
      1. Rule 104(a) ............................................................... 399
      2. Rule 702 ........................................................................ 400
      3. Rule 703 ........................................................................ 408
      4. Rule 706 ........................................................................ 409
      5. Rule 403 ........................................................................ 410
      6. Summary ................................................................. 411
   B. Application of Evidence Rules to Stable Isotope Ratio Analysis 411
      1. Introduction ............................................................... 411
      2. Inauspicious Debuts ..................................................... 412
      3. Rule 104(a) and the Daubert/702 Analysis .................... 414
      4. Rule 702 Analysis ....................................................... 414
      5. Rule 703—Facts or Date Relied Upon—Otherwise Inadmissible 429
      6. Rule 706—Court-Appointed Experts .............................. 430
      7. Rule 403—Exclusion Based on Prejudice, Confusion, or Duplication ... 431
III. FORENSIC IDENTIFICATION, DNA PROFILING, AND STABLE ISOTOPE RATIO ANALYSIS 432
   A. DNA Profiling and Traditional Forensic Techniques .................. 432
   B. DNA Profiling—Brief Overview ............................................ 435
   C. DNA Profiling and Stable Isotope Ratio Analysis .................. 436
      1. Some Similarities and Differences .................................. 437
      2. Principles and Methods ................................................. 439
      3. Application of Principles and Methods ............................. 440
      4. DNA Experts and Isotope Experts .................................. 440
      5. Helping—Not Confusing—the Jury .................................. 440
CONCLUSION ................................................................. 441

* © 2010 James R. Ehleringer, Distinguished Professor of Biology, University of Utah.
** © 2010 Scott M. Matheson, Jr., Professor of Law, Hugh B. Brown Presidential Endowed Chair in Law, University of Utah.

We wish to thank the following for their review and comments on drafts of this Article: Tom Brenna, Teneille Brown, John Casale, Lynn Jorde, Dan Medwed, and Libby Stern. We also wish to thank Bentley Mitchell and Christine Poleshuk for their research assistance.
INTRODUCTION

Forensic investigation and courtroom fact-finding seek to determine how seemingly identical evidence specimens are related or where certain evidence came from. For example, does the bullet removed from the victim match the ammunition belonging to the suspect? Do both of these recovered bombs share a common source? Where were these drugs grown or produced? Where was this counterfeit money made? Where was this unidentified body located prior to death?

To be more specific: Were the explosives seized separately from terrorist shoe bombers Richard Reid in 2001 and Saajid Badat in 2003 of common origin? Was natural or synthetic testosterone detected in the urine samples collected from world-class cyclist Floyd Landis during the 2006 Tour de France? What laboratory location may have produced the anthrax that was mailed in 2001 to various targets in the tense aftermath of the September 11 attacks?

Stable isotope ratio analysis was used in each of these investigations. In the shoe bomber case, Mr. Badat pleaded guilty before his trial began. In the cycling case, the Court of Arbitration for Sport relied on stable isotope evidence to conclude that Mr. Landis should be banned from cycling for two years. In the anthrax case, the suspected perpetrator committed suicide before the case was presented to a court. The anthrax case will be used as an example in this Article.

Measuring the abundances of naturally occurring chemical stable isotopes can help answer forensic evidence questions. Quantitative measurements of stable isotopes have been used extensively in the fields of biology, chemistry, ecology, geology, and oceanography for decades. The extension of stable isotope analysis into forensic investigation and identification is more recent. Its presentation in the courtroom is inevitable. Judges and lawyers need to understand the utility and limitations of this technical measurement and potentially significant tool.

What do judges and lawyers need to know about stable isotope data? Is stable isotope evidence relevant? Reliable? Admissible? What should the proponent of stable isotope evidence be required to prove? What challenges to it should be made? How strong are factual claims based on this evidence? How can jurors evaluate it with fairness and understanding?

Rule 702 of the Federal Rules of Evidence and the landmark U.S. Supreme Court case of Daubert v. Merrell Dow Pharmaceuticals, Inc., assign judges the

duty to evaluate the reliability and validity of scientific evidence and expert testimony. To do so, judges and lawyers must become knowledgeable about scientific methods in general and about specific theories and techniques applicable to particular cases.

The purpose of this Article is to assist courts and counsel to understand measurements of stable isotopes and the admissibility issues this evidence presents. It attempts to:

- explain stable isotope abundance measurement and its many potential forensic applications;
- identify the range of evidentiary claims that reasonably can be made about stable isotope evidence, both to show the powerful potential of this methodology but also its limitations;
- suggest the threshold presentation that proponents of stable isotope evidence should make and what judges should expect;
- offer questions that opponents of this evidence should raise and that judges should consider;
- analyze the reliability issues based on Daubert and other factors;
- describe the qualifications that should be required of an expert witness on stable isotope evidence; and
- propose the manner in which stable isotope evidence can be presented to the trier of fact so that it can be fairly understood.

After presenting background on stable isotope analysis and expert evidence admissibility standards, we apply admissibility requirements to this methodology. We consider the use of stable isotope analysis in the investigation of the anthrax attacks of 2001 to illustrate a specific application. We urge courts to apply the rigorous reliability scrutiny to stable isotope analysis that already has been applied to DNA profiling evidence. Accordingly, this Article includes discussion comparing and contrasting stable isotope analysis with DNA identification evidence.

Our goal is to help judges, lawyers, and juries to evaluate stable isotope evidence in a manner that meets the high standards of scientific integrity and due process that our society should demand in resolving critical issues in the courts.

The recent National Academy of Sciences report on forensic science, *Strengthening Forensic Science in the United States: A Path Forward* (“NAS Report”), underscores the importance of achieving this goal. Among its significant conclusions, the report states: “The bottom line is simple: In a number of forensic

---

5 “If scientific insights are going to play a supportive role in the legal process, they must be expressed in a language that legal actors can understand.” ROBIN FELDMAN, THE ROLE OF SCIENCE IN LAW 175 (2009).

6 Professor Ehleringer participated with other scientists in applying stable isotope analysis in the Amerithrax case, the name given to the investigation of the 2001 anthrax attacks.

7 NATIONAL RESEARCH COUNCIL OF THE NATIONAL ACADEMIES, STRENGTHENING FORENSIC SCIENCE IN THE UNITED STATES: A PATH FORWARD (2009) [hereinafter NAS REPORT].
science disciplines, forensic science professionals have yet to establish either the validity of their approach or the accuracy of their conclusions, and the courts have been utterly ineffective in addressing this problem.8

We seek a better bottom line for forensic use of stable isotope analysis. With solid research, strict adherence to protocols and standards, accurate reporting of results and within-specimen variance, and effective courtroom presentation, stable isotope analysis can contribute to the search for truth in our justice system.

I. BACKGROUND ON STABLE ISOTOPES AND FORENSIC APPLICATIONS

The scientific literature on stable isotope analysis is extensive. The following provides basic background information.

What is an isotope? Atoms of the same element having different numbers of neutrons are called isotopes. A unique characteristic of each chemical element is the number of protons in its nucleus. That number corresponds to the element’s place on the Periodic Table of Elements. As examples, a hydrogen atom has one proton, a carbon atom has six protons, and an oxygen atom has eight protons. Atoms of the same element typically, though not always, have the same number of neutrons in their nuclei as they have protons. Some atoms of the same element have greater or fewer neutrons than the most common form of the element.9

Virtually all elements on the Periodic Table of Elements have multiple isotopes. We denote different isotopes by listing the number of protons and neutrons as a superscript to the left of the element’s symbol. For example, most hydrogen atoms have one proton and no neutrons (\(=^1\)H), but a few have one proton and one neutron (\(=^2\)H); even fewer have one proton and two neutrons (\(=^3\)H). As a second example, most carbon atoms have six protons and six neutrons (\(=^{12}\)C), but other naturally occurring isotopes of carbon can have seven (\(=^{13}\)C) or eight neutrons (\(=^{14}\)C).

---

8 Id. at 53.
9 “[T]he nuclei of atoms are made up of various combinations of protons and neutrons, and . . . within atoms of a given element, the ratio of neutrons to protons may vary from isotope to isotope . . . .” Allen v. United States, 588 F. Supp. 247, 270 (D. Utah 1984), rev’d on other grounds, 816 F.2d 1417 (10th Cir. 1987), cert. denied, 484 U.S. 1004 (1988).
Isotopes come in two forms: stable and radioactive. Isotopes that persist in their same elemental form are stable isotopes. That is, over time, these atoms will not change and will not decay into another element. Most elements on Earth are stable isotopes. In contrast, trace amounts of elements on Earth are radioactive. Radioactive isotopes are not stable and will decay over time from one element to another as parts of the nucleus leave the atom in a sequence of radioactive decay. The lifetime of radioactive elements may be as short as a few nanoseconds for laboratory-produced radioactive isotopes or as long as many billions of years (e.g., Rubidium-87 ($^{87}$Rb)). In between is the familiar example of $^{14}$C dating, a powerful tool for dating materials that are in the age ranges of 300-50,000 years\(^{10}\) or of 1-40 years.\(^{11}\) Naturally occurring stable isotopes are the focus of this paper.

For an element, how are the abundances of stable isotopes expressed? As an introduction, the isotope abundances of an element can be expressed as percentages. Consider carbon: 98.89% of carbon is $^{12}$C, while 1.11% is $^{13}$C, and $^{14}$C amounts to less than one ten-billionth of 1%. Lead is a more complicated

---


example, with 52.4% of lead as $^{208}\text{Pb}$, 22.1% as $^{207}\text{Pb}$, 24.1% as $^{206}\text{Pb}$, and 1.4% as $^{204}\text{Pb}$.

Expressing abundance as a percentage provides only a “low-resolution” approximation of abundance. Instead, the preferred means of expressed natural abundances of stable isotopes is the ratio of the rare-to-common isotope forms (e.g., $^2\text{H}/^1\text{H}$, $^{13}\text{C}/^{12}\text{C}$, $^{15}\text{N}/^{14}\text{N}$, $^{18}\text{O}/^{16}\text{O}$, $^{87}\text{Sr}/^{86}\text{Sr}$, $^{204}\text{Pb}/^{208}\text{Pb}$). For the heavy elements in the Periodic Table of Elements (e.g., Sr and Pb), it is sufficient to describe abundance as the ratio of the rare-to-common isotope forms. However, for commonly analyzed light isotopes (e.g., H, C, N, O, Cl, S), the ratio of the rare-to-common isotope forms is not sufficient because the rare forms are rare enough that the ratio is usually less than 0.01. In response, the scientific community has adopted the “delta notation” ($\delta$) to describe the isotope ratios of light elements. Here, the ratio of the heavy-to-light stable isotopes of a sample ($R_{\text{sample}}$) is compared to the equivalent ratio of an internationally recognized standard ($R_{\text{standard}}$) as

$$\delta(\text{‰}) = \left( \frac{R_{\text{sample}} - R_{\text{standard}}}{R_{\text{standard}}} \right) \times 1000 \text{‰}$$

where ‰ (per mil) is interpreted as the difference, in parts per thousand, between the sample and the international reference standard. Per mil is per thousand, which is the same as ten times percent (parts per hundred).

The $\delta$ value of an international reference standard is defined as 0‰. For hydrogen, we express the delta notation hydrogen isotope ratio as $\delta^2\text{H}$, for carbon as $\delta^{13}\text{C}$, and so forth. Note that the delta notation isotope ratio value may have a positive or negative value, depending on how much of the rare isotope is in the sample versus the standard. International standards are available through the National Institute of Standards and Technology (NIST), which is part of the U.S. Department of Commerce, and through the International Atomic Energy Agency (IAEA) in Vienna, Austria.

Laboratories create working reference materials (also known as working standards) for daily analytical uses through comparisons of working reference materials with international standards. Because the international standards are precious materials and available only in limited quantities, it is an accepted practice to rely on calibrated laboratory-based or association-based working reference materials. Laboratories exchange their working reference materials with other laboratories to verify that the materials have a correct value relative to the international standards. Round-robin exchange tests among laboratories (also known as ring tests) are also a common practice to ensure that analyses conducted in one laboratory will produce the same results when performed in a different laboratory.

How are stable isotope ratios measured? High-precision measurements are required because the differences in stable isotope ratio values among samples can be small. Light elements in the Periodic Table of Elements, such as H, C, N, O, Cl,
and S, are measured on an isotope ratio mass spectrometer (IRMS).\textsuperscript{12} There are other established and emerging techniques for measuring isotopic variability in nature. For H and O isotopes in water\textsuperscript{13} or C isotopes in carbon dioxide,\textsuperscript{14} stable isotope ratios can also be measured with laser spectroscopy.

For some applications, isotopic variability can be measured by nuclear magnetic resonance (NMR) techniques, often referred to as \textit{site-specific natural isotope fractionation} (SNIF-NMR). Strictly speaking, SNIF-NMR does not measure isotope ratios but does measure natural isotopic variability within molecules in liquids, and it is most commonly applied to the detection of adulteration in foods.\textsuperscript{15} Traditional structural or organic mass spectrometer (MS) approaches will not provide the accuracy required for a high-precision isotope ratio measurement.

For both IRMS and laser-spectroscopy measurement methods, the sample is converted into a gas, and the different isotopes in the gases are measured in separate detectors. With IRMS, other instruments are often coupled in front of the IRMS to convert the sample from its original form into the gas, which is measured in a continuous analytical process. Quality-control samples with a known isotopic composition are analyzed before and after the sample is analyzed to ensure the long-term accuracy of an observation.\textsuperscript{16} For heavy elements in the Periodic Table of Elements, such as Sr and Pb, the samples are measured on a thermal ionization

\begin{footnotesize}

\textsuperscript{13} See Willi A. Brand et al., Letter to the Editor, \textit{Cavity Ring-Down Spectroscopy Versus High-Temperature Conversion Isotope Ratio Mass Spectrometry; A Case Study on $\delta^2$H and $\delta^{18}$O of Pure Water Samples and Alcohol/Water Mixtures}, 23 \textit{RAPID COMM. MASS SPECTROMETRY} 1879 (2009).

\textsuperscript{14} See Ed. H. Wahl et al., \textit{Applications of Cavity Ring-Down Spectroscopy to High Precision Isotope Ratio Measurement of $^{13}$C/$^{12}$C in Carbon Dioxide}, 42 \textit{ISOTOPES ENVTL. & HEALTH STUD.} 21 (2006).


\end{footnotesize}
mass spectrometer (TIMS) or a multi-collector inductively coupled plasma mass spectrometer (MC-ICP-MS).\textsuperscript{17}

Stable isotope analyses can be applied to samples that consist of the entire organism or mixture (bulk level), specific compounds within an organism or within a mixture (compound-specific level), and specific locations within a complex molecule (position-specific or intramolecular level), reflecting increasing levels of chemical specificity. Samples are most commonly measured at the bulk level; that is, the isotope ratio is determined without chemical separation in a usually complex chemical mixture. Compound-specific isotope analysis is required in some contexts. For instance, it is used to establish whether an athlete’s testosterone is naturally produced by the body or is instead a supplemental hormone taken to stimulate athletic performance. Finally, intramolecular isotope ratios, often determined by SNIF-NMR, report on isotopic structure within molecules.

What is an “isotope fingerprint”? “Isotope fingerprint” is a commonly used term in the isotope field to describe combinations of stable isotope ratio observations from a particular specimen. It is an imperfect metaphor because the isotopic characteristics of elements in a sample are not marks or impressions left at a crime scene, and they are analyzed very differently than fingerprints.\textsuperscript{18} Nonetheless, the term is used to describe the stable isotope ratios of the chemical elements in a sample taken from a particular specimen, and they can be compared to the observations from one or more other specimens.

Why are stable isotope ratios useful in forensic identification? The various forensic identification techniques in criminal investigation and prosecution share the goal “of matching a sample associated with the defendant (or victim) to a sample found at the crime scene.”\textsuperscript{19} Comparable applications are made in civil and administrative proceedings as well. Forensic identification is based on differentiation. One challenge that can arise is the ability to distinguish between specimens that are chemically identical. That is, going beyond traditional chemical identification approaches, is it possible to distinguish between two or more specimens that are known to contain the same compound (e.g., the same explosive compound or the same sugar)? Stable isotope ratio analysis may help by providing an additional piece of information about the specific compounds.

If two chemically identical specimens share a common origin or relationship, then we would expect these compounds to have similar stable isotope ratios for


\textsuperscript{18} See Alexandra J. Roberts, Everything New Is Old Again: Brain Fingerprinting and Evidentiary Analogy, 9 Yale J.L. & Tech. 234, 242–56 (2007) (arguing that analogizing novel evidence to more-established forms can hinder understanding of the former).

each of the elements in the compound. The source of a sample may be based on production considerations or on geographical location.\textsuperscript{20} If a substance of unknown origin appears to be isotopically the same as a substance of known origin, it is at least possible and perhaps probable that the unknown substance and the known substance have the same origin.\textsuperscript{21} If the unknown substance and the known substance are different isotopically, then they likely have different origins—geographical origin and/or production method—unless one of the two specimens was treated in such a way as to cause its isotope ratio to change.

Isotope measurement is therefore used to determine whether a sample or specimen is excluded from or consistent with a known source and perhaps even to reach a conclusion that it is a highly probable match with a known source. Stable isotope analysis can be used to reinforce or corroborate other supporting information. It can be used to distinguish among or eliminate other specimens as possibly related to the specimen of interest.

The relative amounts of heavy and light stable isotopes for a given element vary in nature based on physical and biological processes.\textsuperscript{22} This naturally occurring variation allows for forensic identification.\textsuperscript{23} The ratios may vary based on origin, including geographic location,\textsuperscript{24} thereby allowing analysis of whether a sample containing a particular chemical came from a particular location. Ratios also can vary as a result of adulteration, in which chemically similar compounds are substituted for one another, such as substituting a cheaper compound for a more expensive compound.\textsuperscript{25} For any given unknown substance, there may be


\textsuperscript{22} See ZACHARY SHARP, PRINCIPLES OF STABLE ISOTOPE GEOCHEMISTRY 7 (2006); Gabriel J. Bowen et al., Isoscapes to Address Large-Scale Earth Science Challenges, 90 EOS 109 (2009); Wolfram Meier-Augenstein & Ray H. Liu, Forensic Applications of Isotope Ratio Mass Spectrometry, in ADVANCES IN FORENSIC APPLICATIONS OF MASS SPECTROMETRY 149, 151-52 (Jehuda Yion ed., 2004).

\textsuperscript{23} See James R. Ehleringer et al., Forensic Science Applications of Stable Isotope Ratios, in FORENSIC ANALYSIS ON THE CUTTING EDGE: NEW METHODS FOR TRACE EVIDENCE ANALYSIS 399, 401–05 (Robert D. Blackledge ed., 2007) [hereinafter Ehleringer et al., Applications]; Meier-Augenstein & Liu, supra note 22, at 150–53; see generally MEIER-AUGENSTEIN, supra note 1.

\textsuperscript{24} See Jason B. West et al., Stable Isotopes as One of Nature’s Ecological Recorders, 21 TRENDS IN ECOLOGY & EVOLUTION 408 (2006).

\textsuperscript{25} See R. Fügel et al., Quality and Authenticity Control of Fruit Purées, Fruit Preparations and Jams—A Review, 16 TRENDS IN FOOD SCI. & TECH. 433 (2005); Jonathan W. White, Internal Standard Stable Carbon Isotope Ratio Method for Determination of C4
more than one chemical element to choose from for stable isotope analysis, enabling a higher level of confidence with regard to the conclusion.

What are the forensic applications of stable isotope analysis? The ability of stable isotope analysis to match or distinguish evidence specimens points to three key applications. The first is to determine whether two specimens have the same isotopic composition and therefore may have a common origin. This application describes the classic forensic identification task of comparing an unknown sample or questioned item from a crime scene to a known sample or exemplar taken from a subject or victim. The first step is to compare the items and determine whether they appear to match. If they do, the second step is to determine the probability that the items came from the same source.26 The reference and test materials that are analyzed are often called associative evidence.27

The second application is to ascertain the expected stable isotope ratios for specimens from a given location (e.g., the origin or authenticity of a food product) or a production method (e.g., an explosive compound). Then the stable isotope ratio of a sample of unknown origin can be compared to the collection of known observations from the given location, such as data contained in a database of authentic observations.

