VIRGINIA LAW REVIEW

VOLUME 95 MARCH 2009 NUMBER 1

ARTICLES

INVALID FORENSIC SCIENCE TESTIMONY AND WRONGFUL CONVICTIONS

Brandon L. Garrett[*] *and Peter J. Neufeld*[**]

THIS is the first study to explore the forensic science testimony by prosecution experts in the trials of innocent persons, all convicted of serious crimes, who were later exonerated by post-conviction DNA testing. Trial transcripts were sought for all 156 exonerees identified as having trial testimony by forensic analysts, of which 137 were located and reviewed. These trials most commonly included testimony concerning serological analysis and microscopic hair comparison, but some in-

[*] Associate Professor, University of Virginia School of Law.

[**] Co-Founder and Co-Director, The Innocence Project.

For their invaluable comments, we thank Kerry Abrams, Edward Blake, John Butler, Paul Chevigny, Simon Cole, Madeline deLone, Jeff Fagan, Stephen Fienberg, Samuel Gross, Eric Lander, David Kaye, Richard Lewontin, Greg Mitchell, John Monahan, Erin Murphy, Sinead O'Doherty, George Rutherglen, Stephen Schulhofer, William Thompson, Larry Walker, and participants at the Harvard Criminal Justice Roundtable, a UVA School of Law summer workshop, and the Ninth Annual Buck Colbert Franklin Memorial Civil Rights Lecture at the University of Tulsa College of Law. The authors thank the participants at the Fourth Meeting of the National Academy of Science, Committee on Identifying the Needs of the Forensic Sciences Community for their useful comments, and we thank the Committee for inviting our participation at that meeting. Preliminary study data were presented as a report to the Committee on February 1, 2008. We give special thanks to all of the attorneys who graciously volunteered their time to help locate and copy exonerees' criminal trial transcripts, and particularly Winston & Strawn, LLP for its work in locating and scanning files from DNA exonerees' cases. The fine research assistants whose contributions made this study possible are: Josephine Liu, T.J. Parnham, Kerry Shapleigh, and Jason Shyung. For excellent reference library assistance, we thank Jon Ashley, Michelle Morris and Kent Olson. Finally, we thank the Olin Program at University of Virginia Law School for its research support.

cluded bite mark, shoe print, soil, fiber, and fingerprint comparisons, and several included DNA testing. This study found that in the bulk of these trials of innocent defendants—82 cases or 60%—forensic analysts called by the prosecution provided invalid testimony at trial—that is, testimony with conclusions misstating empirical data or wholly unsupported by empirical data. This was not the testimony of a mere handful of analysts: this set of trials included invalid testimony by 72 forensic analysts called by the prosecution and employed by 52 laboratories, practices, or hospitals from 25 states. Unfortunately, the adversarial process largely failed to police this invalid testimony. Defense counsel rarely cross-examined analysts concerning invalid testimony and rarely obtained experts of their own. In the few cases in which invalid forensic science was challenged, judges seldom provided relief. This evidence supports efforts to create scientific oversight mechanisms for reviewing forensic testimony and to develop clear scientific standards for written reports and testimony. The scientific community can through an official government entity promulgate standards to ensure the valid presentation of forensic science in criminal cases and thus the integrity and fairness of the criminal process.

INTRODUCTION

Over the past two decades, DNA testing technology has both enhanced and eroded the status of forensic science in criminal cases. Traditional forensic disciplines were unable to identify a perpetrator with any great discrimination. For example, conventional serology analysis of blood group substances was widely used in sexual assault cases through the 1980s. The underlying method was sound and frequencies of the A, B, and O blood types were derived from well-established and scientifically valid databases. While serology could exclude or place an individual within a per-

centage of the population with a given blood type, it could not distinguish particular individuals with any greater specificity.

Forensic science had advanced dramatically by 1989, when Gary Dotson became the first innocent person in the United States exonerated by post-conviction DNA testing. A jury convicted Dotson in 1979 of rape, and he was sentenced to 25–50 years in prison.[1] In 1988, DNA testing was conducted after the Governor of Illinois had denied Dotson a pardon, despite the victim's recantation in which she stated that she had fabricated her accusation to conceal consensual intercourse with her boyfriend.[2] Edward Blake, who pioneered the forensic application of the polymerase chain reaction ("PCR") technology, conducted the testing. He found that the DNA results excluded Dotson as the source for the male genetic profile, but that the victim's boyfriend was included.[3] Based on those findings, Dotson's conviction was vacated.[4]

Blake also found that the State's forensic analyst's testimony at Dotson's trial was misleading. The analyst had testified that both Dotson and the semen donor possessed the B blood type, a type shared by only eleven percent of Caucasians. The problem was not with the methods used in the laboratory but with the testimony in the courtroom. While on the witness stand, the analyst did not tell the jury that the victim was also Type B and that her fluids were mixed in the sample. The Type B substances observed in the sample could have come entirely from the victim. Her genetic markers could have overwhelmed, or "masked," those from the semen; as Blake put it, "no genetic information was obtained about the semen donor."[5] Thus, based on the testing methods available at the time, any male could have been the donor. It was misleading to suggest to the jury that a subset (11%) of the population including

---

[1] Dotson's accuser later wrote a confessional book explaining her false testimony titled *Forgive Me*. Cathleen C. Webb & Marie Chapian, Forgive Me (1985).

[2] See Edward Connors, Thomas Lundregan, Neal Miller & Tom McEwan, Nat'l Inst. of Justice, Convicted by Juries, Exonerated by Science: Case Studies in the Use of DNA Evidence to Establish Innocence After Trial 51–52 (1996), available at http://www.dna.gov/case_studies/convicted_exonerated/dotson [hereinafter Convicted by Juries, Exonerated by Science].

[3] Affidavit of Edward T. Blake at 23, D. Crim., People of the State of Illinois v. Gary E. Dotson, No. P.C. 4333 (July 29, 1985).

[4] Convicted by Juries, Exonerated by Science, supra note 2.

[5] Affidavit of Edward T. Blake, State v. Gary E. Dotson, supra note 3, at 13.

Dotson could have been the donor and imply that 89% of the population was excluded.[6]

Thus, scientific advances led to Dotson's exoneration, but invalid forensic science testimony had also supported his conviction. Two hundred thirty-two innocent persons have now been exonerated by post-conviction DNA testing.[7] Several of those exonerations, like Dotson's, have triggered scrutiny of the use of forensic science.[8] Scandals involving faulty work at some of our nation's preeminent crime laboratories, including several arising from exoneration cases, have led to investigations, audits, and efforts to provide independent oversight.[9] At the same time, scientists, legislators, and lawyers have raised questions concerning the validity and reliability of certain forensic science techniques. The American Bar Association issued a set of reform principles,[10] and courts increasingly scrutinize forensic evidence in criminal cases.[11] Such efforts, unlike this study, chiefly focus on either the reliability of forensic science techniques or whether the underlying methodology is sound.

---

[6] Id. Blake, who generously offered comments on this paper, conducted post-conviction DNA analysis in several other exonerees' cases. Blake published extensively on conventional serology prior to his groundbreaking DNA work. See Curriculum Vitae, Edward T. Blake, http://www.fsalab.com/etb_cv.htm#publications.

[7] See The Innocence Project Home Page, http://www.innocenceproject.org (providing count of U.S. post-conviction DNA exonerations; the number as of January 29, 2009, is 232). The first author of this Article conducted a study of how the first 200 DNA exonerees fared during appeals and post-conviction proceedings. See Brandon L. Garrett, Judging Innocence, 108 Colum. L. Rev. 55, 58–59 (2008).

[8] See Robert Bazell, DNA Acquittals Shaking Up Forensic Science, NBC News, Feb. 12, 2008, http://www.msnbc.msn.com/id/23113417; Jane Campbell Moriarty, "Misconvictions," Science, and the Ministers of Justice, 86 Neb. L. Rev. 1 (2007).

[9] See Margaret A. Berger, The Impact of DNA Exonerations on the Criminal Justice System, 34 J.L. Med. & Ethics 320, 321–22 (2006) (noting investigations of "mistakes due to the incompetence or fraud of particular analysts," some of which "have gone on for years"); Paul C. Giannelli, Wrongful Convictions and Forensic Science: The Need to Regulate Crime Labs, 86 N.C. L. Rev. 163 (2007) (describing news reports of scandals and oversight efforts); Maurice Possley et al., Scandal Touches Even Elite Labs: Flawed Work, Resistance to Scrutiny Seen Across U.S., Chi. Trib., Oct. 21, 2004, § 1, at 1.

[10] Andrew E. Taslitz, Convicting the Guilty, Acquitting the Innocent: The ABA Takes a Stand, 19 Crim. Just. 18–19 (2005).

[11] See, e.g., House v. Bell, 386 F.3d 668, 708 (6th Cir. 2004) (Merritt, J., dissenting) ("High on the list of the causes for mistakes are the kinds of errors we see in this case: the misinterpretation or abuse of scientific evidence . . . ."); United States v. Bentham, 414 F. Supp. 2d 472, 473 (S.D.N.Y. 2006) ("False positives—that is, inaccurate incriminating test results—are endemic to much of what passes for 'forensic science.'").