The third application is to determine a probable geographic source location of an unknown specimen.28 These source attribution applications enable forensic assignment of evidentiary material to a geographic region of origin by comparing the isotopic composition of the sample to geospatial mapping of predicted stable isotope ratios.29 Maps have been developed for the predicted isotopic composition of water throughout the world and serve as a principal basis for geographic identification of a wide range of stable isotope values in plants, animals, and microbes based on water as a substrate.30

What are some examples of stable isotope ratio analysis applications to topics of forensic interest? Some of the many possible examples include:
- distinguishing between real and adulterated food products through carbon isotope analyses;31

---

26 See Saks & Koehler, supra note 21, at 199.
27 See Thornton & Peterson, supra note 20, at 62.
30 See id. at 40–43.
31 See MEIER-AUGENSTEIN, supra note 1, at 37–42; Ehleringer et al., Applications, supra note 23, at 406–08; Meier-Augenstein & Liu, supra note 22, at 163–65.
- distinguishing between natural and manufactured fertilizers associated with the production of natural foods;32
- distinguishing between vegetable products grown with natural versus manufactured fertilizers;33
- distinguishing between fruit juices that are natural versus juices that are adulterated by sugar substitutes like high fructose corn syrup;34
- distinguishing between real versus manufactured flavorings, such as vanillin;35
- determining the region or manufacturer of illicit drug samples;36
- determining the source of precursor materials for the manufacture of methamphetamine;37
- matching the isotopic composition of an organism with its food and water diet, given the influence of variations in the isotopic composition of water at various geographic locations around the world;38
- ascertaining the geographic origins or movements and dietary patterns of individuals based on the hydrogen, oxygen, carbon, and nitrogen isotope analyses of human hair, bones, fingernails, and/or teeth;39

33 See Karyne M. Rogers, Nitrogen Isotopes as a Screening Tool to Determine the Growing Regimen of Some Organic and Nonorganic Supermarket Produce from New Zealand, 56 J. AGRIC. & FOOD CHEMISTRY 4078 (2008).
36 See John Casale et al., Stable Isotope Analyses of Heroin Seized from the Merchant Vessel Pong Su, 51 J. FORENSIC SCI. 603 (2006); Ehleringer et al., Applications, supra note 23, at 408–11; Ehleringer et al., Spatial Considerations, supra note 16, at 49–50; Meier-Augenstein & Liu, supra note 22, at 168–70.
37 See Michael Collins et al., \( \delta^{13}C, \delta^{2}N \) and \( \delta^{2}H \) Isotope Ratio Mass Spectrometry of Ephedrine and Pseudoephedrine: Application to Methylamphetamine Profiling, 23 RAPID COMM. MASS SPECTROMETRY 2003 (2009).
analyzing the isotopic composition of lead bullets and primers to ascertain which ammunition and/or firearm caused a particular gunshot entry;\(^{40}\)

- predicting the geographic origin and movement of humans based on the stable isotope composition of their scalp hair;\(^{41}\)

- tracing the origin of counterfeit currency based on isotopic differences arising from the cotton used to produce the security paper;\(^{42}\) and

- identifying the starting materials or manufacturing processes by comparing the differentiated isotopic characteristics of organic peroxides in improvised explosives used by terrorists.\(^{43}\)

**How precise and accurate can stable isotope ratio analysis be in identifying the source of an evidence sample?** The resolving power of stable isotope ratio analysis can vary depending on many factors, including (a) the precision of the analytical measurement, (b) the heterogeneity within a sample or among a common population, and (c) the nature and quality of the reference database. The results of this forensic method range from matching a specimen to a specific source with a high level of confidence to linking a sample only with a large geographic region and perhaps even more than one region.

This point can be understood through the example of recent research on the potential for use of strontium isotope ratios ($^{87}$Sr/$^{86}$Sr) to determine the geographic origins of marijuana, such as seized samples.\(^{44}\) Strontium isotopes were considered in region-of-origin analyses, independent of the use of hydrogen isotopes to predict region of origin. Strontium isotope ratios from plants cultivated in seventy-nine counties throughout the United States were compared with the ratios expected from bedrock-based ages contained in U.S. Geological Survey data.\(^{45}\)

The results showed that the marijuana strontium isotope ratios retained a primary geologic signal that would facilitate geologic sourcing.\(^{46}\) But how precise the sourcing can be in a particular case will vary. A second approach is to predict region of origin using hydrogen isotopes. A survey of hydrogen isotope ratios of

---


\(^{43}\) See Sarah J. Benson et al., *Forensic Analysis of Explosives Using Isotope Ratio Mass Spectrometry (IRMS)—Preliminary Study on TATP and PETN*, 49 SCI. & JUST. 81 (2009). For a recent overview of forensic applications of stable isotope analysis, see MEIER-AUGENSTEIN, supra note 1, at 145–221.

\(^{44}\) See Janet M. Hurley et al., *Stable Isotope Models to Predict Geographic and Cultivation Conditions of Marijuana*, 50 SCI. & JUST. 86 (2010); Jason B. West et al., *The Stable Isotope Ratios of Marijuana. II. Strontium Isotopes Relate to Geographic Origin*, 54 J. FORENSIC SCI. 1261 (2009).

\(^{45}\) See West, supra note 44, at 1262–63.

\(^{46}\) See id. at 1263–67.
marijuana from all fifty states was compared to maps of the water isotopes that are known to exhibit a coherent, repeatable pattern of differences across the United States.\(^{47}\) Two independently measured stable isotopes can be used to predict region of origin. A Venn-diagram space can represent the region of origin with the highest probability of being the source region for the material of interest.

The study showed that some regions would be indistinguishable without more detailed information about the particular regions. On the other hand, where the strontium bedrock and hydrogen water isotope values vary considerably between two locations, isotope ratio analysis may be able to predict whether a specimen originated from one of two possible locations with a high degree of confidence.\(^{48}\) This example not only illustrates the isotope method’s range of precision in resolving the source of an unknown sample, but also shows how information from other investigative techniques (in the example here, narrowing the possible source locations to two) can influence that range. This example also shows the ability of the technique to address questions about the distribution of controlled substances, such as the origins of marijuana seized in different regions of the United States.\(^{49}\)

What are the limitations of stable isotope ratio analysis? The principal limitations are both unique to stable isotope analysis and similar to other scientific applications. Five examples illustrate these points.

1. In some cases, the heterogeneity within a sample results in a variance that is large enough to make it difficult to distinguish among samples that have similar stable isotope ratios.\(^{50}\)

2. As in any laboratory analysis in which there are multiple steps in a sequence, each step in the preparation, isolation, and purification of a specimen prior to its analysis can decrease the overall precision, even though the precision at any one step in the process can be high. This concern applies to other forensic identification techniques as well.\(^{51}\)

3. The quality of isotopic databases can vary based on the availability of authentic materials, the history and preservation of materials prior to analyses, and the confidence as to the exact origins of the materials. For some elements at some locations, the databases can be substantially


\(^{48}\) See West, *supra* note 44, at 1267.


\(^{50}\) See Ehleringer et al., *Applications, supra* note 23, at 405.

complete and very reliable. For other elements and other locations, the databases may not be as extensive.

(4) The limits of valid statistical analysis and sampling can affect the certainty levels of this methodology. It is important that expert witnesses in this field avoid the exaggerated testimony that has been criticized in other fields. As the NAS Report explained: “The insistence by some forensic practitioners that their disciplines employ methodologies that have perfect accuracy and produce no errors has hampered efforts to evaluate the usefulness of the forensic science disciplines.”

(5) Issues of contamination and degradation of samples as well as accurate evidence handling and storage apply to stable isotope samples, just as they do to other forensic sample evidence.

We will explore the precision and limitations of stable isotope ratio analysis further in our later discussion of the reliability considerations regarding this evidence.

II. ADMISSIBILITY STANDARDS AND STABLE ISOTOPE RATIO ANALYSIS

Although stable isotope ratio analysis has been used to facilitate investigations, it has not yet become a frequent source of expert evidence in the courtroom. We expect that it will. This Part discusses rules of admissibility applicable to stable isotope ratio analysis, with particular emphasis on the issues of scientific validity under the Daubert framework. We then turn to application of these standards to stable isotope ratio analysis. The goal is to develop the analytical framework to aid courts and litigants in framing and addressing the issues of admissibility under “the exacting standards of reliability such evidence must meet.”

A. The Federal Rules and Daubert

The admissibility of expert testimony depends on two key elements: the test used to evaluate the proposed evidence and the judge’s role in applying that test. Although both continue to differ from state to state, the Federal Rules of Evidence on expert testimony and the U.S. Supreme Court’s interpretation of what those rules require have substantially shaped federal and state evidence law. Many states have adopted some version of the Daubert reliability standard. This Article uses

---

52 NAS REPORT, supra note 7, at 47.
54 See 1 MODERN SCIENTIFIC EVIDENCE—THE LAW AND SCIENCE OF EXPERT TESTIMONY 3–4 (David L. Faigman et al. eds., 2009–10 ed.).
the *Daubert* framework for admission of expert testimony in federal court to analyze stable isotope ratio evidence.

As the following recounts, in federal courts the responsibility for threshold evaluation of the validity of scientific evidence is placed on the judges, who make the admissibility decision. If scientific evidence is admitted, the trier of fact determines its weight. Current evidence law and practice leaves the search for and selection of experts primarily to the parties and anticipates that experts will testify in an adversarial system through which their views may be challenged by opposing experts and on cross-examination as biased and unreliable.\(^56\) The legal framework for the admission of expert testimony in federal trials is found primarily in Federal Rule of Evidence 702, but other Federal Rules also play an important role. Rules 104(a), 702, 703, 706, and 403 will be addressed here.\(^57\)

1. Rule 104(a)

The Federal Rules governing expert testimony divide responsibility between the judge and jury. The judge is responsible for determining the admissibility of expert testimony, including the qualifications of the expert witness and the helpfulness and reliability of the evidence. This is reflected in Rule 104(a), which provides:

Preliminary questions concerning the qualification of a person to be a witness, the existence of a privilege, or the admissibility of evidence shall be determined by the court, subject to the provisions of subdivision (b) [concerning conditional admissions]. In making its determination it is not bound by the rules of evidence except those with respect to privileges.


\(^{57}\) Federal Rules of Evidence 704 and 705 address expert testimony on an ultimate issue and disclosure of facts or data underlying an expert opinion, respectively. Although these rules may be pertinent to expert stable isotope testimony in a particular case, they are not central to the discussion presented here.
The proponent of expert evidence must establish these matters by a preponderance of proof.\(^{58}\) Trial judges have the discretion to make Rule 104(a) determinations based on briefs and other materials submitted before trial, a pretrial evidentiary hearing, or testimony presented at trial.\(^{59}\) Courts often dispense with a pretrial \textit{Daubert} hearing if the expert evidence is well established and no novel challenge is raised.\(^{60}\) If the judge determines that the expert witness is qualified to testify and that the other expert evidence requirements are met, then the evidence is admitted, and the jury’s role is to evaluate the weight of that evidence.\(^{61}\)

2. \textit{Rule 702}

Federal Rule of Evidence 702 provides:

If scientific, technical, or other specialized knowledge will assist the trier of fact to understand the evidence or to determine a fact in issue, a witness qualified as an expert by knowledge, skill, experience, training, or education, may testify thereto in the form of an opinion or otherwise, if (1) the testimony is based upon sufficient facts or data, (2) the testimony is the product of reliable principles and methods, and (3) the witness has applied the principles and methods reliably to the facts of the case.

The rule calls for consideration of scientific evidence based on three criteria: relevancy, qualifications, and reliability. The reliability component can in turn be divided into three issues: sufficiency of the facts or data, reliability of the principles and methods, and proper application of the principles and methods. With respect to the principles and methods, the reliability of the theory (principles) and the technique (methods) are distinct issues.\(^ {62}\)

\(^{58}\) See \textit{Daubert v. Merrell Dow Pharms., Inc.}, 509 U.S. 579, 592 n.10 (1993); \textit{Bourjaily v. United States}, 483 U.S. 171, 175–76 (1987); \textit{Lewis v. CITGO Petroleum Corp.}, 561 F.3d 698, 705 (7th Cir. 2009).

\(^{59}\) See \textit{Millenkamp v. Davisco Foods Int’l}, Inc., 562 F.3d 971, 979 (9th Cir. 2009); \textit{In re Scrap Metal Antitrust Litig.}, 527 F.3d 517, 532 (6th Cir. 2008); \textit{United States v. Kenyon}, 481 F.3d 1054, 1061 (8th Cir. 2007).

\(^{60}\) See \textit{United States v. Pena}, 586 F.3d 105, 111 n.4 (1st Cir. 2009).


\(^{62}\) See PAUL C. GIANNELLI & EDWARD J. IMWINKELRIED, 1 SCIENTIFIC EVIDENCE 2 (4th ed. 2007).
(a) Relevancy

The “assist the trier of fact” language of Rule 702 is sometimes referred to as a “helpfulness” standard and is regarded as a relevance test for expert testimony.63 As the Daubert Court noted, Rule 702’s reference to “assist the trier of fact” “goes primarily to relevance.”64 Even if scientifically valid, the expert testimony must “fit”—it must relate to a disputed issue in the case.65 This test is worded less rigorously than the “beyond the ken” standard that many courts applied before the Federal Rules were adopted.66 The “beyond the ken” test only allowed expert testimony that would provide information which an ordinary person would not otherwise know or understand.67 The “assist” or “helpfulness” threshold is commensurate with and complementary to the general relevance standard of Rule 401, which asks whether evidence makes a fact of consequence to the case more or less probable.68

(b) Qualifications

Experts must be qualified to testify. Rule 702 provides that an expert witness can be qualified “by knowledge, skill, experience, training, or education.”69 This language encompasses a wide array of experts who have developed their expertise through a variety of means, which suggests that courts can approach the qualification issue with flexibility.70 For scientific testimony, the conventional approach is to scrutinize an expert’s credentials, training, and experience.71 Under

---

64 Daubert, 509 U.S. at 591.
65 See id. at 591–92.
66 See United States v. Joyce, 511 F.2d 1127, 1131 (9th Cir. 1974) (“To warrant the use of expert testimony . . . the subject of the inference must be so distinctively related to some science, profession, business or occupation as to be beyond the ken of the average layman.”).
68 FED. R. EVID. 401 (“‘Relevant evidence’ means evidence having any tendency to make the existence of any fact that is of consequence to the determination of the action more probable or less probable than it would be without the evidence.’”). Courts have justified exclusion of proposed expert evidence as not helpful under Rule 702 if it falls within the jury’s “common knowledge.” See Persinger v. Norfolk & W. Ry. Co., 920 F.2d 1185, 1188 (4th Cir.1990); Mueller v. Auker, No. CIV 04-399-S-BLW, 2010 WL 2265867, at *3 (D. Idaho June 10, 2010); Meemic Ins. Co. v. Hewlett-Packard Co., No. 09-10155, 2010 WL 1949750, at *11-12 (E.D. Mich. May 13, 2010).
69 FED. R. EVID. 702.
70 See Watkins v. Telsmith, Inc., 121 F.3d 984, 988 (5th Cir. 1997); Raymond v. Raymond Corp., 938 F.2d 1518, 1526 (1st Cir. 1991).
71 See United States v. Parra, 402 F.3d 752, 758 (7th Cir. 2005).
Rule 104(a), the judge must determine whether the proposed expert is qualified to present the proffered testimony. The jury can evaluate the relative strength of each expert’s qualifications and credibility.  

(c) Reliability

The reliability determination is based on the U.S. Supreme Court’s decision in *Daubert*, which determined “the proper standard for admission of expert testimony” under Federal Rule of Evidence 702. The case was decided in the context of growing concern and controversy over the perceived flood of “junk science” admitted in courtrooms.

The *Daubert* Court agreed that adoption of the Federal Rules of Evidence superseded “the dominant standard for determining the admissibility of novel scientific evidence at trial”—the 70-year-old “general acceptance” test based on *Frye v. United States*. Under the *Frye* test, scientific evidence “must be sufficiently established to have gained general acceptance in the particular field in which it belongs.”

At the time of *Daubert*, Rule 702 read: “If scientific, technical, or other specialized knowledge will assist the trier of fact to understand the evidence or to determine a fact in issue, a witness qualified as an expert by knowledge, skill, experience, training, or education, may testify thereto in the form of an opinion or otherwise.” The Court found nothing in this language or the rule’s drafting history to establish “general acceptance” as the exclusive test for admissibility. It did find in the rules that “the trial judge must ensure that any and all scientific testimony or evidence admitted is not only relevant, but reliable.” Justice Blackmun, writing for the Court, found the basis for this obligation mainly in Rule 702, which requires “that an expert’s testimony pertain to ‘scientific knowledge,’” a term that “establishes a standard of evidentiary reliability.”

Rule 702 and *Daubert* shifted responsibility for determining whether new scientific and technological innovations should be admitted as courtroom evidence.

---

72 See *In re Scrap Metal Antitrust Litig.*, 527 F.3d 517, 529–31 (6th Cir. 2008); *Diestel v. Hines*, 506 F.3d 1249, 1268–69 (10th Cir. 2007).
75 *Daubert*, 509 U.S. at 585–89.
76 *Frye v. United States*, 293 F. 1013, 1014 (1923).
77 *Id.*
78 *Daubert*, 509 U.S. at 588–89.
79 *Id.* at 589.
80 *Id.* at 590.
from the scientists to the judges. Justice Blackmun stressed the “gatekeeping role” of the trial judge to determine whether the expert will testify about (1) scientific knowledge that (2) will assist the trier of fact to understand or determine a fact in issue. He wrote, “We are confident that federal judges possess the capacity to undertake this review.” He pointed to Rule 104(a) as calling upon judges to make this admissibility determination.

For expert testimony, this task calls for a “preliminary assessment of whether the reasoning or methodology underlying the testimony is scientifically valid and of whether that reasoning or methodology properly can be applied to the facts in issue.” This inquiry focuses on the expert testimony’s scientific foundation and rests on a preponderance of the evidence standard. Although many factors may bear on whether expert testimony is based on sound methods and principles, the Court offered five considerations that are integral to the scientific method to determine “whether a theory or technique is scientific knowledge that will assist the trier of fact.”

First, can the theory or technique be and has it been tested under the scientific method of “generating hypotheses and testing them to see if they can be falsified”? The judge must assess the research methods used to test the hypothesis in question. Justice Blackmun explained that the “focus, of course, must be solely on principles and methodology, not on the conclusions that they generate.” The gatekeeper judge therefore must concentrate on the reliability and validity of the general principles or methods on which an expert’s conclusions are based.

The Daubert Court referred to the scientific method as the cornerstone of scientific knowledge and cited the esteemed philosopher of science Karl Popper. But no single scientific method applies to all theories, all fields, and all applications, as Popper himself observed. The common element is that science be open to criticism and revision, the concept he called “falsifiability,” one of the analytical factors the Court emphasized in Daubert.

---

82 Daubert, 509 U.S. at 592, 597.
83 Id. at 593.
84 Id. at 600.
85 Id. at 592 n.10.
86 Id. at 593–94.
87 Id. at 593 (quoting Michael D. Green, Expert Witnesses and Sufficiency of Evidence in Toxic Substances Litigation: The Legacy of Agent Orange and Bendectin Litigation, 86 NW. U. L. REV. 643, 645 (1992)).
88 Daubert, 509 U.S. at 595.
89 See GIANNELLI & IMWINKELRIED, supra note 62, at 52–57.
90 Daubert, 509 U.S. at 593.
92 See id. at 279.
93 Daubert, 509 U.S. at 593. But see FELDMAN, supra note 5, at 128–31 (noting criticism of falsification as “hallmark of true science”).
Second, has the theory or technique “been subjected to peer review and publication”? The Court explained that peer-reviewed publication makes detection of substantive flaws more likely. Although peer-reviewed publication does not guarantee scientific validity, and lack of peer-reviewed publication does not mean a theory or method is unsound, this factor can facilitate critical evaluation of the principles and methods underlying the expert evidence. It contemplates refereed articles, scrutinized by an appropriate peer group, that report the testing of principles and methods and the underlying data so that other scientists can evaluate and verify. Judges should be cognizant that both the quality of publications and of reviewers can vary and that the scientific community’s reactions to the publications are part of assessing this factor.