Meanwhile, Congress tasked the National Academy of Sciences ("NAS") with examining ways to improve the quality of forensic sciences.[12] The Committee's landmark report emphasized that a wide range of forensic disciplines lack validity, where "[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 individual or source."[13] The NAS report noted that "[n]ew doubts about the accuracy of some forensic science practices have intensified with the growing numbers of exonerations resulting from DNA analysis (and the concomitant realization that guilty parties sometimes walk free)."[14] The report recommended wholesale reforms to improve not just the reliability and accuracy of forensic science, but also its presentation, including the creation of an independent federal agency—a "National Institute of Forensic Science"—to establish and enforce the use of "standard terminology" for report writing and testimony.[15] Those latter recommendations are important—and the trials of the exonerated show why.

This study raises a neglected question: even assuming that a particular forensic technique has been validated and deemed reliable for casework, how do we ensure that the data will be interpreted, reported, and testified to within appropriate scientific parameters? Traditionally, there has been almost no oversight of what scientists *say* in the courtroom once the court deems the method used valid and reliable. To look at the problem of forensic science testimony in the courtroom, this Article will examine for the first time a set of criminal trial transcripts in the cases of DNA exonerees. The study asks whether forensic science testimony in exonerees' trials com-

---

[12] See Comm. on Identifying the Needs of the Forensic Science Cmty., Nat'l Research Council of the Nat'l Acads., Strengthening Forensic Science in the United States: A Path Forward (2009), available at http://www.nap.edu/catalog.php?record_id=12589 [hereinafter Strengthening Forensic Science]; see also Comm. on Sci., Tech., and Law, The Nat'l Acads., Identifying the Needs of the Forensic Sciences Community, available at http://www8.nationalacademies.org/cp/projectview.aspx?key=48741. That Committee asked the authors to examine the incidence of invalid forensic science testimony in trials of persons exonerated by post-conviction DNA testing.

[13] Strengthening Forensic Science, supra note 12, at S-5.

[14] Id. at 1-2.

[15] Id. at S-14–S-19, 6-3–6-5, 7-19.

ported with valid scientific principles.[16] Throughout this Article, "invalid" testimony denotes a conclusion not supported by empirical data.[17] This study does not examine reliability—that is, whether a forensic methodology produces consistent results.[18] Nor does it examine whether in a particular case, an examiner made a mistake or engaged in misconduct in the laboratory. Instead, this study ex-

[16] The use of forensic science in criminal trials more generally remains little studied. Simon Cole, Where the Rubber Meets the Road: Thinking About Expert Evidence as Expert Testimony, 52 Vill. L. Rev. 803, 819–24 (2007) (describing the many challenges of conducting such research, particularly where "American trial practice" is "incompletely and sporadically reported"). Several reports note particular examples of forensic science that supported wrongful convictions. The first such study was *Convicted by Juries, Exonerated by Science*, supra note 2, issued by the National Institute for Justice ("NIJ") in 1996, which examined 28 DNA exonerations. Id. at xii. That study did not address whether in the 28 cases forensic science was presented in a valid manner. The second author co-authored a book that described several cases involving use of "junk science," but not how many of the first 86 DNA exonerations involved invalid forensic science. See Barry Scheck, Peter Neufeld & Jim Dwyer, Actual Innocence 158–71 (2000). Others discuss examples of invalid use of forensic science. See Samuel R. Gross et al., Exonerations in the United States: 1989 Through 2003, 95 J. Crim. L. & Criminology 523, 543 (2005) (stating that in 24 exoneration cases, including non-DNA cases, a forensic expert committed perjury). The first author identified cases among the first 200 exonerated by post-conviction DNA testing in which forensic science supported the conviction and then studied appellate and post-conviction proceedings challenging that evidence. See Garrett, supra note 7, §§ II.A–II.B.

[17] Daubert v. Merrill Dow Pharmaceutical, 509 U.S. 579, 590 n.9 (1993) ("[S]cientists typically distinguish between 'validity' (does the principle support what it purports to show?) and 'reliability' (does application of the principle produce consistent results?)."). This study does not examine the reliability or validity of methods used, but rather the validity of conclusions reached by forensic scientists. Similarly, the *Reference Manual on Scientific Evidence* defines a valid measurement as one that "measures what it is supposed to." See David H. Kaye & David A. Freeman, Reference Guide on Statistics, *in* Reference Manual on Scientific Evidence 103 (2d ed. 2000). John Monahan and Laurens Walker define validity in the context of inferences drawn from research projects as "whether the methods and analyses employed were sound enough to justify the inferences drawn by the researcher." John Monahan & Laurens Walker, Social Science in Law 60 (6th ed. 2006). This Article does not refer to any of the more specialized uses of validity in the sciences, such as internal or external validity of research. Id. at 60. The word "invalid" is used not only because of its accepted meaning in the sciences, but also because this Article examines only whether testimony was supported by data, and does not in any way characterize the ethics or the state of mind of the analysts who testified.