Third, what is the “known or potential rate of error” for the scientific technique? Courts have not developed much guidance on how judges should approach this issue. Daubert did not specify an allowable error rate, suggesting a balancing analysis that accounts for the costs of mistakes. In criminal cases, false positive error rates are particularly important because a false positive is evidence of identification that supports a conviction. In applied science, the test or the tester produces false positives and false negatives over a quantity of tests. Errors in applied scientific testing can arise in many ways, including sample size, the nature of the sample that is studied, and the equipment that is used. Errors can be random or systematic. Scientific reliability emerges through replication of studies that recognize their own limitations.

Fourth, what are the standards controlling the scientific technique’s operation? The Daubert Court tied the error rate factor to the “existence and maintenance of standards controlling the technique’s operation.” The quality of the standards and such factors as standardization of procedures and laboratory accreditation are relevant to this part of the reliability analysis.

94 Daubert, 509 U.S. at 593.
95 See id.
96 See id.
97 See id.
98 See GIANNELLI & IMWINKELRIED, supra note 62, at 45–46.
99 See 1 MODERN SCIENTIFIC EVIDENCE, supra note 54, at 63–66; State v. Brown, 687 P.2d 751, 769–70 (Or. 1984) (noting abundance of polygraph literature but “availability of this mass of literature may or may not be relevant in any given case”).
100 Daubert, 509 U.S. at 594.
102 See 1 MODERN SCIENTIFIC EVIDENCE, supra note 54, at 63.
103 Daubert, 509 U.S. at 594.
104 Id.
105 See Mitchell, 365 F.3d at 241.
106 See United States v. Prime, 363 F.3d 1028, 1034 (9th Cir. 2004), vacated on other grounds, 543 U.S. 1101, 1101 (2005).
Fifth, is there “general acceptance” of the theory or technique in the relevant scientific community? The general acceptance standard was the exclusive test under Frye, but under Daubert it is just one of several factors a court may consider in the reliability analysis to determine admissibility. The weight of this factor should vary directly with the rigor of scientific testing in the field. Some fields are more rigorous than others in their assessment of hypotheses. Widespread general acceptance should not guarantee admissibility if the evidence comes from a field with lax research methodology; conversely, expert opinion based on methodology that has been extensively and rigorously tested may be admissible despite not yet achieving general acceptance.

The Court stressed that the Rule 702 inquiry should be “flexible” and aimed at “scientific validity—and thus the evidentiary relevance and reliability” of the principles and methodology relied upon to generate conclusions. The expert testimony must have “a reliable foundation” and be “relevant to the task at hand.” This calls for “[p]ertinent evidence based on scientifically valid principles.” After Daubert, federal district court judges have held “Daubert hearings” pursuant to Rule 104(a) to assess the validity of scientific evidence. Even when no Daubert hearing is held, the trial court’s gatekeeping duties call for development of a sufficient record on the basis for admissibility to facilitate meaningful appellate review.

The Supreme Court insisted in Daubert and ensuing cases that the reliability determination can but need not be based on the factors just described. No specific factor is dispositive on the reliability of an expert’s testimony. For any given scientific evidence, the five-factor Daubert analysis is the starting point. As a practical matter, experts should assume courts and counsel will expect them to address these factors. But other reliability factors should be considered depending

---

107 Daubert, 509 U.S. at 594.
108 See 1 Modern Scientific Evidence, supra note 54, at 47.
110 Daubert, 509 U.S. at 594–95 (1993). Justice Blackmun noted “that scientists typically distinguish between ‘validity’ (does the principle support what it purports to show?) and ‘reliability’ (does application of the principle produce consistent results?).” Id. at 590 n.9. He explained that the Court’s concern is “evidentiary reliability—that is, trustworthiness.” Id. He concluded that in cases “involving scientific evidence, evidentiary reliability will be based upon scientific validity.” Id. (emphasis in original).
111 Id. at 597.
112 Id.
on the nature of the evidence. Drawing from court opinions both before and after *Daubert*, the advisory committee note to the 2000 amendment of Rule 702 mentions additional factors, which include whether the expert:

1. proposes to express an opinion based on research conducted independently of the litigation or based on research conducted for the purpose of testifying;
2. “has unjustifiably extrapolated from an accepted premise to an unfounded conclusion”;
3. “has adequately accounted for obvious alternative explanations”;
4. has been as careful in developing and presenting expert courtroom testimony as the expert would be in his or her regular professional work; or
5. relies on a field of expertise that produces reliable results for this type of expert opinion testimony.116

Because reliability is the touchstone for admissibility, the admissibility presentation should be tailored according to the theory and method of expertise and to the means by which it has been applied to the particular issue in the case at hand.

Two Supreme Court decisions following *Daubert* answered two critical questions. First, in *General Electric Co. v. Joiner*,117 the Court held that the standard for appellate review of a trial court’s application of Rule 702 and *Daubert* to expert testimony is abuse of discretion, thereby solidifying the trial judge’s strategic role as a gatekeeper for scientific evidence by insulating admissibility rulings from a more exacting standard of review.

Second, in *Kumho Tire Co. v. Carmichael*,118 the Court held that the *Daubert* gatekeeping function on evidentiary reliability applies not only to scientific expert testimony but also to expert testimony based on nonscientific technical or other specialized knowledge—in other words, all categories of expertise listed in Rule 702. The Court said that “whether *Daubert*’s specific factors are, or are not, reasonable measures of reliability in a particular case is a matter that the law grants the trial judge broad latitude to determine.”119

*Kumho* also underscored the importance of analyzing factors that help assess the reliability of the wide variety of expertise underlying expert testimony.120 Perhaps even more significant, *Kumho* re-emphasized the trial judge’s “task of ensuring that an expert’s testimony both rests on a reliable foundation and is relevant to the task at hand.”121 The “task at hand” analysis calls for a reliability

---

116 FED. R. EVID. 702 advisory committee’s note to 2000 amendment (citations omitted).
119 Id. at 153.
120 See id. at 149–53; 1 MODERN SCIENTIFIC EVIDENCE, supra note 54, at 71–72.
121 526 U.S. at 141 (quoting Daubert v. Merrell Dow Pharms., Inc., 509 U.S. 579, 597 (1993)).
determination of the expertise as applied to a specific issue in a given case.122
Courts and lawyers have been criticized for failing to follow this task-specific
approach.123

In 2000, Rule 702 was amended in response to Daubert and its case progeny:124

If scientific, technical, or other specialized knowledge will assist the trier
of fact to understand the evidence or to determine a fact in issue, a
witness qualified as an expert by knowledge, skill, experience, training,
or education, may testify thereto in the form of an opinion or otherwise,
if (1) the testimony is based upon sufficient facts or data, (2) the
testimony is the product of reliable principles and methods, and (3) the
witness has applied the principles and methods reliably to the facts of the
case.

The three elements added in 2000 at the end of the rule break down the
reliability analysis in terms of “sufficient facts and data,” “reliable principles and
methods,” and reliable application of the “principles and methods.” The advisory
note states that “[t]he term ‘data’ is intended to encompass the reliable opinions of
other experts” and can include hypothetical facts that are supported by the
evidence.125

The rule makes clear that judges must scrutinize not only principles and
methods used by the expert but also whether those principles and methods have
been properly applied.126 The application issue raises questions of the accuracy and
condition of the instruments that are used, compliance with protocols, and the
competence of both the individuals performing the technique and the people
interpreting the outcomes.127

122 See D. Michael Risinger, Defining the “Task at Hand”: Non-Science Forensic
Evidence After Kumho Tire Co. v. Carmichael, 57 WASH. & LEE L. REV. 767, 773–75, 778
(2000).
123 See D. Michael Risinger, Goodbye to All That, or A Fool’s Errand, by One of the
Fools: How I Stopped Worrying about Court Responses to Handwriting Identification (and
“Forensic Science” in General) and Learned to Love Misinterpretations of Kumho Tire v.
124 See FED. R. EVID. 702 advisory committee’s note to 2000 amendment.
125 Id.
126 Although Justice Blackmun wrote in Daubert that “[t]he focus, of course, must be
solely on principles and methodology, not on the conclusions that they generate,” 509 U.S.
at 595, the Court in General Electric Co. v. Joiner questioned this methodology/conclusion
distinction, explaining that “[a] court may conclude that there is simply too great an
analytical gap between the data and the opinion proffered.” 522 U.S. 136, 146 (1997). The
2000 amendment to Rule 702 made clear that reliability must be established not only for
the “principles and methods” underlying expert testimony but also their application “to the
facts of the case.” See FED. R. EVID. 702 advisory committee’s note to 2000 amendment.
Although the principles and methods underlying some forms of scientific evidence are so well established that a court may take judicial notice of reliability,\textsuperscript{128} scientific evidence based on novel theories and techniques is an unlikely candidate for judicial notice when it first offered at trial.\textsuperscript{129} Stable isotope analysis has been applied in a variety of fields for some time, and the \textit{Daubert} Court did not limit Rule 702’s application to novel scientific techniques.\textsuperscript{130} Moreover, the use of stable isotope ratio analysis in forensic investigation is relatively recent,\textsuperscript{131} and it has not been widely presented in U.S. courts. Accordingly, even if the stable isotope theory and analytical techniques are well established, we think trial judges should and will insist on a reliability showing contemplated under \textit{Daubert}/702.

3. Rule 703

Rule 703 provides:

The facts or data in the particular case upon which an expert bases an opinion or inference may be those perceived by or made known to the expert at or before the hearing. If of a type reasonably relied upon by experts in the particular field in forming opinions or inferences upon the subject, the facts or data need not be admissible in evidence in order for the opinion or inference to be admitted. Facts or data that are otherwise inadmissible shall not be disclosed to the jury by the proponent of the opinion or inference unless the court determines that their probative value in assisting the jury to evaluate the expert’s opinion substantially outweighs their prejudicial effect.

Rule 703 allows an expert to base an opinion on firsthand personal knowledge of facts or data and on facts or data in the evidentiary record “made known to the expert at or before the hearing.” Evidence law has long recognized these two bases for expert opinion.\textsuperscript{132} Rule 703 added a third basis: facts not necessarily admissible in the record “[i]f of a type reasonably relied upon by experts in the particular field in forming opinions or inferences upon the subject.”

Rule 703’s third basis should be understood in the context of the Rule 702/\textit{Daubert} framework. As the Advisory Committee on the Rules of Evidence explained, the 2000 amendment to Rule 702 “makes clear that the sufficiency of the basis of an expert’s testimony is to be decided under Rule 702” because

\begin{itemize}
  \item \textsuperscript{128} See Daubert v. Merrell Dow Pharm., Inc., 509 U.S. 579, 593 n.11 (1993); \textsc{Giannelli & Imwinkelried}, supra note 62, at 3–8.
  \item \textsuperscript{130} Daubert, 509 U.S. at 592 n.11.
  \item \textsuperscript{132} See \textsc{Fed. R. Evid.} 703 advisory committee’s note.
\end{itemize}
sufficiency analysis is integral to Rule 702’s “overarching” reliability requirement.\footnote{FED. R. EVID. 702 advisory committee’s note to 2000 amendment.} Rule 703’s reasonable reliance requirement applies to inadmissible facts or data and therefore “is a relatively narrow inquiry.”\footnote{Id.} As one court described it, “Rule 702 examines the expert’s testimony as a whole,” while “Rule 703 governs the inquiry into the reliability of particular data underlying expert testimony,”\footnote{United States v. W.R. Grace, 504 F.3d 745, 759 n.7 (9th Cir. 2010). The close relationship of these analyses is indicated by the Supreme Court’s citation of both Rule 702 and 703 to support this statement: “an expert is permitted wide latitude to offer opinions, including those that are not based on firsthand knowledge or observation.” Daubert, 509 U.S. at 592.} data that must be “of a type reasonably relied upon by experts in the particular field” but “need not be admissible in evidence.”\footnote{FED. R. EVID. 703.}

If there is an issue about an expert’s reliance on inadmissible data for an opinion, the trial judge must make a Rule 104(a) gatekeeper decision as to whether such reliance is reasonable under Rule 703,\footnote{See, e.g., United States v. Corey, 207 F.3d 84, 88–89 (1st Cir. 2000).} and, if the answer is yes, whether such data can be disclosed to the jury. In 2000, the last sentence of Rule 703 was added to prevent the rule from allowing pervasive and unrestricted disclosure to the jury of evidence, especially hearsay, that would otherwise be inadmissible under the evidence rules.\footnote{See FED. R. EVID. 703 advisory committee’s note to 2000 amendment. If the judge decides that an expert reasonably relied on inadmissible evidence and that this evidence can be disclosed to the jury, the judge would generally need to give a limiting instruction that such evidence can only be used to help the jury understand the expert’s opinion. See W.R. Grace, 504 F.3d at 759 n.7.}

4. Rule 706

Rule 706 provides, in part:

The court may on its own motion or on the motion of any party enter an order to show cause why expert witnesses should not be appointed, and may request the parties to submit nominations. The court may appoint any expert witnesses agreed upon by the parties, and may appoint expert witnesses of its own selection.

In his opinion concurring in part and dissenting in part in Daubert, Chief Justice Rehnquist expressed concern about the Court’s reliability factors and the trial court’s gatekeeping function as “impos[ing] on [judges] either the obligation or the authority to become amateur scientists in order to perform that role.”\footnote{Daubert, 509 U.S. at 600–01 (Rehnquist, C.J., concurring in part and dissenting in part).}
apparent response, Justice Blackmun noted that “Rule 706 allows the court at its discretion to procure the assistance of an expert of its own choosing.”¹⁴⁰ Given the range of scientific and technical issues that enter the courtroom, and given the responsibility that Rule 702 and Daubert place on judges to assess the validity of expert testimony, the Rule 706 option would seem to be a logical and practical course.

Judges have been slow to use independent court-appointed experts for assistance even though the judiciary has received much post-Daubert urging. Justice Breyer has “strongly encouraged” judges “to make greater use of their inherent authority . . . to appoint experts.”¹⁴¹ A court-appointed Rule 706 expert must advise the parties of his or her opinions, may be deposed by either party, and may be called to testify at trial.¹⁴²

Another option is for judges to appoint science advisors pursuant to their inherent authority,¹⁴³ a practice that has started to catch on in some district courts after Daubert.¹⁴⁴

5. Rule 403

Rule 403 provides:

Although relevant, evidence may be excluded if its probative value is substantially outweighed by the danger of unfair prejudice, confusion of the issues, or misleading the jury, or by considerations of undue delay, waste of time, or needless presentation of evidence.

The Daubert Court confirmed that Rule 403 permits the exclusion of evidence even if the expert testimony has cleared the other rules’ hurdles. Justice Blackmun quoted Judge Weinstein: “‘Expert evidence can be both powerful and quite misleading because of the difficulty in evaluating it. Because of this risk, the judge in weighing possible prejudice against probative force under Rule 403 of the present rules exercises more control over experts than over lay witnesses.’”¹⁴⁵

¹⁴⁰ Id. at 595 (majority opinion).
¹⁴³ See Ass’n of Mexican-Am. Educators v. California, 231 F.3d 572, 590–92 (9th Cir. 2000) (en banc); Reilly v. United States, 863 F.2d 149, 156 (1st Cir. 1988); see also Note, Improving Scientific Gatekeeping: Technical Advisors and Scientific Evidence, 110 HARV. L. REV. 941, 952–58 (1997).
¹⁴⁴ See, e.g., Hall v. Baxter Healthcare Corp., 946 F. Supp. 1387, 1392 n.8 (D. Or. 1996); see also 1 MODERN SCIENTIFIC EVIDENCE, supra note 54, at 120–23.
As with other forms of evidence, judges retain discretion under Rule 403 to balance the probative value of expert evidence against substantial dangers of unfair prejudice, misleading the jury, or the other Rule 403 factors. For example, courts have relied on Rule 403 to exclude polygraph evidence, citing its potentially overwhelming impact on the jury and the risk that it may mislead.\footnote{See United States v. Gilliard, 133 F.3d 809, 815–16 (11th Cir. 1998) (upholding district court’s exclusion of polygraph evidence under Rule 403 as potentially confusing); United States v. Pitner, 969 F. Supp. 1246, 1252–53 (W.D. Wash. 1997) (holding that even if polygraph evidence satisfied Rule 702, it “would still be excluded under Fed.R.Evid. 401 and 403. . . . [T]here is a substantial risk that the jurors will substitute the examination results for their own judgment.”); 1 MODERN SCIENTIFIC EVIDENCE, supra note 54, at 115–16.}

The task of expert testimony proponents is to assure the judge of the evidence’s probative value and to mitigate the potentially prejudicial, confusing, and misleading aspects.

6. Summary

The foregoing discussion describes the analytical framework under the rules to consider the admissibility of expert testimony. Further guidance can be found in Daubert and its progeny as well as the concerns expressed in the NAS Report on forensic science. We will now turn to application of these standards to stable isotope ratio analysis as courtroom evidence.

B. Application of Evidence Rules to Stable Isotope Ratio Analysis

1. Introduction

Although the basic science has been tested and applied in leading laboratories for many years, forensic use of stable isotope analysis has been relatively recent, and each forensic application may have its distinctive design, methods, and limits of confidence. When this evidence is offered in court, the challenge for the testifying expert is to identify and explain the basic principles and methods, the specific forensic application, and the limits of the particular application. The challenge for the lawyers is to elicit and contest this information. The challenge for the judge is to develop sufficient understanding of the evidence to make the admissibility decision based on the relevance and reliability standards.

A Rule 104(a) hearing applying the Daubert/702 standard to proffered expert testimony appears to lead to an all-or-nothing outcome: admissibility or non-admissibility.\footnote{See Andrew Jurs, Judicial Analysis of Complex & Cutting-Edge Science in the Daubert Era: Epidemiologic Risk Assessment as a Test Case for Reform Strategies, 42 CONN. L. REV. 49, 75 (2009).} But the intersection of law and science is rarely that straightforward.\footnote{See Susan Haack, Irreconcilable Differences? The Troubled Marriage of Science and Law, 72 L. & CONTEMP. PROBS. 1, 15–21 (2009).} The allure of science and the failure to understand its
limitations in resolving legal issues can lead to stronger conclusions than the applied science can support. Scientists recognize that their work is often contingent, subject to retesting and verification, and includes therefore an element of uncertainty. Perhaps this helps explain why the touchstone for admissibility—reliability—is a matter of degree in contrast to the categorical question of whether an expert witness is allowed to testify.

The Daubert Court addressed this issue: “[T]here are important differences between the quest for truth in the courtroom and the quest for truth in the laboratory. Scientific conclusions are subject to perpetual revision. Law, on the other hand, must resolve disputes finally and quickly.” Judges have some tools to screen the evidence and assist the fact-finder in ways that acknowledge scientific contingency. For example, the admissibility ruling can account for the uncertainties of science through limits on the scope and content of expert testimony, instructions to counsel, and instructions to the jury. But these corrective measures assume that judges have developed the capacity to understand the science and the limits of its applications. One purpose of the Rule 104(a) hearing is to help develop that understanding. Courts also rely on cross-examination and testimony from opposing experts to reveal weaknesses and enable the fact-finder to reach more accurate conclusions.

2. Inauspicious Debuts

Although widely used for investigative purposes, forensic stable isotope ratio analysis has not yet been offered as courtroom evidence on any regular basis. This Article anticipates that it will and aims to assist courts and counsel when it is. One attempt in 2003 in the United States and another in 2006 in the United Kingdom are noteworthy here.