[18] This Article does not examine the reliability of a particular discipline or field or the validity of forensic science methods. See Samuel R. Gross & Jennifer L. Mnookin, Expert Information and Expert Evidence: A Preliminary Taxonomy, 34 Seton Hall L. Rev. 141, 146–47 (2003) (describing distinction between field and method validity). For discussion of critics and defense of such disciplines, see infra Section II.D–E.

amines the validity of testimony—that is, whether what analysts said in court was supported by empirical data.

Examining forensic science testimony in the cases of DNA exonerees has several important limitations. Of the persons exonerated by post-conviction DNA testing, 156 had testimony concerning forensic evidence at their criminal trials. One advantage of looking at these cases is that relevant trial materials could be readily obtained. Most exonerees had forensic science testimony in their cases, because almost all were convicted at trial, and most of the cases involved rapes for which there was preserved crime scene evidence that could later be tested for DNA. Many also had post-conviction lawyers seek DNA testing who retained copies of the trial records. However, the same features that made this set an attractive subject for study also make this set unrepresentative of typical criminal cases. The data set consists entirely of erroneous outcomes, or innocent people convicted at trial. In addition, most exonerees were convicted of rape, since in such cases DNA evidence can often be highly probative to the issue of identity. Very few criminal defendants are convicted at a trial, where most plead guilty, and fewer are convicted of felony rape.[19] Most exonerees were also convicted in the 1980s, before DNA testing was common.

As a result, one cannot determine from these data whether invalid forensic science testimony was common in the past two decades or is today. These data cannot provide information about forensic testimony in other types of far more common criminal cases. Invalid forensic science testimony in wrongful conviction cases might be the tip of a much larger iceberg, but it also might not. To answer that question, a broader inquiry into testimony in other types of cases and current cases is necessary. Such an inquiry, though desirable, faces practical difficulties, as no entity systematically collects or examines forensic science testimony in criminal cases. The purpose here, having obtained data from this group of innocent convicts, is simply to describe the testimony in these trials. That testimony provides examples suggesting a worrisome problem. Without reaching any conclusions about the size of the problem, these data point to the need to further investigate the content of forensic sci-

[19] See Garrett, supra note 7, § I.A.

ence testimony, particularly where the conclusions expressed by forensic scientists on the stand are largely unregulated.

Trial transcripts were obtained for 137 of the 156 exonerees identified as having testimony by forensic analysts called by the prosecution at their trials.[20] This study observed invalid forensic science testimony in the bulk of these trials. In 82 cases, or 60%, forensic analysts called by the prosecution provided invalid testimony. This invalid testimony chiefly involved serological analysis and microscopic hair comparison, but also other forensic techniques, such as bite mark, shoe print, and fingerprint comparisons. Three additional cases involved withholding of exculpatory forensic evidence. Moreover, the invalid testimony was not the product of just a few analysts in a few states, but of 72 forensic analysts employed by 52 laboratories or medical practices in 25 states.

Two basic types of invalid science testimony occurred in these cases: (1) the misuse of empirical population data, and (2) conclusions regarding the probative value of evidence that were unsupported by empirical data. The Dotson case was an example of the first type. The analyst testified that Dotson was included in 11% of the population that could have been the semen donor, when in fact 100% of the population could have been the donor. An example of the second type of invalid testimony was in Timothy Durham's case, where the analyst opined that the particular reddish-yellow hue of his hair and the crime scene hair were found in "about 5 percent of the population."[21] No empirical data exist on the frequency of hair characteristics, and thus that statement was totally unsupported.

As courts have long recognized, forensic expert testimony can play an important role in criminal trials. Juries may give special weight to testimony by forensic scientists; the Supreme Court has cautioned that "[e]xpert evidence can be both powerful and quite misleading because of the difficulty in evaluating it."[22] These crimi-

---

[20] The findings regarding each transcript are summarized in the Appendix, which along with the transcripts themselves, has been made available online.

[21] See infra note 157.

[22] *Daubert*, 509 U.S. at 595; see also United States v. Frazier, 387 F.3d 1244, 1263 (11th Cir. 2004) ("[E]xpert testimony may be assigned talismanic significance in the eyes of lay jurors, and, therefore, the district courts must take care to weigh the value of such evidence against its potential to mislead or confuse."); United States v. Hines, 55 F. Supp. 2d 62, 64 (D. Mass. 1999) ("[A] certain patina attaches to an expert's tes-

nal trials all involved serious charges, typically rape and murder, and ten resulted in death sentences. This study makes no causal claims, however, regarding the degree to which invalid testimony contributed to wrongful convictions. Not only do we not know how jurors reached their verdicts, but these convictions were almost always supported by evidence in addition to the forensic evidence.

The advent of DNA technology has not solved the problem of invalid forensic testimony. DNA has replaced some, but not most, traditional forensic methods. Although DNA testing is now widely available in the kinds of sexual assault cases chiefly examined here, it is used in a small minority of criminal investigations. In a robbery, there is typically no semen deposited by the thief; in a drive-by shooting, no blood from the shooter may be left behind. In the overwhelming majority of cases, laboratories utilize additional forensic individualization disciplines other than DNA, some which are not unlike those that are the main subject of this study. Only two percent of law enforcement requests to crime labs involve requests for DNA analysis.[23] Nor is DNA analysis immune from inaccurate presentation of results. Several recent exonerations in our study set involved invalid trial testimony concerning DNA testing. Furthermore, this study describes only trial testimony. The incidence of faulty use or mischaracterization of the underlying data cannot be known without retesting or reexamination of the underlying forensic evidence.[24] Similarly, this study makes no conclusions about the state of mind of these analysts, which also cannot typically be known.

Unfortunately, our criminal system may not be well situated to prevent unscientific testimony. The adversarial system largely failed to police the invalid testimony during these trials. Defense counsel rarely cross-examined analysts concerning invalid testi-

---

timony unlike any other witness; this is 'science,' a professional's judgment, the jury may think, and give more credence to the testimony than it may deserve.").

[23] See Joseph L. Peterson & Matthew J. Hickman, Bureau of Justice Statistics, Census of Publicly Funded Forensic Crime Laboratories 6 (2002), http://www.ojp.usdoj.gov/bjs/pub/pdf/cpffcl02.pdf; see also Cal. Comm. on the Fair Administration of Justice, Report and Recommendations Regarding Forensic Science Evidence 3 (2007) (referencing Barry Fisher, Director, Crime Laboratory for the Los Angeles County Sheriff Department).

[24] See infra Subsection II.F.2 (discussing few cases where retesting uncovered errors).

mony and rarely retained experts, since courts routinely deny funding for defense experts. Prosecutors, moreover, presented erroneous accounts of the forensic evidence during closing arguments. In a few cases in which the defense challenged invalid forensic science, judges seldom provided relief. Courts do not typically review testimony after finding the underlying methodology reliable and permitting the forensic analyst to take the stand. As the NAS Report explained, "the legal system is ill-equipped to correct the problems of the forensic science community."[25]

For those reasons, the scientific community is a crucial source for both research and reform. Future research should examine the incidence of invalid testimony in cases beyond the cases examined, such as cases not involving DNA exonerations, cases involving more recent criminal trials, cases in which there is no DNA available to exonerate or confirm guilt, and cases involving different forensic disciplines. More important, the scientific community should respond in a forward-looking way by not just revisiting old cases, but also by issuing national standards for written reports and testimony in the future. Currently, no national or widely accepted set of standards for forensic science written reports or testimony exists. No entity promulgates such standards or ensures that all analysts adhere to standards for permissible scientific conclusions regarding forensic evidence. The NAS Committee report examining the needs of the forensic science community provides an important opportunity for legislators, lawyers, and scientists to implement such oversight mechanisms to ensure the accurate use of forensic science in the courtroom.

This Article will proceed as follows. Part I will summarize the findings and describe both the study method and background legal and ethical principles involved. Part II will present the findings by examining each type of invalid forensic science testimony and analysis, beginning with findings regarding conventional serology and microscopic hair comparison and proceeding to findings related to additional forensic science disciplines. Part III will describe the roles of defense counsel, prosecutors, and courts, and then conclude by recommending the adoption of national standards and

[25] Strengthening Forensic Science, supra, at 1-14; see also id., ch.3.

oversight mechanisms to ensure that forensic science reports and testimony adhere to valid scientific standards.

I. SUMMARY OF FINDINGS, METHODOLOGY, AND PRINCIPLES

A. The Study Set and Summary of Findings

In 137 exonerees' trials—the group referred to below as the "study set"—trial transcripts were obtained in which forensic analysts were called to testify by the prosecution. The study set is a subset of the DNA exonerees as a whole. A total of 232 people have now been exonerated by post-conviction DNA testing.[26] One hundred fifty-six exonerees were identified as having trials in which forensic evidence was presented (three more pleaded guilty).[27] Efforts were made to obtain trial transcripts for the first 220 exonerees by contacting post-conviction attorneys, innocence projects and court clerks.[28] Of the 156 exonerees identified as having had fo-

[26] The study group only includes individuals who were exonerated, meaning that their conviction was vacated by a court or they received an executive pardon after DNA test results excluded them, and they were not retried. That group does not include non-DNA exonerations, including persons exonerated based on non-DNA forensic science. For additional discussion of the meaning of the term "exoneration," see Garrett, supra note 7, at 64 n.33.