An unreported federal district court opinion rejecting expert testimony on a particular application of isotope analysis demonstrates potential reliability challenges to this evidence. Mejdreh v. Lockformer Co. was a class action case alleging that defendant companies had exposed plaintiffs to trichloroethylene (TCE) in violation of federal environmental laws. The plaintiffs moved to exclude a defense expert from testifying about an isotopic comparison between volatile organic compounds taken from plaintiff locations and the TCE found on defendant Lockformer’s site. The defense expert compared the isotopic ratios of

---

149 See FELDMAN, supra note 5, at 140.
152 509 U.S. at 596–97.
153 See FELDMAN, supra note 5, at 147.
155 Id.
carbon and chlorine in soil and groundwater samples.156 He concluded that the differences in the chlorine isotope ratios showed that the groundwater in the plaintiff areas did not originate from the defendant Lockformer source.157

The judge concluded that this evidence did not meet the reliability requirements of Rule 702 and Daubert.158 Relying in part on the plaintiff’s stable isotope ratio expert, who challenged the defense expert’s work, the court explained that the methodology, by using large water samples, departed from a peer-reviewed procedure. The method had not been tested or subjected to peer review or publication, appeared to have a high potential error rate, and was not generally accepted in the relevant scientific community.159 A further problem was the expert’s disregard of counter-findings in his research that showed matching carbon isotope ratios from the comparison areas.160 The defense expert also had recanted an opinion about the origin of the TCE found on the plaintiff sites.161

A similar disposition occurred a few years later in the United Kingdom. Isotope analysis surfaced in Ullises Shipping Corp. v. Fal Shipping Co. The owners of a tanker sued the time charterers of the vessel. The tanker had been detained, confiscated, and sold at public auction in the United Arab Emirates based on the vessel holding Iraqi oil in violation of United Nations sanctions applicable to Iraq.162 A central issue at trial was whether the oil on board originated in Iraq. An independent testing laboratory performed four levels of analysis of samples taken from the tanker, including a process using biomarkers and isotope analysis. The lab concluded that the tanker probably contained Iraqi oil.163

The court said a reliability evaluation of these analyses was difficult because no witness from the lab testified. Experts for the parties argued for and against the reliability of the testing methods. The court was not persuaded that the lab’s methodology provided reliable conclusions about the likely origin of the oil, in part because the oil produced in southern Iraq comes from the same geologic formation as oil produced in southern Iran and Kuwait, and there was no reference sample used from Iran or Kuwait. The defense expert also pointed out that seawater contamination in the samples would affect the isotope analysis and that the lab did not follow the methodology set forth in a paper it cited on oil spill fingerprinting.164

In both Mejdruch and Ullises Shipping, the courts did not question the general principles and methods of stable isotope ratio analysis. They both found that the isotope evidence was not sufficiently reliable based upon shortcomings in the

156 Id.
157 Id.
158 Id. at *3.
159 Id.
160 Id.
161 Id. at *2.
163 Id. at 546–50.
164 Id. at 551–54.
specific applications of the science in those cases. The decisions therefore stand as examples of the need for courts under Daubert and Rule 702 not only to evaluate the reliability of the principles and methods of scientific evidence but also to determine whether the expert has applied the principles and methods reliably to the case at hand. Neither case challenges the basic stable isotope ratio theory or technique.

Stable isotope ratio expert testimony has been involved in other cases. In addition to the Floyd Landis doping arbitration, one of the most prominent was the trial of the six men accused of plotting to set off explosions in 2005 in the London subway system. An isotope expert was called to rebut the lead defendant’s claim that he diluted the hydrogen peroxide used in the bombs with London tap water. The expert said this was impossible based on his comparison of the isotopic profiles of the bomb residue and samples of London tap water.

Against this modest backdrop of attempted use of stable isotope forensic evidence in court, we turn to a discussion about admissibility issues regarding stable isotope ratio analysis.

3. Rule 104(a) and the Daubert/702 Analysis

Formal implementation of the judge’s gatekeeper role is the admissibility decision under Rule 104(a). Rule 104(a) does not require a pretrial hearing, but when novel applications of scientific evidence are proposed for admissibility, as in the early years of DNA identification evidence, parties can be expected to seek a pretrial ruling through a motion in limine. A pretrial Daubert hearing would generally be regarded as a reasonable and advisable step.

We expect courts will conduct Daubert pretrial hearings on the admissibility of stable isotope evidence. Whether or not such a hearing is held, the Rule 104(a) admissibility determination depends on how the stable isotope ratio analysis measures up under the Daubert/702 framework and related rules. Stable isotope evidence is based on the application of principles and methods that are amenable to testable hypotheses. Accordingly, we will focus our attention on the Daubert factors, which contemplate scientific methods of analysis. We will consider other reliability considerations as well, which we think are at least equal in importance to the traditional Daubert factors.

4. Rule 702 Analysis

The range of stable isotope forensic applications is so wide and varied that a general discussion applying Rule 702 standards will only provide threshold

---


assistance for a particular case. Even the methods of stable isotope analysis will vary from application to application, thereby requiring a reliability showing for a particular methodology. We are mindful of Rule 702’s requirement to show that stable isotope analysis has been reliably applied to the facts of the specific case.

For the foregoing reasons, and in acknowledgement of the Supreme Court’s admonition to address the Daubert analysis to the “task at hand,” we will include discussion of the application of stable isotope analysis to the investigation of the anthrax attacks of 2001 (referred to as Amerithrax), which the Federal Bureau of Investigation (FBI) recently closed. In this case of a bio-weapons attack, spores from Bacillus anthracisis (anthrax) were contained within letters mailed to news media offices across the United States and to two U.S. Senators. As a result of the anthrax attack, five people were killed and an additional seventeen individuals were afflicted and survived. The chemical and biological analyses of spores recovered from the letters sought to determine the origins of the spores and the identity of the perpetrator(s) behind the attack. Among the many diverse tests employed were stable isotope ratio analyses of spores, culture medium, water used along with culture medium, and envelopes.

(a) Relevancy

The relevance and helpfulness of scientific evidence depends, of course, on the facts that are consequential to a particular case. Rule 702 asks whether expert evidence “will assist the trier of fact to understand the evidence or to determine a fact in issue.” Stable isotope evidence can help the fact-finder determine the likelihood that two or more evidence specimens are consistent with having originated from a common source, or the likelihood that a specimen is consistent with having been associated with a particular geographic location. If the likelihood of an evidence specimen having been associated with a specific location is low, stable isotope ratio analyses can, in many cases, provide data on the likelihood that an evidence specimen could have been associated with other geographic locations. If, in any of these instances, the isotope information would help the fact-finder in making the determination of a consequential fact more or less probable, the stable isotope evidence should meet the “assist the trier” test of Rule 702.

Five applications of stable isotope ratio analysis were relevant to the Amerithrax case: (1) identification of the culture medium most likely to have been used to culture the anthrax; (2) characterization of the geographic region(s) most likely to have been associated with the culturing of the anthrax; (3) similarity comparisons among spore specimens recovered from different anthrax-containing

---

169 FED. R. EVID. 702.
170 FED. R. EVID. 401.
letters; (4) similarity comparisons between the anthrax spore evidence and the
different culture medium used to culture bacteria (including anthrax); and (5)
similarity comparisons of the cellulose composition of letters used to mail the
spores. 171

The following questions address two fundamental aspects of stable isotope
analysis relevance: similarity and location.

First, were the Amerithrax specimens isotopically indistinguishable from each
other and therefore more likely to be related and be consistent with having a
common origin?

Second, were the isotopic compositions of the anthrax spore specimens
consistent with bacterial spores that had been cultured in a particular geographic
region? Additionally, based on the stable isotope ratio measurement, could some
geographic regions be excluded as origin-of-culture possibilities?

Third, were the isotopic compositions of the anthrax spore specimens
consistent with observations expected for the culture of the bacteria grown using a
particular culture medium? Additionally, based on the stable isotope ratio
measurement, could some culture medium be excluded from further consideration
as culture medium possibilities?

These questions presume that reliable data could be obtained and that
appreciable stable isotope ratio variations existed in culture water and culture
medium to allow a meaningful interpretation of the data.

Answers to these questions should achieve a more probable understanding of
the origin and location of the anthrax recovered from the 2001 letters and should
assist a trier of fact. If stable isotope analysis assists to provide such answers, then
it should be considered relevant and helpful under Rule 702.

(b) Qualifications

Rule 702 provides that the stable isotope testifying expert must be “qualified
... by knowledge, skill, experience, training, or education.” Stable isotope ratio
analysis is a specialty field, and experts in this field become so through training
and experience with stable isotope ratio measurement techniques. Often the
strongest experts are familiar with both the stable isotope methodology and have
the chemical or biological expertise to apply the technique to the specific
applications. Examples of the diverse expertise areas might include diet
reconstruction, pollution, explosives, food, and drugs.

In the case of stable isotope analysis, the judge should consider both the
ability of the expert to comment on or interpret the stable isotope methodology and
the limitations to the quality of the data presented. The judge should further assess
the expert’s ability to address the relevance of stable isotope ratio data to a

171 See Juske Horita & Arpad A. Vass, Stable-Isotope Fingerprints of Biological
Agents as Forensic Tools, 48 J. FORENSIC SCI., 1, 4–5 (2003); Helen W. Kreuzer-Martin &
Kristin H. Jarman, Stable Isotope Ratios and Forensic Analysis of Microorganisms, 73
particular topic. Stable isotope ratios have been applied extensively across the sciences, but not all stable isotope experts may be capable of interpreting isotope ratio data in such diverse fields as geochemistry, biology, food quality and origin, and biochemistry.

A stable isotope expert who would be qualified to testify about analyses used in the anthrax investigation should have training and experience in one or more of the relevant fields of biology and chemistry where this technique is practiced. The expert should have established credentials through advanced graduate education and peer-reviewed publications. There should be a relationship between the expert’s publication and laboratory-based qualifications and the nature of the evidence being considered.

The most beneficial expertise would be experience with the techniques of handling and analyzing this particular type of biological evidence, ability to evaluate the data on the basis of a testable hypothesis, familiarity with the strengths and any weaknesses of specific methodologies involved in the isotope analyses of the evidence, and the skills to comment on the reliability and error rate associated with the analyses. The expert should be able to address both the relevance and strengths of the isotope measurements and also the necessary constraints when interpreting the data.

(c) Reliability

The heart of the Rule 702/Daubert analysis is the evaluation of reliability, which involves three inquiries: sufficiency of the facts and data, reliability of the principles and methods, and reliability of the application of the principles and methods to the facts of the case.

i. Facts and Data

This part of Rule 702 calls for careful consideration of what the facts and data are and whether they are sufficient to support the expert’s testimony. For stable isotope evidence, the analysis will turn on (1) whether the samples meet the level of quantity and quality for accurate measurement; (2) whether the mass spectrometry or other appropriate instrumentation can produce interpretable results; (3) whether the databases are sufficiently populated to permit meaningful comparisons and reasonable conclusions; and (4) whether the mathematical models and relationships used to predict geospatial patterns are sufficiently differentiated to permit meaningful comparisons and reasonable conclusions.

The samples and data produced through the isotope measurement and comparison processes must be adequate to permit scientific interpretation. Accordingly, the sufficiency determination should take account of interpretation constraints. The expert should assist the court by identifying and explaining the strengths and limits associated with interpreting the stable isotope ratio data. Courts should focus on the closely intertwined issues of the differentiation of the
data and the interpretability of the data in analyzing Rule 702’s “sufficient facts or data” requirement. The following questions and discussion address these issues.

First, are the differences in stable isotope ratio data sufficiently different in value that two or more specimens can be distinguished? The expert can assist the court with known information about the role of chemical complexity and heterogeneity affecting stable isotope ratio values of materials that are similar to the evidence under consideration. The expert can further assist the court with information on the ranges of stable isotope ratio values to be expected for material(s) that are similar to the evidence under consideration. The court may choose to recognize the value of established databases that contain observations quantifying the nature and extent of stable isotope ratio variations among materials/organisms related to the specimen evidence presented to the court.

Second, is there an accepted scientific basis to interpret and explain the stable isotope ratio measurements that are proposed to be offered as evidence in court? If so, to what extent can this basis be used to inform the court about the appropriateness of interpretation possibilities? The possibilities could include likely relatedness of specimens, likely region(s) of geographic origin(s) of one or more specimens, likely factors influencing the production of specimens, and/or likely relatedness between two or more materials such as starting materials and a finished product. The scientific basis to make each of these interpretations should be established through one or more peer-reviewed publications. The expert should address the strengths and limits of interpreting the stable isotope ratio data in these contexts.

As applied to the Amerithrax example, consider data (facts) and interpretation as two distinct but related issues as described above.

First, are there established methods to obtain data that would allow an expert to make interpretable measurements on Amerithrax spores? The answer may depend on the ability to isolate the spores from background matrix materials that could affect the measurement of an isotope ratio. For instance, the preparation and analysis method should be able to purify the spore from the growth medium to a sufficient level that trace amounts of residual growth medium do not affect the measurement of the spore isotope ratio. That is, are acceptable methodologies in place to provide sufficient confidence that repeatable and meaningful isotope ratio data can be obtained from the evidence specimen(s)?

In the Amerithrax case, sufficient information appeared to be available to establish data accuracy and reliability. As to the ability to measure reliable stable isotope ratio values for bacteria and bacteria spores, numerous publications on this topic have been published as peer-reviewed studies in the microbiological forensic science literature. A recent review of this specific topic by Kreuzer-Martin and Jarman describes no less than ten independent scientific studies involving forensic-related measurements of stable isotope ratios in bacteria and bacteria spores.172

This issue is distinguishable from whether isotope ratios can be measured using instrumentation commonly used for isotope ratio analyses in the biological

---

172 Kreuzer-Martin & Jarman, supra note 171.
and chemical fields of study. Stable isotope ratio analyses are now a commonly accepted measurement. For Amerithrax evidence specimens, hydrogen, carbon, nitrogen, and oxygen isotope ratios were measured using an isotope ratio mass spectrometer. This instrumentation has been in place and accepted as an appropriate measurement instrument for more than five decades.\textsuperscript{173} Preparation and analysis methodologies to analyze bacterial spores were refined from accepted practices in the field. In a series of peer-reviewed publications, Kreuzer-Martin and colleagues demonstrated how meaningful stable isotope ratio data could be acquired for bacteria and spores cultured under different combinations of source water, culture medium, culture practice, and geographic location.\textsuperscript{174} If the preparation and analysis methodologies are accepted, then a reliable stable isotope ratio measurement can be obtained for the evidence specimen(s) that could be used for forensic investigation and possibly introduced as evidence in a court.\textsuperscript{175}

Second, are there established methods to interpret the stable isotope ratio data? At this stage the expert should be able to identify the published experimental studies providing a framework for interpreting the data; describe the breadth, relevance, and limitations of any databases that are relevant to interpreting the data; and explain the strengths and limitations of mathematical/geospatial models and other mathematical relationships that are essential to interpreting the stable isotope ratio data. The expert should be able to evaluate testable hypotheses and provide a likelihood of accepting or rejecting these hypotheses.

\textit{ii. Principles and Methods}

The principles and methods of stable isotope analysis have been described earlier in this paper. We will not repeat those descriptions here, but will focus on the reliability of those principles and methods. We will address the Daubert reliability factors and suggest some other reliability considerations.

\textit{a. Testability—}Stable isotope evidence is based on the application of principles and methods resulting in stable isotope ratio data (facts) that are then amenable to testable hypotheses.\textsuperscript{176} Although the conclusions reached from interpreting stable isotope ratio data can, in some limited circumstances, be influenced by the methodology, any rare influence is more likely to be based on

\textsuperscript{173} See \textit{SHARP}, supra note 22, at 1–39; Budzikiewicz & Grigsby, supra note 12.


\textsuperscript{175} See Kreuzer-Martin & Jarman, supra note 171.

\textsuperscript{176} See Benson et al., supra note 131.
the sample-preparation methodology used than on the instrumentation methodology. Extensive research on analyses of hydrogen, carbon, nitrogen, and oxygen isotope ratios has been conducted in respected laboratories at many universities in the United States and other countries.177

The peer-reviewed literature on this subject is extensive and confirms this measurement approach as a reliable analytical instrumentation method applied to many different fields of science and medicine.178 Testability (falsifiability) of the data should be examined to eliminate the possibility that a sample-preparation methodology can influence the data and therefore influence the conclusion. Where multiple established and commonly used sample-preparation techniques exist for stable isotope analyses of evidence, the expert should establish the reliability and validity of one method versus the other, the potential of differences in the testability of data from these approaches, and an evaluation of the likelihood that one methodology versus the other could lead to a different conclusion.

b. Peer Review and Publication—The testing of stable isotope ratio principles and methodology and the reduction of systematic and random errors are achieved through research, the peer-review process, and the availability of accepted methods and approaches in the open literature. The general utility of stable isotope ratio analysis as a valid, accurate, and reliable scientific measurement tool has been established for decades in diverse disciplines within natural, physical, and social sciences.179 By extension, the utility of isotope analysis as a valid and reliable forensic approach has been established in a wide number of peer-reviewed publications that have appeared over the past two decades in different scientific journals, including the Journal of Forensic Sciences, which is the primary publication outlet of the American Academy of Forensic Sciences.180

The application of stable isotope ratio analysis to a particular case may require the use of specific sample-preparation methodologies with the general stable isotope ratio measurement instrumentation. Although many of the sample-preparation methodologies are now routinely accepted as valid, accurate, and reliable, the expert should establish that the sample-preparation methodology applied in a particular case is valid and reliable, and that a similar methodology has appeared in a peer-reviewed publication.

c. Error Rate—As in all areas of high-precision instrumentation science, measurements to obtain stable isotope ratio data and their interpretation are subject to various sources of error. These are not proficiency errors; they are the small potential errors associated with the accuracy and precision of an analytical instrument.181 The error rates of stable isotope ratio measurements tend to be small because specimens are analyzed along with well-accepted international standards

---

177 See id.
178 See supra note 174.
179 See Jochen HoeFs, Stable Isotope Geochemistry (3d ed. 1987); Sharp, supra note 22.
180 See, e.g., Casale et al., supra note 36.
181 See Benson et al., supra note 131.
multiple times in each analytical run. In addition, an analytical run includes blind reference materials and additional reference materials appropriate to the kinds of evidence specimens being analyzed, which help to improve the analytical quality of the acquired data.

It is nonetheless important to recognize potential error rates associated with stable isotope ratio data. Small random errors can arise because of heterogeneity in the specimens being analyzed. That is, because the stable isotope ratio analysis uses such small sample sizes, it is possible that heterogeneity (degree of mixture) of a sample could affect the accuracy of a measurement. Error arising from heterogeneity can be minimized by increasing the number of replicate analyses of the specimen.

As applied to the Amerithrax case, potential random errors were small based on replication of stable isotope analyses, variations in and among culture medium (within batch heterogeneity and batch-to-batch variation), and water used to culture bacteria. Systematic errors in the analyses of stable isotope ratio data would likely be small if they were based on incorrect or biased sample-preparation methodologies. This latter point would relate to the appropriate application of stable isotope measurements rather than error rates associated with a measurement.

Another type of potential error can occur with the interpretation of stable isotope ratio data. These potential errors concern issues of statistical error and assessment of the data. Two approaches are considered for interpretation of hypothesis testing of stable isotope data: (1) significance tests and (2) likelihood ratios.

1. Significance tests—Significance testing of data associated with specimens and/or between data and expected values in models uses both parametric and nonparametric statistics, as appropriate for the individual data. These statistical tests are used as the independent evaluation method to accept or reject a hypothesis under consideration. The significance testing error rate can theoretically increase as greater complexity is considered. This approach should be considered when interpreting stable isotope ratio data, but such an error rate is typically not an issue with biological organisms. In biological systems, such as bacterial spores, the variation in stable isotope ratio data among analyses will likely exceed the analytical error if there was heterogeneity, but not by much. Therefore, the potential propagation of the measurement errors does not typically exceed twice the error associated with an individual observation. For stable isotope data, these errors tend to be small.