[27] This is a higher proportion of exonerations than previously reported. The first author's *Judging Innocence* study of the first 200 post-conviction DNA exonerations identified 113 cases supported by forensic science, or 57% of that sample set. See Garrett, supra note 7, at 81. There are several reasons why far more cases were identified in this study. First, cases were identified among the most recent exonerations.

Second, once trial transcripts were obtained and reviewed, new cases were identified that contained forensic science testimony. *Judging Innocence* did not examine such trial records, but rather judicial decisions and news reports. See id. at 66. Those sources did not mention that there was forensic science testimony during some of these trials.

Third, *Judging Innocence* examined only cases in which forensic evidence supported the state's case, because there the focus of the study was on whether that evidence was challenged post-trial—obviously an exoneree would not challenge exculpatory evidence. This study, because it focuses on the trial testimony itself, also includes 19 cases in which the state introduced forensic evidence at trial, even though that evidence did not support the state's case, but was rather non-probative or exculpatory. This study includes such cases to present a balanced picture of the testimony concerning forensic science. After all, many of the cases with invalid testimony should properly have been cases in which the forensic science was presented as non-probative or exculpatory.

[28] The authors stopped making systematic efforts to locate additional materials as this Article approached publication in October 2008. The authors note that one additional transcript not included in the study set has been obtained since that time: that

rensic testimony at their trial, trial transcripts for 137 exonerees are studied here (14 were not located).[29] The 137 exonerees were convicted of the following crimes: rape (95 individuals), both rape and murder (33), murder (8), and attempted murder (1). Thus, the vast majority (128, or 93%) of the cases in the study set involved a sexual assault.

The testimony of forensic analysts in the 137 trials in the study set was reviewed, as summarized in the Appendix;[30] the transcripts have been made available online.[31] In each of these 137 trials, forensic analysts were called to testify by the prosecution. Most of those analysts were employed by state or local law enforcement crime laboratories.[32] There are over 350 crime laboratories in the United States.[33] The vast majority are operated by law enforcement

---

of David J. Bryson; that transcript included invalid testimony concerning hair comparison. At Bryson's trial, analyst Joyce Gilchrist testified that "it would be impossible not to be able to distinguish hairs from two different individuals," in effect asserting that human hairs are microscopically unique. Trial Transcript at 341, State of Oklahoma v. David Johns Bryson, No. CRF-82-5031 (Okla. D. Ct., Feb. 7, 1983).

The authors thank Michelle Morris and Kent Olson, reference librarians at the University of Virginia School of Law, for their extraordinary efforts to locate trial transcripts. The authors also thank Winston & Strawn, LLP for parallel efforts to locate each exoneree's trial transcripts.

[29] The 14 transcripts could not be obtained because no transcript was prepared on appeal, the transcript had been lost, or the case has been sealed. An additional 5 exonerees' trials involved only defense experts, whose testimony focused on exculpatory evidence. As discussed further, 19 trials had at least some defense expert testimony, and all 19 testified properly, often identifying flaws in testimony by analysts called by the state. See infra Subsection I.A.2.

[30] Throughout, this Article uses the term "forensic analysts" to refer generally to persons providing expert testimony regarding forensic evidence at trial. Those experts may be doctors, dentists, criminalists, police examiners, or have other professional titles.

[31] Exoneree Trials: Testimony by Forensic Analysts, http://www.law.virginia.edu/html/librarysite/garrett_exoneree.htm. The Appendix, trial materials, and other reports associated with exonerees' cases cited in this paper can all be found at the webpage. The authors thank Jon Ashley and Mary Wood for their invaluable assistance in creating the webpage.

[32] The exceptions were in cases in which FBI employees or analysts from private firms testified. For a history of the development of crime laboratories in the United States, see Paul C. Giannelli, Regulating Crime Laboratories: The Impact Of DNA Evidence, 15 J.L. & Pol'y 59, 61–67, 72 (2007).

[33] Id. at 70; see also Peterson & Hickman, supra note 23, at 2.

agencies as state or regional laboratories, though some are operated by local governments in large metropolitan areas.[34]

In conducting a review of these 137 exonerees' trial transcripts, this study found invalid forensic science testimony was not just common but prevalent. This study found that 82 cases—60% of the 137 in the study set—involved invalid forensic science testimony.