In the Amerithrax case, such errors were likely to have been small with, for example, carbon isotopes in bacteria spores having a precision of ±0.2‰ relative to

---

182 See Kreuzer-Martin & Jarman, supra note 171.
183 See Thomas & Jamin, supra note 15.
184 See Kreuzer-Martin & Jarman, supra note 171.
185 See id.
187 See Benson, supra note 131.
an overall 95% confidence interval precision of ±0.4‰ that includes specimen heterogeneity.\textsuperscript{188}

Another potential source of significance testing error can be associated with the interpretative precision of models (e.g., linear models and geospatial models) that relate stable isotope data to a known pattern or a geospatial distribution. These interpretative errors may exceed random analytical errors. The magnitude of an interpretive error will depend on the size and representativeness of any database used for assignment comparisons, and on the strengths and reliability of models used to interpret the data in a geospatial context. These systematic errors are reduced by considerations of larger and more representative databases and by development and testing of new geographic models used to interpret the data.\textsuperscript{189}

Kreuzer-Martin and Jarman,\textsuperscript{190} as well as Jarman and colleagues,\textsuperscript{191} considered these issues with respect to the Amerithrax case and concluded that the errors were small and very unlikely to influence interpretation of the data. In addition, Jarman and colleagues provided a data interpretation framework based on Bayesian statistics that lends itself well to error rate analyses, considering stable isotope ratio analyses in conjunction with other chemical analyses of the evidence.\textsuperscript{192} In each case, the errors are recognized here for the sake of completeness, but when incorporated individually and/or collectively, the overall error does not significantly affect the strengths of conclusions that can be derived from data through significance testing.\textsuperscript{193}

2. **Likelihood ratios**—An emerging error rate analysis approach for quantitative scientific data is the application of likelihood ratios in addition to significance testing. A likelihood ratio analysis considers the data in the context of two contrasting hypotheses.\textsuperscript{194} The likelihood ratio is the ratio of the probability that hypothesis one is correct relative to the probability that hypothesis two is correct. This can be viewed as the ratio of the true positive rate to the false positive rate.\textsuperscript{195}

In the most straightforward approach, the application of likelihood ratios to the Amerithrax case would be, “What is the ratio of the probability that the anthrax spores originated from region 1 relative to the probability that they originated from any geographic region other than region 1?” The result would be a number that varied from near zero to a large number in excess of a million. At a likelihood value of 1, there is an equal probability that the “recovered” and “control” evidence did or did not share a common origin. The lower the likelihood ratio is from one, the greater the probability that “recovered” and “control” evidence do

\begin{itemize}
  \item[\textsuperscript{188}] Kreuzer-Martin, *Part 2*, supra note 174.
  \item[\textsuperscript{189}] See Kreuzer-Martin & Jarman, *supra* note 171.
  \item[\textsuperscript{190}] See id.
  \item[\textsuperscript{191}] See id.
  \item[\textsuperscript{192}] See id.
  \item[\textsuperscript{193}] See id.
  \item[\textsuperscript{194}] See Aitken & Lucy, *supra* note 186.
\end{itemize}
not share a common origin. In contrast, the more the likelihood ratio exceeds one, the higher the likelihood that the “recovered” and “control” evidence do share a common source.196

Farmer and colleagues applied likelihood ratio analyses to stable isotope ratio analyses of fifty-one specimens of white paint. The authors arbitrarily selected one of the paints as “recovered” and another as “control.” They were able to calculate three results: the likelihood that “recovered” and “control” were from a common source, the likelihood of a false positive, and the likelihood of a false negative. This approach can be applied to all stable isotope ratio data, complementing traditional statistical testing approaches.197

d. Standards—In stable isotope ratio analyses, standards play an important role in enabling the investigator to determine the accuracy and precision of the measurements and to compare similar analyses conducted in different laboratories. This Daubert factor is therefore especially important to the reliability analysis of stable isotope evidence.

International standards have been recognized and made available to laboratories.198 In each analytical run, it is typical for the laboratory to include working reference materials calibrated against international standards,199 “blind” quality-control samples of known isotope ratio values, and replicate samples of the same piece of evidence (assuming sufficient material exists).200 As we discussed in Part I of this Article, laboratories participate in round-robin exchange tests to ensure that different laboratories conducting the same stable isotope analysis will reach the same results.201

If the evidence specimen must be pretreated in any way prior to isotope ratio analysis, the principle of identical treatment should always be followed. That is, reference materials and “blind samples” should undergo the same pretreatment as the evidence specimen. By this means, it is possible to know the accuracy, precision, and error rate of an analytical run associated with analysis of evidence.202

The introduction of stable isotope evidence into a court proceeding requires confirmation of several practices that are common to the introduction of other scientific data. Two standard practices establish the quality and accuracy of the stable isotope ratio data.

First, acceptable technology and methodology must be used in the analysis of the evidence. The former can be accomplished by use of instruments that have sufficient precision and that are used in this field of science today for stable isotope

196 See N. Farmer et al., Stable Isotope Analysis of White Paints and Likelihood Ratios, 49 SCI. AND JUST. 114 (2009).
197 See id.
198 See M. Boner & H. Förstel, Stable Isotope Variation as a Tool to Trace the Authenticity of Beef, 378 ANALYTICAL AND BIOANALYTICAL CHEMISTRY 301 (2004).
199 See, e.g., N. Farmer et al., supra note 196.
200 See Thomas & Jamin, supra note 15.
201 See id.
202 See id.
ratio measurements. The laboratory analyzing the evidence specimens must have written protocols to ensure that the instrumentation is in correct operating condition. These written practices must have been in effect at the time the stable isotope ratio measurements were conducted on evidence specimens.

These quality assurance and quality control protocols are typically written in a step-by-step fashion. Included in these practices is the use of laboratory reference gases to convert the electrical signals from the instrument associated with analysis of a specimen into raw stable isotope ratio values. The laboratory reference gases that are part of the analysis process must have been calibrated against internationally accepted reference standards (often referred to as primary standards). This can be accomplished by using methods appropriate to the evidence in hand and accepted in the scientific discipline through peer-reviewed publication(s). These methodological protocols are typically written in a step-by-step fashion for the analytical technician and based on peer-reviewed publication(s) to ensure quality, accuracy, and repeatability.

Second, acceptable (1) quality control practices and (2) reference and calibration materials must be used in the analytical run associated with the analyses of all evidence specimens. The former is achieved by determining whether the analytical run (typically up to 100 analyses that include specimens, calibration materials, and blind reference materials) met the threshold requirements established by written laboratory protocols and data templates that are used to correct or normalize the data from instrument voltage signals into stable isotope ratio values on an international scale.

The latter is achieved by use of laboratory reference materials that have been calibrated against internationally accepted reference standards. The known stable isotope ratio values of laboratory reference materials are used to make final calibration adjustments to the raw stable isotope ratio data of a specimen or blind reference sample. Most laboratories will also include blind reference materials that have been calibrated against established reference materials as part of the analytical run. This serves as an independent check on the accuracy of the measurements in the analytical run. The reference materials and blind reference materials should be selected to be as similar in composition as is feasible to the specimen(s) being evaluated for court presentation.

These practices are used to ensure that the stable isotope ratio analyses can be repeated in the same manner in another laboratory and obtain similar results.

203. General Acceptance—Stable isotope analyses are commonly used to trace the origins and movements of biological and non-biological materials in fields such as environmental science, ecology, atmospheric science, anthropology, food science, and oceanography. The technique to measure stable isotope ratios of light elements by use of an isotope ratio mass spectrometer is widely accepted.

203 See, e.g., Ehleringer et al., Isoscapes, supra note 39, at 369–82; Stable Isotopes in Ecological Research (P.W. Rundel et al. eds., 1989); Stable Isotopes in Ecology and Environmental Science (Robert Michener & Kate Lajtha eds., 2d ed. 2007); Martin et al., Interpretation of Combined, supra note 15, at 62–63.
within the scientific community. The extension of isotope analyses into forensic sciences is robust, and results should be readily interpretable by forensic scientists and the courts.

Although the stable isotope analysis technique is accepted and well-respected for scientific investigation, its application to the forensic analysis of specific types of biological or nonbiological materials may be questioned in some unusual situations that may therefore require peer-reviewed publication in the scientific literature. This concern illustrates the general point that reliability of a theory or method must be established for the particular application in the case at hand. For stable isotope ratio analysis, this concern relates especially to evidence that can be lost over time. Is the material being analyzed stable or not over time? Put another way, does the evidence persist, or does it change or disappear over time (such as by evaporation)?

Examples of stable materials appropriate for stable isotope analysis are water in a closed container or a non-evaporating material. Examples of substances that can still be analyzed but possibly challenging to interpret are materials that evaporate over time unless sealed in a container. Two examples of such materials would be alcohol in an open jar allowed to evaporate continuously, or triacetone triperoxide (TATP) allowed to decompose or sublimate in an open room. The latter is the explosive used in the 2005 London subway bomb attacks. With evidence of the physical processes that influence the isotope ratios of evaporating or decomposing materials, these materials can still be interpreted in forensic cases.

f. Other Reliability Considerations—The Supreme Court stressed in Daubert and Kumho that the factors identified in Daubert for the reliability determination are illustrative. The reliability analysis should be tailored to fit the principles and methodology underlying the expert testimony and the application of those methods to the particular case. Although our discussion of the five Daubert reliability factors is important to assess stable isotope ratio evidence, additional or refined considerations should also be evaluated, such as whether:

- The evidence analyzed was in a stabilized form and did not change in amount or abundance between the time that evidence was obtained and analyzed.
- The principle of identical treatment was applied to evidence, reference materials, and quality-control samples.
- The instrumentation used was appropriate and capable of achieving the high precision required for an isotope ratio analysis at natural abundance levels.

204 See I.T. PLATZNER, MODERN ISOTOPE RATIO MASS SPECTROMETRY (1997); J. Thomas Brenna et al., High-Precision Continuous-Flow Isotope Ratio Mass Spectrometry, 16 MASS SPECTROMETRY REVIEWS 227 (1997); Meier-Augenstein & Liu, supra note 22, at 150; Muccio & Jackson, supra note 12, at 213.

The instrument had achieved quality assurance requirements established by the laboratory and found to be broadly acceptable by the scientific community.

- The stable isotope ratio analyses of quality-control samples were within the range of acceptable values.
- The variances among replicate analyses of stable isotope ratio analyses of the evidence were within acceptable ranges.

An expert who is interpreting stable isotope ratio evidence should consider the following before reaching a conclusion:

- An understanding of the extent, if any, that heterogeneity of composition of the evidence plays a role in interpreting the stable isotope ratio values. Significant sample heterogeneity must be considered in the final interpretation of the data.
- When two or more evidence samples are compared, knowledge that the evidence has been treated similarly since it was obtained. If evidence specimens were not handled in an identical fashion, then the potential consequence of this variance must be considered when interpreting the data.
- When interpreting stable isotope ratio observations where two or more evidence specimens are compared, an understanding of the typical variance affecting the precision of the isotope ratio measurement (e.g., handling, purity of preparation, and reductions in overall measurement precision associated with multiple preparation steps prior to isotope ratio analyses).
- An understanding that the conclusions to be reached may only answer whether the observations are “consistent with” or “not consistent with” a specific hypothesis or question. At this level, it may be appropriate to ask the expert to assign probabilities to each interpretation based on the facts provided to the expert.

iii. Application

Rule 702 requires the trial judge to “scrutinize not only the principles and methods used by the expert, but also whether those principles and methods have been properly applied to the facts of the case.”206 The “task at hand” analysis emphasized in Daubert and Kumho calls for a reliability assessment of the specific application. Even if courts have previously accepted a scientific method as a valid basis for expert testimony, the reliability of the method used in a given case must be assessed by the reasonableness of applying it to the facts of the case and by the validity of how conclusions are drawn from the data.207 We already have used the

---

206 FED. R. EVID. 702 advisory committee’s note.
207 See Kumho Tire Co., 526 U.S. at 153–54; Hendrix v. Evenflo Co., 609 F.3d 1183, 1195 (11th Cir. 2010).
Amerithrax case to illustrate aspects of the reliability analysis. We will expand here on the application of stable isotope ratio analysis to the anthrax attacks.

A court presentation in the Amerithrax case would require a showing that the stable isotope ratio analyses were based on and consistent with peer-reviewed and validation studies in the field that demonstrate the testability and reliability of the principles and methods applied. It would also call for demonstrating a sufficiently small risk of error to allow for reliable conclusions. Reliability also would depend on a showing that laboratory protocols were followed, that the instrumentation was properly calibrated and used, that measurements were accurately recorded and interpreted, and that samples were properly handled. The level of certainty for conclusions about common source of origin and location of origin would need to be supported as set forth earlier.

To understand this application, consider the life cycle of bacteria. The spores that were mailed in the Amerithrax case represented the dormant or resting stage of bacterial growth.208 Spores are a cell form produced by some, but not all, bacteria in the growth-phase transition from active vegetative growth, nutrient-rich bacteria cultures to a much slower growth. The transition to a dormant spore form occurs as medium becomes depleted in the essential nutrients to maintain the actively dividing bacteria form. During all stages of the bacteria growth cycle, the bacteria take up nutrients from their external environment (known as medium when bacteria are actively cultured).

The application of stable isotope methodology allows for the establishment of reliable patterns and for the testing of specific hypotheses. During bacteria growth, the compounds that are part of each bacterium are built within the bacteria cell based on the uptake and chemical conversion of nutrients derived from the external medium (e.g., water, salts, carbon source, nitrogen source). “You are what you eat” is an appropriate phrase to describe the stable isotope composition of both microbial and animal systems. There are precise, predictable, and reliable relationships between the carbon and nitrogen isotope ratios of the nutrient medium and those of the cell walls, proteins, and carbohydrates that characterize the bacteria.209 Thus, from measurements of the carbon and nitrogen isotope ratios of a bacterial spore, it is possible to reconstruct the carbon and nitrogen isotope ratios of the growth medium used to culture the bacteria.

The studies by Kreuzer-Martin and colleagues additionally showed that different growth medium often had distinctive carbon and nitrogen isotope ratios. Thus, from analyses of the carbon and nitrogen isotope ratios of bacteria spores, it is possible to reconstruct or predict the carbon and nitrogen isotope ratios of the growth medium and, in some cases, to predict the specific culture medium used to produce the bacterial spores. From such analysis, relevance is clear. Carbon and nitrogen isotope ratio analyses allow the investigator to compare among samples to see whether they possibly share a common growth medium and to predict the

208 See AMERITHRAX INVESTIGATIVE SUMMARY, supra note 168, at 2, 13–16.
growth medium type used to culture the bacteria. This in turn could provide a significant lead to determine whether the growth media in one or more laboratories are consistent with or not consistent with the growth media that were likely to have been used to cultivate the anthrax.

During bacterial growth, hydrogen and oxygen atoms from water in the culture medium are incorporated into the proteins, carbohydrates, and outer cell wall complex of the bacteria being cultured. There are very predictable relationships between the hydrogen and oxygen isotope ratios of bacteria spores and the hydrogen and oxygen isotope ratios of the water used to culture the bacteria. The geographic patterns of hydrogen and oxygen isotope ratios of local waters differ across a continental landscape and exhibit predictable and reliable spatial patterns. This observation allowed Kreuzer-Martin and colleagues to determine whether the isotope ratios of a spore were consistent with or not consistent with growth using a particular water source in different geographic regions of the United States. The predictability of the pattern was tested by culturing the same bacteria in different parts of the United States using the same medium and reliably reconstructing the region from which the different bacteria had been grown. The combination of experimental studies in the laboratory and then field validation of the technique to identify different geographic regions is possible for most stable isotope studies.

The methods used in the anthrax analysis included five sets of key observations, which established that:

1. The hydrogen and oxygen isotope ratios of Bacillus spores were distinctly, linearly, and predictably related to the hydrogen and oxygen isotope ratios of the local water source used to culture the bacteria.
2. There were distinctive and predictable spatial zones of hydrogen and oxygen isotope ratios of local water sources and geographic regions across the United States.
3. The carbon and nitrogen isotope ratios of the Bacillus spores were distinctly, linearly, and predictably related to the carbon and nitrogen isotope ratios of the growth medium.
4. There were distinct and predictable differences in the relationships between hydrogen and oxygen isotope ratios of Bacillus spores that allowed determination of their culture in liquid vessels versus agar.

---

210 Kreuzer-Martin et al., Microbe Forensics, supra note 38.
212 Kreuzer-Martin et al., Microbe Forensics, supra note 38.
213 Id.
plates, and if grown on agar plates, a timeline of spore harvest from agar plates.215

(5) The patterns observed for *Bacillus* spores from several taxa applied to both virulent and non-virulent *Bacillus anthracis* (anthrax) spores.216

Although this specific example shows the relevance and reliability of stable isotope analyses for establishing bacterial relationships, the same approach also can be applied to animal-production systems, such as cattle fed at a feedlot. An investigation may attempt to link the location of that feedlot with the types of grain the animals were fed. The isotope ratios of many biological products (e.g., beef) record useful information related to the region of growth of that product.217 Stable isotope ratio analysis may be useful when illegal import of animal products is alleged or when location and growth type need to be distinguished. In the future, courts are likely to face cases where the authenticity or purported origin of a food product is challenged or the designation of “natural” is challenged.

5. Rule 703—Facts or Date Relied Upon—Otherwise Inadmissible

As noted earlier, Rule 703 lists three bases of facts or data for an expert opinion: those facts or data (1) perceived by the expert at or before the hearing, and (3) of a type reasonably relied upon by experts in the particular field that need not be admissible in evidence. The third basis primarily concerns us here.

The 2000 amendment to Rule 702 “makes clear that the sufficiency of the basis of an expert’s testimony is to be decided under Rule 702.”218 The Advisory Committee explained: “Rule 702 sets forth the overarching requirement of reliability, and an analysis of the sufficiency of the expert’s basis cannot be divorced from the ultimate reliability of the expert’s opinion.”219 Accordingly, whether the expert “reasonably relied” on otherwise inadmissible evidence in “forming opinions or inferences”220 is a “relatively narrow inquiry.”221

Under this inquiry, the stable isotope expert can only rely on such evidence if it is “a type reasonably relied upon by experts in the particular field” of stable

---

218 Fed. R. Evid. 702 advisory committee’s note to 2000 amendment.
219 Id.
220 Fed. R. Evid. 703.
221 Fed. R. Evid. 702 advisory committee’s note to 2000 amendment.
isotope ratio analysis.\textsuperscript{222} As with the other admissibility requirements discussed here, this one is very case-specific. However, because stable isotope analysis is the product of precise laboratory measurement and analysis, reliance on inadmissible hearsay or other inadmissible evidence to form an opinion should be rare.

The use of geospatial mapping of isotope abundance in evaluating laboratory measurements of specific samples may call for Rule 703 analysis if admissibility issues arise regarding this basis of the expert evidence. Whether treated under Rule 702 or 703, the issue comes down to the reliability of this source of data as a basis for expert opinion, and a showing would be needed that stable isotope experts reasonably rely upon this source of facts or data.

6. Rule 706—Court-Appointed Experts

Stable isotope analysis has a long tenure in the laboratory but not in the courtroom. It will be new to most judges, lawyers, and jurors. At least with its initial forays into the courtroom, stable isotope evidence may be a good candidate for a court-appointed expert or technical adviser to assist the judge in making the reliability assessment and the admissibility decision. Technical advisers (or scientific consultants) have a less formal role than court-appointed experts. They provide advice and guidance to the judge and do not typically testify and undergo cross-examination at deposition or trial.\textsuperscript{223} We think that either an appointed expert or technical adviser could serve a judge well in the early days of stable isotope courtroom evidence.

The U.S. judicial system relies largely on parties to call experts who explain, challenge, and defend scientific evidence. But even with the benefit of testimony from party-called experts, judges face challenges in assessing the reliability of complex expert testimony. Justice Breyer has observed that "most judges lack the scientific training that might facilitate the evaluation of scientific claims or the evaluation of expert witnesses who make such claims."\textsuperscript{224} Many state court judges reported in a survey that they lacked adequate training to evaluate all of the expert evidence presented in their courtrooms.\textsuperscript{225} Another survey of judges showed a dearth of training in math and science.\textsuperscript{226}

\textsuperscript{222} FED. R. EVID. 703.


\textsuperscript{224} Stephen Breyer, \textit{Introduction} to FEDERAL JUDICIAL CENTER REFERENCE MANUAL ON SCIENTIFIC EVIDENCE 1, 4 (2d ed., 2000); see also NAS REPORT, \textit{supra} note 7, at 12.


A court-appointed expert or technical adviser could educate and prepare the judge to address expert issues in a particular case, and may help judges exclude or deter unreliable, inaccurate, or biased testimony.\(^{227}\) Just the possibility of using this approach can discourage biased testimony and lead to more accurate results.\(^{228}\)

It is likely for the near term that courts will need to rely on experts who are familiar with the strengths and weaknesses of isotope ratio analytical measurement, the application of stable isotope analyses to a particular discipline (e.g., biochemistry, anthropology, food science, explosives) rather than simply familiarity with the instrumentation, and the constraints on the interpretability of isotope ratio data. Until this type of evidence is well established within the courts, judges should consider relying on court-appointed experts or technical advisers who have attained a PhD degree in one of the sciences, have familiarity with operating a stable isotope ratio laboratory, and have demonstrated experience in the field broadly associated with the evidence being presented to the court.

Judges seeking expert assistance on stable isotope analysis of anthrax should look for isotope scientists who fit the qualifications presented previously for a testifying expert on the application of stable isotope techniques to the anthrax case.

7. Rule 403—Exclusion Based on Prejudice, Confusion, or Duplication

Even if expert evidence meets Rule 702 reliability and all the other foregoing requirements, it can still be excluded under Rule 403 if “its probative value is substantially outweighed by the danger of unfair prejudice, confusion of the issues, or misleading the jury, or by considerations of undue delay, waste of time, or needless presentation of cumulative evidence.”\(^{229}\)

In response to appropriate and relevant questions, testimony interpreting stable isotope ratio evidence can and should be clear, not misleading, and not confusing to the court or jury. Similarly, an expert who recognizes the limits and weaknesses of an application of this technique should not present a danger of unfair prejudice. Critical elements to avoid these problems include calling an expert to testify who has the appropriate qualifications and experience, and having counsel who can elicit responses from the expert that are understandable to the court and jury.

Interpretation of stable isotope data is usually a matter of looking at two data values and asking whether the values are the same or different. An advantage of this approach is that multiple elements or molecules can be analyzed for their stable isotope ratios. Two evidence specimens may have the same carbon isotope ratio value, but for them to be consistent with each other, they should be consistent in stable isotope ratios of all of the elements in each specimen presented as

---

\(^{227}\) See Cooper & Tomlin, supra note 223, at 225–26; Jurs, supra note 147, at 86.


\(^{229}\) FED. R. EVID. 403.
evidence. This basic comparative presentation should not confuse or mislead the fact-finder.

Associated with the stable isotope ratio value of an element is the precision and reliability of that observation. When two or more specimens are being compared, it should be clear these specimens are different if they differ in their values. Again, the training and experience of the expert and the ability to explain this point should avoid Rule 403 problems.

When the expert is asked to provide an interpretation of the evidence, the expert’s comments should address whether the observations are consistent with a stated question or hypothesis. The expert may provide a level of precision to that statement. If such a dialogue is approached when presenting stable isotope ratio data to the court, the interpretation should not induce prejudice or mislead the jury.

It is possible that other evidence in a case has been or can be presented to establish the identification or associative evidence that stable isotope analysis would also establish. For example, stable isotope analysis may be relied upon in an investigation to assist in identifying the geographic origin of an evidence sample. If that information is relevant at trial, and if it can be established more easily by other means (e.g., statements of admission by a party), the stable isotope evidence, just like other information from the investigative phase of a case, may be challenged under Rule 403 as presenting cumulative evidence at trial.

III. FORENSIC IDENTIFICATION, DNA PROFILING, AND STABLE ISOTOPE RATIO ANALYSIS

A. DNA Profiling and Traditional Forensic Techniques

In the post-Daubert era, forensic identification has received both praise and criticism, depending on the identification technique in question. Traditional applications such as analysis of bite marks, hair samples, handwriting, fingerprints, and firearms have long been accepted in criminal proceedings. But, recently called “the problem children of forensic science,” they have been criticized for lacking sufficient empirical research to satisfy rigorous reliability and validity analysis. This criticism has led some to diagnose a crisis in forensic science and to predict

---

231 “These [techniques] have not undergone the type of extensive testing and verification that is the hallmark of science elsewhere.” Donald Kennedy & Richard A. Merrill, Assessing Forensic Science, ISSUES SCI & TECH., Fall 2003, at 33, 34; see also 1 MODERN SCIENTIFIC EVIDENCE, supra note 54, at 89–94; NAS REPORT, supra note 7, at 43–44, 127–82; Paul C. Giannelli, Forensic Science: Under the Microscope, 34 OHIO N.U. L. REV. 313, 353–35 (2008); D. Michael Risinger & Michael J. Saks, A House with No Foundation, ISSUES SCI & TECH., Fall 2003, at 35.
232 See generally JIM FISHER, FORENSICS UNDER FIRE: ARE BAD SCIENCE AND DUELING EXPERTS CORRUPTING CRIMINAL JUSTICE? (2008); KELLY M. PYREK, FORENSIC SCIENCE UNDER SIEGE: THE CHALLENGES OF FORENSIC LABORATORIES AND THE MEDICO-
a paradigm shift based on growing recognition that certain forensic claims of discernable uniqueness for comparison matches may not stand up to proficiency testing or empirical research.233

These issues have been aired prominently with the 2009 release of the much-anticipated NAS Report on forensic science.234 Shortly thereafter, Justice Scalia cited the report in a majority opinion. He wrote that “[s]erious deficiencies have been found in the forensic evidence used in criminal trials,”235 and that “[f]orensic evidence is not uniquely immune from the risk of manipulation.”236 The president of the American Academy of Forensic Sciences has called for validation studies on a variety of forensic applications—e.g., analysis of bite marks, tool marks, handwriting, and latent fingerprints—that courts routinely have admitted for some time.237

Daubert has prompted some trial judges to increase their scrutiny of expert credentials.238 Attorneys have become more aggressive in attacking the admissibility of forensic evidence,239 which has required courts to “confront[] challenges to testimony . . . whose admissibility had long been settled.”240 Most of these challenges have failed to exclude evidence, but some have exposed empirical weaknesses of common forensic techniques.241 Despite the stringent standard of proof in criminal proceedings, various evidence commentators have suggested that trial judges have been more rigorous Daubert gatekeepers in scrutinizing expert evidence in civil cases than in criminal cases.242

Many traditional forensic techniques were developed largely within the setting of crime laboratories’ efforts to aid criminal investigation and prosecution.243 Daubert challenges to these techniques ask judges to reassess not

---

234 See NAS REPORT, supra note 7, at 5–8.
236 Id. at 2536.
238 See GIANNELLI & IMWINKELRIED, supra note 62, at 39.
only the reliability and validity of expert evidence but also the strength of claims that can be made based on application of these methods to particular cases. The risk of overreliance on expert evidence that lacks rigorous scientific grounding can carry over to techniques developed through basic scientific research. To whatever extent the principles and methods underlying expert evidence are based on empirical research and testing, forensic science witnesses should guard against claiming too much. If they attempt to overstate, Daubert expects judges to perform their gatekeeping role.

In contrast to many of the traditional forensic methods, DNA profiling is the product of extensive basic science research. It often is called the “gold standard” of forensic identification and has “revolutionized forensic investigations.” Indeed, the NAS Report states that “no forensic method other than nuclear DNA analysis has been rigorously shown to have the capacity to consistently and with a high degree of certainty support conclusions about ‘individualization’ (more commonly known as ‘matching’ of an unknown item of evidence to a specific known source).” Like DNA, stable isotope analysis was developed in academic science settings and relies on analysis of measurements from high-precision instruments.

Exonerations based on DNA analysis often have corrected wrongful convictions based on forensic science testing errors. The forensic geneticist seeks to identify the source of a biological sample. The scientific community has devoted substantial attention to establishing the accuracy and reliability of DNA profiling for forensic uses. Unlike their approach to certain other methods of

244 See id.

245 See id. at 32–33; Saks & Koehler, supra note 21, at 202 (contending that forensic scientists routinely overstate individualization claims for identification evidence and that individualization is not scientifically valid). But see David H. Kaye, Probability, Individualization and Uniqueness in Forensic Science Evidence: Listening to the Academics, http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1261970 (last visited June 1, 2010) (arguing that individualization claims—that a trace sample came from a person or object—can be scientifically defensible in courtroom explanation of forensic identification when random match probability is negligible).

246 See, e.g., BEECHER-MONAS, supra note 51, at 103; Berger, supra note 19, at 1126; Michael Lynch, God’s Signature: DNA Profiling, The New Gold Standard in Forensic Science, 27 ENDEAVOR 2, 93 (2003); Thornton & Peterson, supra note 20, at 82; see also NAS REPORT, supra note 7, at 40–41.


248 NAS REPORT, supra note 7, at 87.

249 See Saks & Koehler, supra note 233, at 892; NAS REPORT, supra note 7, at 42.

250 Jobling & Gill, supra note 247, at 739.

251 The National Research Council produced two book-length reports for the National Academy of Sciences on DNA forensic identification, including recommendations for strengthening the rigor of laboratory practices and improving courtroom presentation of the evidence. COMMITTEE ON DNA TECHNOLOGY IN FORENSIC SCIENCE, NATIONAL RESEARCH COUNCIL, DNA TECHNOLOGY IN FORENSIC SCIENCE (1992); COMMITTEE ON DNA
forensic identification, courts have insisted on a strong showing of scientific validity for admission of DNA evidence, including written protocols and proficiency testing.\textsuperscript{252} Indeed, DNA has received more extensive judicial scrutiny than any other area of forensic criminal investigation.\textsuperscript{253}

The NAS Report observed: “Among existing forensic methods, only nuclear DNA analysis has been rigorously shown to have the capacity to consistently, and with a high degree of certainty, demonstrate a connection between an evidentiary sample and a specific individual or source.”\textsuperscript{254} The report’s Recommendation 3 calls for research to determine “accuracy, reliability, and validity in the forensic science disciplines.”\textsuperscript{255} The report concluded that many forensic methods lack sufficient empirical research support and called for further research on such familiar techniques as fingerprint examination, handwriting comparison, firearms identification, bite mark identification, and hair analysis.\textsuperscript{256}

\textbf{B. DNA Profiling—Brief Overview}

When DNA profiling evidence is presented in the courtroom, the usual question is whether an evidentiary DNA sample matches a known DNA sample taken from a victim or a suspect. If a match is seen, a random match probability must be calculated to assess the likelihood that the evidentiary sample also matches somebody else in the general population (i.e., whether the DNA profile is unique to the suspect). This number is routinely so low that its reciprocal often exceeds the world’s entire population.\textsuperscript{257}

As noted previously, forensic identification is based on differentiation. Variations in genetic material differentiate individuals from each other. The extreme probabilities found in DNA profiling are based on the extraordinary range of genetic variation. The scientific and legal literature on this subject is

\textbf{FORENSIC SCIENCE: AN UPDATE, NATIONAL RESEARCH COUNCIL, THE EVALUATION OF FORENSIC DNA EVIDENCE (1996).}


\textsuperscript{254} NAS REPORT, supra note 7, at 100.
\textsuperscript{255} Id. at 22–23, 190.
\textsuperscript{256} See id. at 8, 42–44, 136–76.
\textsuperscript{257} Jobling & Gill, supra note 247, at 743; see also People v. Nelson, 185 P.3d 49, 52 (Cal. 2008) (“The prosecution presented evidence that the odds that a random person unrelated to defendant . . . could have fit the profile of some of the crime scene evidence are one in 930 sextillion (93 followed by 22 zeros).”).
extensive. A brief explanation is presented here to allow comparison with stable isotope ratio analysis as a forensic identification tool.

Within each cell (except red blood cells), the nucleus contains a person’s entire genetic code in forty-six chromosomes. Twenty-three chromosomes from the mother combine with twenty-three from the father. About 20,000 to 25,000 genes composed of DNA sequences are located on the chromosomes. DNA’s basic material consists of four nucleotide bases consisting of sugar and phosphate compounds joining in base pairs—adenine with thymine, cyclosine with guanine. Of the approximately three billion base pairs, about three million (0.1 percent) vary (except identical twins), which allows us to distinguish individuals from one another. This differentiation is found mostly in the noncoding DNA. Only about one percent of our DNA encodes to enable transcription to RNA and the production of proteins.

DNA identification analysis focuses on the variation in noncoding DNA. The profiling process uses a primer to find a given location (or locus) of noncoding DNA. People vary in how many times a sequence of base pairs repeats itself at the locus. This variation is called polymorphism. The repeating sequence at the locus is called an allele. The frequency in the population of an allele having a certain number of base pair sequence repeats at a given locus has been determined through compilation of genetic databases. A person’s genotype or genetic profile is based upon which alleles are present at the various chosen loci.

The frequencies of a person’s DNA sequences at up to thirteen locations are multiplied against each other to produce the probability of a random match, an infinitesimally small number. The random match probability is based on the product rule, which yields a probability that a series of independent facts—in this case, the frequencies of alleles at typically thirteen different loci on the DNA strand—will occur. It is derived by multiplying each of the frequencies against each other.

C. DNA Profiling and Stable Isotope Ratio Analysis

Stable isotope analysis, like DNA profiling, originated in mainstream science laboratories primarily associated with university scientists and the geochemistry-petroleum industry. Like other laboratory-based forensic science, such as toxicology and drug analysis, DNA and stable isotope ratio analysis can be

---

258 For further explanation of DNA profiling, see, e.g., LYNN B. JORDE ET AL., MEDICAL GENETICS 29–56 (3d ed. 2003); John M. Butler, Genetics and Genomics of Core Short Tandem Repeat Loci Used in Human Identity Testing, 51 J. FORENSIC SCI. 253 (2006); Peter Gill, DNA as Evidence—The Technology of Identification, 352 NEW ENG. J. MED. 2669 (2005); David H. Kaye & George F. Sensabaugh, Jr., Reference Guide on DNA Evidence, in REFERENCE MANUAL ON SCIENTIFIC EVIDENCE 485 (2d ed. 2000); see generally JOHN M. BUTLER, FORENSIC DNA TYPING—BIOLoGY, TECHNOLOGY, AND GENETICS OF STR MARKERS (2d ed. 2005).

distinguished from forensic disciplines based on expert interpretation and comparison of observed patterns, such as fingerprints, writing samples, bite marks, and tool marks.

Stable isotope ratio analysis cannot claim the same level of extensive basic research, practical applications, federal oversight, private support for applied research, and national quality assurance and control as has been afforded DNA. However, stable isotope analysis should be subjected to the same rigorous courtroom scrutiny as DNA profiling.

DNA evidence has been compared to and contrasted with a variety of other forms of evidence, such as hair analysis, fingerprint identification, and handwriting comparison, arguably obscuring rather than improving understanding of the nature and the reliability of the non-DNA evidence. Analogizing novel scientific forensic evidence to a well-analyzed form of expertise risks a less strenuous analysis of the former rather than rigorous scrutiny of the novel evidence on its own terms, a risk we seek to avoid in the following discussion.

1. Some Similarities and Differences

We note here some similarities and differences between DNA profiling and stable isotope ratio analysis to help develop better understanding of the latter. Both techniques are used forensically to help answer whether an unknown sample and a known sample have a common origin.

DNA profiling is used to determine the origin of biological evidence. Stable isotope analysis can be used for this purpose, but it also can be used to determine the origin of nonbiological evidence and also the geographic origin of evidence. The two methods can complement each other in an investigation and corroborate each other for proof in court. For example, stable isotope evidence can be used to distinguish among different possible regions from which the same DNA evidence might have been found.

In a case involving an unidentified murder victim found during 2005 in Dublin, Ireland, stable isotope analysis of the victim’s tissue was conducted to determine the body’s geographic point of origin and life history. This analysis helped the investigation to identify the victim by serving as the basis to seek a DNA parental cross-matching between the victim and a child. Once the true identity of the victim was established, two murder suspects were quickly identified, followed by arrest and conviction.

Similar to the use of multiple loci on the DNA strand for DNA analysis, multiple chemical elements from a sample can be measured for stable isotope analysis. Just as the probabilities of multiple DNA loci can be multiplied to

\[ \text{Probability of DNA match} = \prod_{i=1}^{n} \text{Probability of locus } i \]

\[ \text{Probability of isotope match} = \prod_{i=1}^{m} \text{Probability of element } i \]

260 See NAS REPORT, supra note 7, at 101.
261 See Roberts, supra note 18, at 248–56.
262 See id. at 269–70.
263 See Wolfram Meier-Augenstein & Isla Fraser, Forensic Isotope Analysis Leads to Identification of a Mutilated Murder Victim, 48 SCI. & JUST. 153 (2008).
determine a random match probability, the frequencies of the stable isotope ratios from multiple elements can be multiplied to determine a random match probability.

Despite the foregoing, DNA profiling produces a much smaller random match probability than stable isotope ratio analysis and other forensic identification techniques. The probability of a random DNA match is so low that DNA profiling is easily the most powerful forensic identification method to produce evidence supporting conviction or exoneration in criminal cases and identifying unknown victims of crimes, accidents, or natural disasters.

The resolving power of stable isotope analysis is rarely as strong as DNA because there is not as much naturally occurring isotopic variation as there is DNA variation, especially when you consider that the number of DNA loci that can be analyzed is much greater than the number of chemical elements in an evidence sample submitted for stable isotope analysis. DNA identification evidence is exceptionally reliable because people other than identical twins have a unique genetic pattern with extraordinary variation.264

The random match probabilities of DNA profiling are the product of well-developed reference population databases that provide the frequency of various alleles occurring at different locations on the DNA strand. The stable isotope databases are based on sample measurements and are not as expansive as DNA databases. In some cases, two samples that are distinguishable through DNA profiling may have the same stable isotope ratios or be sufficiently close in value so that the samples cannot be distinguished with the stable isotope technique.

This distinction in the resolving power of the two methods points to the importance of recognizing what a particular forensic method can accurately tell us.265 Stable isotope analysis can provide forensic evidence that DNA profiling may not or cannot supply. For example, depending on the nature of the sample and the reference database, stable isotope analysis may be able to identify an unknown sample with a precise geographic source. Short of that, it may provide reliable evidence of a geographic area as a sample’s source or exclude a geographic area as a source or origin. In all of these instances, the evidence may not be otherwise obtainable, including through DNA analysis.

Stable isotope analysis can do something that other forensic identification techniques cannot do: relate or distinguish two pieces of evidence that have identical chemical composition, even samples having identical DNA. For example, blood or hair samples from identical twins will produce a DNA match, but if the twins have been located in different geographic regions and have therefore consumed different water and food, the samples will likely exhibit different stable isotope ratios for one or more different elements.

A further advantage of stable isotope analysis is that it can be applied to samples that do not contain DNA. DNA analysis can, of course, be applied to

264 Berger, supra note 19, at 1126–27.
265 See NAS REPORT, supra note 7, at 188.
plants and animals, but stable isotope analysis can be applied to nonbiological samples. Even if samples do contain DNA, and even if DNA analysis can show specific genetic information about the strain of an organism, stable isotope analysis may supplement that information by identifying the range of regions from which that organism—for example, microbiological materials constituting a bioterrorism threat—may have originated.

2. Principles and Methods

DNA profiling has reached a point where its principles and methods and its laboratory techniques are well accepted in the scientific community and the courtroom. The principles and methods of stable isotope ratio analysis are similarly well established in the scientific community but have not received the rigorous judicial scrutiny for courtroom presentation required under Daubert and Rule 702. Peer-reviewed scientific publications and testimony from experts in the field can demonstrate that an application that is new to the legal system is nonetheless well established in its scientific field.

The DNA profiling methodology has taken various forms, and still more will be developed. Each new technique must be validated under the Daubert/702 framework. The PCR STR technique used in most forensic applications today does not vary significantly from case to case. Its principles and methods have met the reliability test in the courtroom and are routinely accepted as reliable to the point where courts take judicial notice of reliability. This does not preclude, of course, the need to ensure that the technique was properly applied in the particular case.

Although stable isotope analysis has a solid scientific foundation and follows a baseline methodology, the precision of the technique could vary depending on the abundances of substances analyzed. For courtroom presentation, this means that the reliability showing for stable isotope methodology would likely need to be tailored to the particular application or “task at hand.”

The principles and methods of both DNA and stable isotope analysis may require a fresh reliability analysis if a new method is used or an unusual application is involved. For example, the inquiry may concern whether there is

---


267 See Kreuzer-Martin et al., Microbe Forensics, supra note 38, at 818–19.

268 See Wilson v. Sirmons, 536 F.3d 1064, 1102 (10th Cir. 2008) (“Numerous federal and state courts as well as scientific investigators have found that PCR DNA analysis is reliable.”); United States v. Boswell, 270 F.3d 1200, 1204–05 (8th Cir. 2001); United States v. Morrow, 374 F. Supp. 2d 51, 61 (D.D.C. 2005).