This study focused on trial testimony, but noted instances in which it was later uncovered that the analyst withheld exculpatory forensic evidence. Withholding is not apparent from the trial testimony, but in 13 cases the concealment was later uncovered during post-conviction review, investigations, or civil discovery.[35] Three of those cases did not involve invalid testimony; thus a total of 85 cases—or 63% of the 137 cases—involved either invalid testimony or withholding of exculpatory evidence.

The testimony at these 137 exonerees' criminal trials chiefly involved serological analysis (100 cases) and microscopic hair comparison (65), because most of these cases involved sexual assaults for which such evidence was commonly available at the time. Indeed, in many cases, where both hair and semen were recovered from the crime scene, both disciplines were utilized. Some cases also involved testimony concerning: fingerprint comparison (13 cases), DNA analysis (11), forensic geology (soil comparison) (6), forensic odontology (bite mark comparison) (6), shoe print comparison (4), fiber comparison (2), voice comparison (1), and fingernail comparison (1).

In the two main categories of evidence present in the study set, serology and hair comparison testimony, this study found the following: Of the 100 cases involving serology in which transcripts were located, 57 cases, or 57%, had invalid forensic science testimony. Of the 65 cases involving microscopic hair comparison in

[34] There are, in addition, several federal laboratories, most notably the FBI lab, which is "the Nation's largest publicly funded forensic crime laboratory." Peterson & Hickman, supra note 23, at 2.

[35] For a discussion of these cases, see infra Section II.F. Thirteen total cases involved concealment of forensic evidence or analysis. Eleven also involved invalid testimony. These cases do not include at least 5 exonerees' cases in which it was withheld at trial that a prosecution witness had been hypnotized. Those cases are those of E. Honaker, L. Jean, L. Mayes, and G. Woodall. Forensic use of hypnosis involves uses unrelated to the identity of the perpetrator of a crime.

which transcripts were located, 25 cases, or 38%, had invalid forensic science testimony.

Table 1, below, summarizes the incidence of invalid trial testimony by forensic analysts in the cases for which transcripts were located. (Ten cases involved more than one type of invalid testimony.)

Type of Forensic Analysis	Cases with trial transcripts	Cases involving invalid science testimony	Percentage of cases with trial transcripts involving invalid science testimony
Serology	100	57	57
Hair comparison	65	25	38
Soil comparison	6	0	0
Fingerprint comparison	13	1	8
Bite mark comparison	6	4	67
Shoe print comparison	3	1	33
DNA testing	11	3	27
Voice comparison	1	1	100

Table 1: Invalid Forensic Science Testimony by Type of Analysis

The cases involving proper testimony are also useful to examine. Many of those cases involved non-inculpatory testimony. Of the 55 cases in which all testimony was valid, 22 contained the testimony of forensic analysts who presented only evidence that was non-probative (13 cases) or exculpatory (11). Thus, almost half of the valid forensic testimony was not inculpatory and likely did not significantly support the conviction.

In contrast, most of the invalid forensic testimony involved evidence presented as inculpatory. In just 2 of the 82 cases with invalid testimony, the analysts testified that all of the forensic evidence was non-probative or inconclusive; in fact that evidence was exculpatory. The forensic testimony would have played a reduced role in many more of the 82 cases had forensic analysts accurately presented the evidence.

Three additional cases for which materials were obtained involved guilty pleas and no trial transcript. Two of those cases also involved invalid forensic analysis later exhibited in criminal trials

of co-defendants.[36] Those cases should trouble us since the vast majority of criminal cases are resolved through guilty pleas.[37]

B. Study Protocol and Types of Invalid Testimony Identified

The authors established a protocol in advance to review the testimony, and created categories used to evaluate each transcript.[38] The authors were the primary reviewers of these transcripts,[39] but law student research assistants unfamiliar with these cases were trained on the protocol and reviewed each case as well.[40] As noted in the introduction, two basic categories of invalid science testimony recurred in these cases: (1) the misuse of empirical population data and (2) conclusions regarding the probative value of evidence in the absence of empirical data. The study protocol further divided testimony into three sub-types for each of those two categories. The Appendix lists how each case was categorized. The testimony itself is available for review online and Part II describes examples of each type of testimony. Below are the six types of invalid testimony that were identified.

1. Non-Probative Evidence Presented as Probative

The first category is the inaccurate use of empirical population data. The first and most common type of invalid testimony in this

---

[36] The cases of Christopher Ochoa and Bradford Marcellius each contained false confessions and involved invalid serology analysis later introduced in trials of co-defendants. The third case, that of James Ochoa, included DNA analysis and fingerprint analysis excluding him, but also dog scent identifications of him (although it is equivocal whether dog scent identification should be considered a form of forensic analysis).