269 See Kaye & Sensabaugh, supra note 258, at 550.

scientific literature and precedent for the analysis applied to a specimen in a specific case. 271

3. Application of Principles and Methods

The disputed issues in DNA profiling evidence concern specific forensic applications and laboratory techniques—whether the DNA sample was collected, processed, and analyzed properly. The issues include whether evidence samples were mishandled or mislabeled, whether samples that are degraded or contaminated or mixed can still be tested, whether the laboratory followed validated protocols and met quality control and assurance standards, whether the probability of a match between known and unknown DNA samples is valid based on an appropriate reference population and calculation, and whether presentation of the match probability may be unfairly misleading or prejudicial to the fact-finder. 272

For stable isotope ratio analysis, similar issues regarding collection, processing, and analysis of evidence apply, as explained previously in Part II.B.

4. DNA Experts and Isotope Experts

Whether the proposed testifying expert is truly an expert under Rule 702 should be established in every case. DNA profiling testimony may require expertise in several fields—e.g., molecular biology, laboratory procedures, probability and statistics—to establish admissibility or to explain the technique to the jury. 273 Stable isotope ratio analysis also may require expertise in several fields, such as biology, chemistry, and laboratory procedures. In DNA cases where the reliability of principles and methods has been established as a matter of judicial notice, a lab technician may suffice as the expert witness to describe the application and results. 274 Stable isotope cases initially will require expert testimony in all pertinent areas of expertise to establish Daubert reliability. Only when the technique has been accepted in enough cases to set a sound reliability precedent would a lab technician be sufficient.

5. Helping—Not Confusing—the Jury

Even if the proponent of expert testimony satisfies Rule 702 and Daubert, the judge still can exclude the evidence under the Rule 403 balancing test if its probative value is substantially outweighed by its potential to mislead, confuse, or

271 See Kaye & Sensabaugh, supra note 258, at 552.
272 See id. at 489, 503–48; NAS REPORT, supra note 7, at 100; Dan E. Krane et al., Letter to the Editor, Sequential Unmasking: A Means of Minimizing Observer Effects in Forensic DNA Interpretation, 53 J. FORENSIC SCI. 1006 (2008).
273 See Kaye & Sensabaugh, supra note 258, at 489–90.
prejudice the jury. If the judge decides this balance favors allowing admissibility of the evidence, the proponent of the evidence should have ample incentive to present it in a way that is helpful and understandable to the jury.

With DNA evidence, in addition to explaining the principles and methods and their application in the particular case, the expert needs to explain the random match probability. This part of DNA testimony has been challenged as unduly confusing and prejudicial. The expert’s task is to present the random match probability analysis to avoid this risk as much as possible. As we have discussed previously, stable isotope expert testimony faces similar issues.

The NAS Report noted that judicial education programs have been offered for years on DNA identification evidence, but such programs have not focused to any comparable degree on other forensic disciplines. We suspect the same is true for lawyer continuing education. We recommend educational opportunities for judges and lawyers to learn more about forensic science, including stable isotope ratio analysis.

CONCLUSION

Stable isotope analysis has undergone a long gestation in university and other well-respected commercial laboratories. It is based on precise measurements using sophisticated instrumentation. Its reliability and validity for use in many fields and applications are well documented. Like DNA identification analysis, the scientific foundation for stable isotope evidence provides a strong starting point for its use in the courtroom.

Forensic application of stable isotope analysis has been growing in recent years, and its appearance in courtrooms is reasonably certain and just around the corner. As a newcomer to that setting, it should receive judicial scrutiny of its reliability as required under Rule 702 and Daubert. We have written this Article in anticipation of this happening.

We have attempted to explain the science and law that will guide this process. Using the Daubert/702 framework, we have identified the issues that should be addressed. We have stressed that a general showing of reliability of principles and

275 See United States v. Gilliard, 133 F.3d 809, 815–16 (11th Cir. 1998) (upholding district court’s exclusion of polygraph evidence under Rule 403 as potentially confusing); United States v. Pitner, 969 F. Supp. 1246, 1252–53 (W.D. Wash. 1997) (holding that even if polygraph evidence satisfied Rule 702, it “would still be excluded under Fed.R.Evid. 401 and 403 . . . [if] there is a substantial risk that the jurors will substitute the examination results for their own judgment”).

276 One court addressed this issue as follows: “Although I acknowledge that a jury could become confused concerning the meaning and potential significance of a random match probability estimate, I am confident that the risk of confusion is acceptably small if the concept is properly explained.” United States v. Shea, 957 F. Supp. 331, 345 (D.N.H. 1997).

277 See NAS REPORT, supra note 7, at 235.
methods is only the beginning. The expert’s work on the particular case must also satisfy the reliability evaluation.

Because there are so many actual and potential forensic applications of stable isotope analysis, this Article only sets the stage for a judge or lawyer confronted with this evidence. We hope the discussion of the Amerithrax case and the comparisons between stable isotope analysis and DNA identification evidence will be useful. We also hope we have kept faith with the NAS Report by calling for stable isotope evidence to satisfy rigorous reliability scrutiny. Finally, as a scientist and a lawyer, we hope this Article serves the causes of good science and just results.
LONG OVERDUE: UTAH’S INCOMPLETE APPROACH TO EYEWITNESS IDENTIFICATION AND SUGGESTIONS FOR REFORM

Steven J. Joffe

I. INTRODUCTION

On January 17, 1985, Ronald Cotton, a twenty-two-year-old black man, was convicted of rape and sentenced to life in prison.1 That evening, his alleged victim, Jennifer Thompson, “toasted her victory with champagne.”2 Approximately six months earlier, in July of 1984, when Thompson was a twenty-two-year-old college student, a black male had broken into Thompson’s apartment, “held a knife to her throat, and forced her to submit to sexual intercourse.”3 During the course of the horrific attack, Thompson vowed to stay alert so that she could later help police catch her attacker and put him away forever.4 At one point during the rape, she even tricked the man into turning on a light so she could memorize his face and check him for scars, tattoos, and other distinguishing features that could help her identify him.5

Soon after the attack ended, Thompson was admitted to the hospital, where she met with police detectives and created a composite sketch of the face she had seen that night.6 After the sketch was released to the public, tips poured in.7 One of these tips identified Cotton, a convicted felon who worked at a restaurant near Thompson’s home.8 Three days later, police officers asked Thompson to identify her attacker through a photo lineup.9 Although Thompson was unsure at first, after five minutes of staring at the photos, she pointed to a picture of Ronald Cotton (“Cotton”) and identified him as her rapist.10

2 Id.
5 See O’Neill, supra note 1.
6 Id.
7 See Stahl, supra note 4.
8 See id.
9 See id.
10 See id.
Armed with this evidence, prosecutors charged Cotton with rape and brought him to trial. Based largely on Thompson’s eyewitness testimony, the jury found him guilty. After he was sentenced, Cotton was taken to a North Carolina prison, where he was ordered to spend the remainder of his natural life.

Eleven years later, DNA evidence conclusively established that Cotton was not Thompson’s rapist. Instead, the test results conclusively demonstrated that a man named Bobby Poole, a convicted rapist who bore a strong resemblance to Cotton, was the true perpetrator. After eleven devastating years of incarceration, Cotton walked out of prison a free man. Despite her conscious efforts to memorize the face of her attacker and bring him to justice, Thompson had identified the wrong man.

Although the American criminal justice system has long proclaimed that “it is far worse to convict an innocent man than to let a guilty man go free,” Ronald Cotton is not alone. For decades, men and women throughout the United States have been wrongfully convicted of crimes they did not commit, sentenced to lengthy prison terms, and in some cases even put to death. Since 1992, 250 of these wrongfully convicted individuals have been exonerated by DNA evidence; seventeen of them had been sentenced to death. Countless others also have been exonerated through other non-DNA, post-conviction processes. Although the
factors contributing to wrongful convictions are numerous, experts unanimously agree that the leading cause is mistaken eyewitness identification, which has played a role in more than 75 percent of these 250 DNA exonerations.

Although scientific evidence and case examples documenting the inherent unreliability of eyewitness identifications have existed for centuries, state courts and legislatures throughout the country have been slow to respond to the problem. Fortunately, modern scientific studies, combined with public advocacy efforts and common sense, have recently led several states to adopt measures designed to decrease the number of wrongful convictions resulting from mistaken identifications.

The Utah Supreme Court took the state’s first steps toward reform in its 1986 opinion in State v. Long, in which the court mandated that trial courts deliver a cautionary instruction to jurors in cases involving a central issue of eyewitness identification. Despite the court’s good intentions, since Long scientific research consistently has demonstrated that cautionary “instructions do not serve as an


24 See Bethany Shelton, Turning a Blind Eye to Justice: Kansas Courts Must Integrate Scientific Research Regarding Eyewitness Testimony into the Courtroom, 56 U. KAN. L. REV. 949, 953–54 (2008) (“Skepticism of eyewitness testimony accuracy can be traced back thousands of years.”); see also Innocence Project, Eyewitness Misidentification, http://www.innocenceproject.org/understand/EyewitnessMisidentification.php (last visited June 1, 2010) (“As far back as the late 1800s, experts have known that eyewitness identification is all-too-susceptible to error, and that scientific study should guide reforms for identification procedures. In 1907, Hugo Munsterberg published ‘On the Witness Stand,’ in which he questioned the reliability of eyewitness identification. When Yale law professor Edwin Borchard studied 65 wrongful convictions for his pioneering 1932 book, ‘Convicting the Innocent,’ he found that eyewitness misidentification was the leading cause of wrongful convictions.”).

25 See State v. Long, 721 P.2d 483, 491 (Utah 1986) (“[D]espite judicial recognition of the documented unreliability of eyewitness identification, courts have been slow to accord the problem the attention it deserves and to fashion ways of minimizing its potentially unjust effects.”).


27 721 P.2d 483 (Utah 1986).

28 See id. at 492.
effective safeguard against mistaken identifications and convictions . . . .”

Instead, social scientists surmise that reducing mistaken eyewitness identifications involves at least two critical components. First, efforts must begin during an eyewitness’s initial interaction with police because the witness’s memory often becomes contaminated by conscious or unconscious police suggestion. Second, researchers suggest that efforts to curb eyewitness error must continue throughout the trial by permitting expert witnesses to inform jurors of the inherent shortcomings of eyewitness identifications. In the sixteen years following Long, Utah’s courts and legislature have refused to acknowledge this research, instead choosing to rely solely on an outdated and ineffective jury instruction. Recently, however, in State v. Clopten the Utah Supreme Court shifted course and instructed trial courts to admit expert testimony on eyewitness identification in nearly every case in which it is requested. Although the admission of expert testimony fulfills one critical component, Utah has taken no steps to address the impact that police officers’ initial interaction with eyewitnesses can have on mistaken identifications.

This Comment argues that rather than wait another sixteen years to implement further reforms, the Utah legislature should immediately implement legislation requiring Utah’s law enforcement agencies to comply with scientifically proven “best practices” designed to increase eyewitness accuracy and decrease wrongful convictions. Part II of this Comment provides an overview of the leading causes of mistaken eyewitness identifications. Part III explores the past and present steps the Utah Supreme Court has taken to address the eyewitness error problem. Finally, this Comment concludes by discussing the shortcomings of Utah’s current approach and offers a solution to decrease the frequency of eyewitness error and wrongful convictions in the state.

II. THE EYEWITNESS IDENTIFICATION PROBLEM

A. The Memory Formation Process

Contrary to popular belief, the human mind does not operate like a video camera, gathering and recording every detail of an event and accurately retaining it for later playback. Instead, human memories are formed through a highly complex process in which images and details of events are “constantly altered through the integration of new experiences and interpretations.”

---

30 See id. at 823, 842–43.
31 See id. at 842–45.
32 See id. at 823–42.
33 See 2009 UT 84, ¶ 49, 223 P.3d 1103.
34 See Thompson, supra note 22, at 1497.
generally describe this process as including three distinct steps: (1) acquisition; (2) retention; and (3) recall.\textsuperscript{36} During the acquisition stage, an observer witnesses an event “and must decide to which aspects of the visual stimulus he should attend.”\textsuperscript{37} Because such events often involve a vast amount of visual information, during acquisition, the observer must unconsciously determine what information will be stored in their long- or short-term memory systems and what information will be instantly discarded.\textsuperscript{38} Numerous factors such as the observer’s age, race,\textsuperscript{39} maturity, intelligence, personality, and even the environmental conditions present during the event can significantly affect the accuracy of details perceived during acquisition.\textsuperscript{40} These numerous variables, over which law enforcement officials have no control, are generally referred to by experts as estimator variables.\textsuperscript{41}

Next, during the retention stage, the perceived information is encoded and stored in the observer’s memory system for later recall.\textsuperscript{42} Finally, during the recall stage, memories are retrieved from the memory system and restored to the observer’s conscious memory.\textsuperscript{43}

During the retention and recall stages, memories are highly susceptible to manipulation caused by post-event feedback.\textsuperscript{44} For example, the observer may engage in conversations with others, overhear others’ conversations, or read or hear media accounts about the event.\textsuperscript{45} The information and feedback perceived by the observer during these events “can bring about powerful and unexpected changes in the witness’s memory.”\textsuperscript{46}

\section*{B. Police Suggestions & Flawed Procedure}

In addition to estimator variables, researchers also have found that a large amount of post-event feedback occurs during a witness’s initial interactions with police officers following an observed event.\textsuperscript{47} Specifically, because eyewitnesses often feel a strong desire to see that “justice” is done,\textsuperscript{48} they frequently look to the

\begin{itemize}
  \item \textsuperscript{37} See \textit{ELIZABETH LOFTUS, EYEWITNESS TESTIMONY} 21 (1979).
  \item \textsuperscript{38} See Fradella, supra note 36, at 7.
  \item \textsuperscript{39} Researchers suggest that it is more difficult for witnesses to accurately identify individuals of a different race than members of their own race; this scientifically demonstrated phenomenon is referred to as “own-race bias.” See Shelton, supra note 24, at 951–52.
  \item \textsuperscript{40} See Thompson, supra note 22, at 1497.
  \item \textsuperscript{41} See Fradella, supra note 36, at 10.
  \item \textsuperscript{42} See id. at 9–10.
  \item \textsuperscript{43} See id.
  \item \textsuperscript{44} See Shelton, supra note 24, at 972–74.
  \item \textsuperscript{45} See Thompson, supra note 22, at 1498 (quoting \textit{LOFTUS, supra} note 37, at 21–22).
  \item \textsuperscript{46} Id.
  \item \textsuperscript{47} See id. at 1504.
  \item \textsuperscript{48} See \textit{LOFTUS, supra} note 37, at 109 (“[Eyewitnesses] want to see crimes solved and justice done, and this desire may motivate them to volunteer more than is warranted by
officers conducting the investigation, who may unconsciously provide clues to support their own investigative theories and evidence interpretation, for assistance in identifying the “correct” perpetrator.\(^{49}\) Such suggestions can occur in a variety of ways.\(^{50}\) For example, an officer may prepare a live lineup or photo lineup in which the prime suspect stands out from the other individuals,\(^{51}\) or the officer may innocently give subtle suggestions through “innocuous body language” or facial expressions.\(^{52}\) In some cases, officers may even make more direct suggestions by asking a witness to take another look at a specific suspect’s photo or by assuring a witness that he or she has chosen the right person.\(^{53}\) For the reasons discussed above, these subtle suggestions can have a tremendous effect on the accuracy of an eyewitness’s identification.\(^{54}\)

In addition to altering a witness’s memory, or in some cases even creating false memories, police suggestion tends to increase the witnesses’ confidence in the identifications they have made.\(^{55}\) This increased confidence likely accounts for the common phenomenon in which an eyewitness appears unsure of her identification at the time of a lineup, but is then able to take the stand months later at trial and testify with 100 percent certainty that the person she identified is the person she saw commit the crime.\(^{56}\)

Social scientists also have found that the specific structure of the identification techniques used during police interviews may have a significant impact on witnesses’ ability to accurately describe events or identify suspects.\(^{57}\) For example, anyone who has watched a legal drama, such as *Law and Order*, is likely familiar with the “simultaneous” lineup procedure in which an eyewitness is asked to make an identification after viewing six photos or live “suspects”
simultaneously. Despite the traditional acceptance of this process in the law enforcement community, some researchers now argue that this method is flawed and may increase eyewitness error because it “encourages the eyewitness to engage in relative judgment [in which the eyewitness can] compare each lineup individual against the others, to determine which individual most closely resembles the culprit.” The obvious flaw with this approach, however, is that “the person the eyewitness believes most closely resembles the culprit might not be the culprit at all.” In fact, unbeknownst to the witness, “the true culprit might not even be present in the lineup.”

These techniques through which law enforcement officials communicate with eyewitnesses and attempt to retrieve and record information from them are generally referred to by experts as system variables. Unlike estimator variables, “system variables are those that the criminal justice system can and should control” because of the damaging effects they can have on the accuracy of eyewitness identifications.

C. The “Blind” Jury

The problems created through police interactions are further compounded when inaccurate eyewitness testimony is presented in court. In the typical criminal trial, a panel of jurors is asked to evaluate the credibility of evidence—including eyewitness testimony—presented by both the prosecution and defense. Based on these credibility determinations, jurors must ultimately decide whether the evidence presented establishes the defendant’s guilt beyond a reasonable doubt. Although the evaluation of eyewitness testimony by jurors has been a common practice for centuries, today’s research suggests that our system is inherently flawed. The reasons for this conclusion are twofold.

First, according to researchers, jurors tend to have an “‘implicit faith’ in eyewitness identification” and tend to believe such testimony, “despite

---

59 See id. (emphasis omitted); see also Gary L. Wells, Eyewitness Identification Evidence: Science and Reform, CHAMPION, June 2005, at 12, 13–14 (“Witnesses compare one lineup member to the other lineup members, determine who looks most like the perpetrator, and then tend to select that person.”).  
60 Winzeler, supra note 58, at 1599.  
61 Id.  
63 Id. (emphasis added).  
64 See Friedland, supra note 35, at 165–66.  
66 See id. at 165–66.
impeachment, despite aggressive cross-examinations, and despite [even]
cautionary [jury] instructions." In fact, one study suggests that jurors tend to
overestimate “the accuracy of eyewitness identifications by more than 500%.“ In
another study, when only circumstantial evidence was offered at trial, only 18
percent of jurors voted to convict a defendant. However, when the same evidence
plus eyewitness testimony was presented, 72 percent of jurors voted to convict. Perhaps most shocking of all, when jurors were later informed that the eyewitness
was legally blind and was not wearing glasses at the time of the crime, 68 percent
still voted to convict the defendant without any additional corroborating
evidence.

Second, research suggests that jurors are unable to accurately determine the
credibility of eyewitness testimony because they are generally unaware of the
inherent unreliability of eyewitness identifications. Due to this lack of
foundation, jurors tend to make their credibility judgments based on stereotypes
and commonly held misconceptions about human memory. For example, jurors
often focus on “witness confidence, consistency, and memory of specific details.” However, because eyewitnesses generally arrive at trial convinced their
testimonies are true, these factors have been found to have no correlation to their
actual credibility. Similarly, jurors tend to believe that violence in a crime
enhances the clarity with which a memory can be recalled and that the presence of
a weapon during a crime makes a witness’s memory more accurate, when, in fact,
in both cases the opposite is true. Perhaps nothing is more demonstrative of
jurors’ inability to determine eyewitnesses’ credibility than the stories of Ronald
Cotton and the hundreds of other men and women convicted on the basis of such
faulty identifications.

67 Noah Clements, Flipping a Coin: A Solution for the Inherent Unreliability of
Eyewitness Identification Testimony, 40 IND. L. REV. 271, 284 (2007) (quoting ELIZABETH
F. LOFTUS & JAMES M. DOYLE, EYEWITNESS TESTIMONY: CIVIL & CRIMINAL 200 (3d ed.
1997 & Supp. 2004)).
68 Kevin Jon Heller, The Cognitive Psychology of Circumstantial Evidence, 105
MICH. L. REV. 241, 244 (2006).
69 See Robert B. Handberg, Expert Testimony on Eyewitness Identification: A New
70 See id.
71 See id.
72 See Wise et al., supra note 29, at 812 (“Jurors place great faith in eyewitness
testimony because they seemingly believe that perceptual memory is like a videotape that
can be replayed with near perfect fidelity.”).
73 See Clements, supra note 67, at 285 (“The criteria by which jurors judge the
reliability of a witness do not correlate with accuracy.”).
74 See id.
75 See id.
76 See Jules Epstein, Tri-State Vagrancies: The Varying Responses of Delaware, New
Jersey, and Pennsylvania to the Phenomenon of Mistaken Identifications, 12 WIDENER L.
Because wrongful convictions based on mistaken eyewitness identification are created by both the inherent unreliability of human memory and jurors’ inability to accurately assess eyewitness credibility, effective reforms must address both issues. The remainder of this Comment discusses the steps that Utah has already taken to address this problem and suggests further steps to reform the state’s currently incomplete system.

III. UTAH’S APPROACH TO EYEWITNESS TESTIMONY—PAST & PRESENT

A. The “Long” Era: One Step Forward, Two Steps Back

In 1986 the Utah Supreme Court in State v. Long took its first steps toward addressing the inherent unreliability of eyewitness testimony. In Long, the petitioner appealed his convictions of felonious assault and weapons possession arguing his conviction should be reversed based on the trial judge’s refusal to give a cautionary jury instruction regarding the inherent unreliability of eyewitness testimony. Prior to Long, the decision whether to grant such an instruction was “left largely to the discretion of the trial court.” However, after recognizing that “[t]he literature [wa]s replete with empirical studies documenting the unreliability of eyewitness identification” and that “human memory is both limited and fallible,” the Utah Supreme Court abandoned this discretionary approach and required that a cautionary instruction be read to jurors in all future cases involving a central issue of eyewitness identification. After announcing this new rule, the court concluded its opinion by noting that its chosen approach would offer “some protection from false conviction, while [also] ensuring the efficacy of the jury system by providing jurors with the knowledge necessary for sound decision making.”

During the next sixteen years, in a series of cases the court upheld its opinion in Long while simultaneously refusing to take further steps to safeguard defendants against the consequences of eyewitness error. For example, in State v. Butterfield, the court rejected a defendant’s argument that the trial court erred in excluding expert testimony concerning the “limitations inherent in eyewitness identification.” In reaching this decision, the court noted that although it required trial courts to give cautionary instructions, it did not require the instruction “

---

77 See, e.g., Wise et al., supra note 29, at 833.
78 721 P.2d 483 (Utah 1986).
79 See id. at 488.
80 See id. at 484, 487–88.
81 Id. at 487 (citing State v. Tucker, 709 P.2d 313, 316 (Utah 1985)).
82 Id. at 488.
83 Id. at 492.
84 Id.
86 Id. ¶¶ 41, 43.
include additional expert testimony concerning eyewitness identification."\(^{87}\)

Instead, the court explained that decisions whether to admit such testimony would be left to the "sound discretion of the trial court" and would be reviewed only for abuse of discretion.\(^{88}\) Additionally, the justices concluded that the trial court had not abused its discretion because the expert testimony sought to be admitted "did not deal with the specific facts from [the] case [and would therefore] constitute a lecture to the jury about how it should judge the evidence."\(^{89}\) This function, the court reasoned, was fulfilled by a Long instruction which "adequately and thoroughly explain[ed] how to evaluate eyewitness identifications presented at trial."\(^{90}\)

Less than a year later, in State v. Hubbard,\(^{91}\) the court took the reasoning of Butterfield one step further. Specifically, after reviewing a trial court’s cautionary jury instruction and finding it to be insufficient, the court refused to overturn the exclusion of expert testimony holding that the flawed instruction had still satisfied the Supreme Court’s “expressed concerns about the need for cautionary [jury] instructions.”\(^{92}\) Thus, the court concluded that even if a trial court’s decision to not admit expert testimony is erroneous, such error would be considered harmless as long as a Long instruction was given to the jury.\(^{93}\) Finally, in State v. Maestas,\(^{94}\) over the dissenting opinion of Chief Justice Durham, the court affirmed its position that a Long instruction adequately informed the jury of the inherent unreliability of eyewitness identifications and that expert testimony was therefore unnecessary.\(^{95}\)

Although the court in Hubbard claimed that it was not adopting a per se rule concerning the admissibility of expert testimony on eyewitness identifications,\(^{96}\) in practice, this is precisely what occurred. In every case following Long, Utah appellate courts repeatedly upheld trial court decisions excluding expert testimony on the grounds that the Long instruction was an adequate substitute.

**B. Clopten’s Solution: One Half of the Battle**

Fortunately, in December of 2009, at the encouragement of the Utah Court of Appeals to “revisit the boundaries of trial court discretion in excluding expert testimony on the subject,”\(^{97}\) the Utah Supreme Court granted certiorari in State v. Clopten\(^ {98}\) to review the question of “whether expert testimony regarding the

---

\(^{87}\) Id. ¶ 42 (quoting State v. Kinsey, 797 P.2d 424, 427 (Utah Ct. App. 1990)).

\(^{88}\) Id. ¶ 43.

\(^{89}\) Id. ¶ 44.

\(^{90}\) Id.

\(^{91}\) 2002 UT 45, 48 P.3d 953.

\(^{92}\) Id. ¶ 20 (quoting State v. Long, 721 P.2d 483, 495 (Utah 1986)).

\(^{93}\) See id.

\(^{94}\) 2002 UT 123, 63 P.3d 621.

\(^{95}\) See id. ¶¶ 65, 69.

\(^{96}\) See 2002 UT 45, ¶ 14.


\(^{98}\) 2009 UT 84, 223 P.3d 1103.
reliability of eyewitness identification should be presumed admissible." In addressing the issue, the court first noted that “it was never [its] intent . . . to establish cautionary instructions as the sole means for educating juries about eyewitness fallibility.” But it had become “clear that Long [had] actually [created] a de facto presumption against eyewitness expert testimony in Utah’s trial courts.” In recognition of the “[d]ecades of stud[ies] . . . establish[ing] that eyewitnesses are prone to identifying the wrong person as the perpetrator of a crime,” the court concluded that such a presumption against expert testimony was inappropriate. Instead, the court held that the admissibility of eyewitness expert testimony should be determined under Utah Rule of Evidence 702, which, like its federal counterpart, requires that expert testimony “assist the trier of fact.” Although the court stated that it was not adopting an outright presumption in favor of expert testimony, it went on to specifically address each prong of the 702 analysis and ultimately concluded that expert testimony should be routinely admitted “in cases where eyewitnesses are identifying a stranger and where one or more established factors affecting accuracy are present.” Finally, the court modified its approach concerning cautionary instructions during “the Long era” by specifying that a Long instruction would no longer be required when expert testimony concerning eyewitness identification is admitted.

Although the court carefully cautioned that it was “not mandating the admission of eyewitness expert testimony in every case,” this appears to be precisely what it has done. As the dissenting opinion in Clopton recognized, the majority’s conclusion that Rule 702 is satisfied “in every case involving an eyewitness identification of a stranger,” leaves district court judges “no discretion” to exclude eyewitness expert testimony except in extremely rare cases. Thus, in nearly every criminal trial in which eyewitness testimony is presented, Utah defendants will now almost certainly be entitled to an eyewitness expert.

99 Id. ¶ 6.
100 Id. ¶ 12.
101 Id. ¶ 12–13.
102 Id. ¶ 15.
103 See id. ¶ 30.
104 See Fed. R. Evid. 702(a). Utah Rule of Evidence 702 in relevant part specifies that “if scientific, technical, or other specialized knowledge will assist the trier of fact to understand the evidence or to determine a fact in issue, a witness qualified as an expert by knowledge, skill, experience, training, or education, may testify thereto in the form of an opinion or otherwise.” Utah R. Evid. 702(a).
105 See Clopton, 2009 UT 84, ¶ 30.
106 Id. ¶ 31–32.
107 See id. ¶ 34. The Court also noted that “[i]n cases where the defense does not call an eyewitness expert, the holding in Long still applies.” Id. In other cases, where eyewitness testimony is heard, admissibility of such expert testimony “is a matter for the trial judge’s discretion.” Id.
108 Id. ¶ 33.
109 See id. ¶ 53 (Durrant, J., dissenting) (emphasis added).
IV. ISSUES WITH UTAH’S CURRENT APPROACH AND SUGGESTIONS FOR REFORM

A. The Current Problem

The rule adopted in Clopiten is clearly an important victory for criminal defendants, defense attorneys, and wrongful conviction advocates in both Utah and the nation at large. Several questions remain, however, concerning the Clopiten rule’s practical implementation and effectiveness in deterring wrongful convictions. For example, it is currently unclear who will pay for experts called to testify on behalf of defendants utilizing public defenders. Similarly, critics question whether the rule will actually fulfill its purpose or instead merely create a “battle of the experts.”

The answers to these concerns are still somewhat unknown. However, one thing is strikingly clear—the Clopiten rule is an incomplete solution. Quite simply, the admission of expert testimony is not a complete solution. To the contrary, it constitutes only a small Band-Aid on the gaping wound created by eyewitness error in our criminal justice system. To be sure, research has found that “the live testimony of an expert can help dilute a jury’s excessive reliance on eyewitness identifications” by informing jurors of the issues associated with eyewitness error. However, even with the assistance of expert testimony, the ultimate decision concerning an eyewitness’s credibility is still placed in the hands of jurors who, in many cases, are simply incapable of accurately making this determination. Take for example the research study discussed above involving the legally blind witness. On first glance, the study serves as an alarming example of the inability of jurors to determine witness credibility. Perhaps more alarming, however, is the recognition that even after the jurors were informed through “expert testimony” that the eyewitness was legally blind, an astounding 68 percent—only 4 percent less than the 72 percent who originally found the defendant guilty—still voted to convict the defendant. Furthermore, while expert testimony can assist jurors in understanding estimator and system variables, it does nothing to prevent system variables, which the system can and should control, from contaminating a witness’s memory in the first place—long before an innocent person is ever hauled into court.

111 See, e.g., Ric Simmons, Conquering the Province of the Jury: Expert Testimony and the Professionalization of Fact-Finding, 74 U. Cin. L. Rev. 1013, 1060 (2006) (discussing the concern that the admission of expert testimony can lead to a “battle of the experts”).
113 See Handberg, supra note 69, at 1043.
114 Id.
B. The Solution

Given the demonstrated inability of jurors to accurately access the accuracy and credibility of eyewitness identifications—with or without expert testimony—effective reforms to decrease eyewitness error must begin long before an innocent person is ever brought to trial. In recognition of this need for earlier intervention, researchers and agencies throughout the country, including the U.S. Department of Justice, the American Bar Association, and the National Innocence Project, have conducted extensive research into the investigative techniques utilized by law enforcement agencies and the influence these techniques can have on eyewitness identifications. Fortunately, this research has helped experts develop reformed identification procedures—known as “best practices”—designed to reduce the impact of police suggestion on eyewitnesses. These reformed procedures provide specific guidelines for the presentation, composition, and procedure of lineups, show-ups, and other investigative techniques. Specifically, the best practices provide six key principles designed to ensure that an eyewitness’s identification is the “result of her memory of the crime and not the manner in which the identification procedure was conducted.” First, whenever a potential eyewitness is asked to make an identification, through either a lineup or show-up, the eyewitness must be told that “the perpetrator may not be in the lineup and that the investigation will continue regardless of whether an identification is made.” Second, all lineups must be conducted using a “double-blind” technique in which photos or suspects are presented by an “administrator who does not know who the suspect is.” This “double-blind” technique “prevents well-intentioned officials from giving inadvertent clues to the witness as to which person in the lineup is the

---

115 See, e.g., Wise et al., supra note 29, at 833.
118 See Innocence Project, supra note 49, at 11.
120 See Turtle et al., supra note 119, at 12–13.
121 See id. at 12–32.
122 Innocence Project, supra note 49, at 5, 19; see Turtle et al., supra note 119, at 18.
123 Innocence Project, supra note 49, at 4, 20; see Turtle et al., supra note 119, at 20–29.
police suspect.” Third, lineups must include only one suspect, who resembles the witness’s description of the perpetrator, as well as four “fillers” who also resemble the suspect’s description to ensure that a suspect “is not identified simply because of the composition of the lineup.” Fourth, all lineups should be presented in a sequential format in which suspects and fillers “are presented one-by-one . . . instead of side by side,” to reduce the issues created by the simultaneous lineup process previously discussed. Fifth, immediately after an identification a witness must provide a statement in his or her own words indicating their level of confidence. These confidence statements help jurors “to assess the eyewitness evidence appropriately” and help minimize “the effects of reinforcing feedback that can distort the confidence level of an eyewitness between the time of the identification and . . . trial.” Finally, all identification procedures must be videotaped to ensure an accurate record of the identification and investigatory techniques.

Research and real-world application has consistently demonstrated that implementation of these “best practices” not only helps to “maximize the quality, accuracy, and value of eyewitness evidence,” but also minimizes “errors that can potentially waste time in an investigation, cause unnecessary hardship for [the wrongfully accused],” and most importantly, “leave the actual offender at large” in our communities. For example, one study found that solely requiring police officers to give an eyewitness a cautionary instruction that the “perpetrator may not be in the lineup reduced the rate of erroneous identifications by 42%” while only reducing “the rate of accurate identifications . . . by 2%.” Similarly, by eliminating cases based on faulty identifications long before litigation ever commences, implementation of these principles can help increase both the efficiency and resource allocation of our criminal justice system.

Despite the overwhelming research in support of these “best practices” and their implementation in states throughout the nation, including New Jersey,"
Wisconsin,\textsuperscript{134} and North Carolina,\textsuperscript{135} the Utah legislature has taken no steps to ensure compliance with these techniques by state law enforcement officials. Instead, the Utah Code solely specifies that “peace officers conducting a lineup shall not attempt to influence the identification of any particular suspect.”\textsuperscript{136} The code neither suggests nor requires any specific techniques designed to reduce eyewitness errors.\textsuperscript{137} Furthermore, while some state police agencies may have adopted departmental best practices, there currently is no state oversight assuring that these procedures are followed and no consequences imposed when they are not.

To combat the problem of eyewitness error and its devastating effect on the lives of wrongfully convicted individuals, the Utah legislature should amend Title 77 (Criminal Procedure) of the Utah Code to require that all state and county law enforcement agencies comply with the best practices discussed above. A model statute outlining these requirements is included in the final pages of this Comment.

Additionally, to ensure compliance with these best practices, the legislature should provide strict remedies for any violations. Depending on the blatancy of a violation, these remedies could range from complete exclusion of the eyewitness’s testimony, and all evidence discovered through that testimony, to less drastic measures like requiring that the trial court simply consider non-compliance in “adjudicating motions to suppress the eyewitness’s identification.”\textsuperscript{138}

Finally, to ensure proper implementation of these procedures, the Utah Department of Public Safety should create training programs designed to educate state law enforcement officials on the “methods, technical aspects and scientific findings regarding the basis of eyewitness identification practices and procedures” outlined above.\textsuperscript{139} Adoption of these principles will not only allow law enforcement officials to conduct their investigations more efficiently, but will also allow Utah’s prosecutors, citizens, and judges to have increased confidence in the police’s ability to bring the right defendant to trial.

V. CONCLUSION

For centuries, experts and social scientists have warned of the inherent unreliability of eyewitness testimony. The failure of courts and legislatures to address this problem and develop adequate safeguards against it has resulted in the wrongful convictions of hundreds, perhaps thousands, of innocent men and women. Although Utah’s decision to provide jurors with expert testimony

\begin{footnotesize}
\begin{itemize}
\item \textsuperscript{135} N.C. Gen. Stat. § 15A-284.52 (2009).
\item \textsuperscript{136} Utah Code Ann. § 77-8-3 (2009).
\item \textsuperscript{137} See id.
\item \textsuperscript{139} Innocence Project, supra note 49, at 42.
\end{itemize}
\end{footnotesize}
concerning the credibility of eyewitness identification is a strong step in the right
direction, it resolves only a small piece of the eyewitness identification problem
and cannot alone be considered an adequate solution. Instead, given the current
scientific research and the state’s desire to ensure justice and fairness in criminal
proceedings, the Utah legislature should pass legislation mandating that Utah’s law
enforcement agencies use scientifically based “best practices” during all
investigative procedures involving eyewitnesses. Rather than wait another sixteen
years to implement these reforms, as the state did during the Long era, these
reforms should be implemented immediately. Combining these scientifically
proven best practices with the newly granted access to eyewitness expert testimony
not only will increase the ability of Utah’s law enforcement officials and courts to
ensure that justice is done, but also will help prevent eyewitness errors long before
an innocent person is ever hauled into court. Only through such a complete
approach to eyewitness identifications can the state ensure that the Ronald Cottons
of this world no longer sit in prison while the guilty remain free.

VI. MODEL UTAH LEGISLATION

Title 77-8-3—Eyewitness Identification Reform Act

§ A. Eyewitness Identification Procedures
(1) All lineups shall be conducted by an independent administrator or
by an alternative method including automated computer programs
or other methods approved by the Utah Department of Public
Safety.
(2) Individuals or photos shall be presented to witnesses sequentially,
with each individual or photo presented to the witness separately, in
a previously determined order, and removed after it is viewed before
the next individual or photo is presented.
(3) Before a lineup, the eyewitness shall be instructed that:
   a. The perpetrator might or might not be presented in the lineup;
   b. The lineup administrator does not know the suspect’s identity;
   c. The eyewitness should not feel compelled to make an
      identification;
   d. It is as important to exclude innocent persons as it is to identify
      the perpetrator; and
   e. The investigation will continue whether or not an identification
      is made.
(4) Before the lineup is conducted, the eyewitness shall acknowledge
the receipt of the instructions in writing. If the eyewitness refuses to
sign, the lineup administrator shall note the refusal and shall sign
such notation.

140 Adapted from North Carolina’s Eyewitness Identification Reform Act, N.C. GEN.
(5) In a photo lineup, the photograph of the suspect shall be contemporary and, to the extent practicable, shall resemble the suspect’s appearance at the time of the alleged offense.

(6) All lineups shall be composed so that fillers generally resemble the eyewitness’s description of the perpetrator, while ensuring that the suspect does not unduly stand out from the fillers.

(7) If there are multiple eyewitnesses, the suspects shall be placed in a different position in the lineup or photo array for each eyewitness.

(8) In a lineup, no writings or information concerning any previous arrest, indictment, or conviction of the suspect shall be visible or made known to the eyewitness.

(9) In a live lineup, any identifying actions, such as speech, gestures, or other movements, shall be performed by all lineup participants.

(10) In a live lineup, all lineup participants must be out of view of the eyewitness prior to the lineup.

(11) Only one suspect shall be included in a lineup.

(12) Nothing shall be said to the eyewitness regarding the suspect’s position in the lineup or regarding anything that might influence the eyewitness’s identification.

(13) If there is an identification, the lineup administrator shall seek and document a clear statement from the eyewitness, at the time of the identification and in the eyewitness’s own words, as to the eyewitness’s confidence level that the person identified in a given lineup is the perpetrator. The lineup administrator shall separate all witnesses in order to discourage witnesses from conferring with one another before or during the procedure. Each witness shall be given instructions regarding the identification procedures without other witnesses present.

(14) If the eyewitness identifies a person as the perpetrator, the eyewitness shall not be provided any information concerning the person before the lineup administrator obtains the eyewitness’s confidence statement about the selection. There shall not be anyone present during the live lineup or photographic identification procedures who knows the suspect’s identity, except the eyewitness and counsel as required by law.

(15) Unless it is not practicable a video record of live identification procedures shall be made. If a video record is not practical, the reasons shall be documented, and an audio record shall be made. If neither a video nor audio recording is practical, the reasons shall be documented, and the lineup administrator shall make a written record of the lineup.

(16) Whether through video, audio, or in writing, the record shall include all of the following information:

a. All identification and non-identification results obtained during the identification procedure, signed by the eyewitness,
including the eyewitness’s confidence statement. If the eyewitness refuses to sign, the lineup administrator shall note the refusal of the eyewitness to sign the results and shall also sign the notation.

b. The names of all persons present at the lineup.
c. The date, time, and location of the lineup.
d. The words used by the eyewitness in any identification, including words that describe the eyewitness’s certainty of identification.
e. Whether it was a photo lineup or live lineup and how many photos or individuals were presented in the lineup.
f. The sources of all photographs or persons used.
g. In a photo lineup, the photographs themselves.
h. In a live lineup, a photo or other visual recording of the lineup that includes all persons who participated in the lineup.