[37] See discussion in Garrett, supra note 7, at 74 ("All but the nine who pleaded guilty in the innocence group (96%) were convicted at criminal trials. In contrast, 68% of murder convictions and 84% of felony rape convictions were obtained through plea bargaining.").

[38] Edward Blake, a forensic scientist, and scientists including Eric Lander and Richard Lewontin reviewed these categories.

[39] Both authors have represented exonerees included in the study sample. When in law practice, Garrett assisted with civil cases brought by four of these exonerees. Neufeld and the Innocence Project that he co-directs assisted in the exonerations of many of these exonerees.

[40] The students did not conduct a review that was blind to the authors' coding. However, they were instructed to review whether the transcripts were coded properly, and they reviewed the full testimony by each analyst.

category, present in 48 cases, was the interpretation of non-probative evidence as inculpatory evidence. That is, the testimony disregarded that the evidence was non-probative, and instead the analyst provided a statistic purporting to include the defendant and implying that a percentage of the population was excluded. The Dotson case described earlier provides an example of this type of invalid testimony. In a typical rape case, the evidence most likely to provide information about the genetic markers of the rapist is the vaginal swab collected immediately after the rape, when the victim is examined at a hospital. The analyst tested that evidence and testified that both Dotson and the semen donor possessed the B blood type, a type shared by only 11% of Caucasians. Eleven percent of Caucasians possess the B type; well-defined databases, developed over decades, provided the distribution of the four ABO blood group types in various racial and ethnic groups. However, that 11% statistic was invalid in the context of a rape prosecution, for it was not the combined frequency of all possible blood group types potentially possessed by the semen donor. Unlike today's DNA testing that can isolate and amplify very small amounts of genetic material, a major shortcoming of conventional blood grouping was that one could not separate the female contribution from the semen in a mixed stain present in a typical rape case. Therefore, one would generally not know if there was sufficient semen in the sample such that one would expect to detect its genetic markers. If there was not enough semen in the sample, only the victim's genetic markers would be observed. The analyst did not tell the jury that because the victim was also Type B, where her fluids were mixed in the sample, her Type B blood group substances could have masked any substances from the semen. The evidence was totally non-probative. In the Dotson case, the analyst should have told the jury that 100% of males could have been the donor. Part II describes this type of invalid testimony further; it involves the well-known problem of masking and non-quantification.[41]

In a related set of serology cases, moreover, the analysts testified that they observed no blood group substances in the crime scene samples. Rather than conclude that the contributor could have been any type because the evidence was potentially degraded,

[41] See infra Subsection II.A.1.

these analysts testified that a defendant who did not secrete blood group substances was affirmatively included.

2. Exculpatory Evidence Discounted

A second type of invalid testimony occurred in 23 cases in which exculpatory evidence was discounted. For example, in Paul Kordonowy's case, serological tests of the victim's underpants revealed Type A antigens, which neither the victim nor Kordonowy possessed. Rather than testify that Kordonowy was excluded by the finding inconsistent with his type, the analyst told the jury to disregard that exculpatory evidence, and instead made the unsupported claim that bacteria could somehow have changed the reading and produced the Type A antigens.[42] In other serology cases and several cases involving hair comparison, analysts similarly discounted exculpatory results and claimed to reach a non-probative or inculpatory result, often by relying on pure speculation.

3. Inaccurate Frequency or Statistic Presented

In a third type of invalid testimony present in 13 cases, the frequency or statistic presented was erroneous. In several exonerees' trials, analysts falsely divided frequencies in half. For example, in the Perry Mitchell case, the semen was left by a Type O secretor, and Type O secretors comprise 35% of the population. The serologist divided the accurate frequency in half and testified that only 17.5% of men could have contributed the semen and thus 82.5% of the relevant population was excluded.[43] However, population statistics regarding ABO blood group substances are identical for both sexes; 35% of both men and women are Type O secretors; thus, it was erroneous to divide that statistic in half.

4. Statistic Provided Without Empirical Support

The second major category of invalid testimony includes conclusions unsupported by any empirical data. In a fourth type of invalid testimony present in 5 cases, statements were made providing a frequency or probability in the absence of any empirical support.

---

[42] See infra Subsection II.A.2.
[43] See infra Subsection II.A.3.

Some forensic disciplines involve more subjective analyses not premised on empirical population data. For example, in the field of microscopic hair comparison, no adequate empirical data exist regarding the frequency of microscopic characteristics of human hairs. Yet in the Bromgard case, the analyst testified that there was a 1 in 10,000 chance that the two hairs found at the crime scene could come from someone other than Bromgard.[44] Those frequency statistics were simply made up by the analyst.[45]

5. Non-numerical Statements Provided Without Empirical Support

In a fifth type of invalid testimony present in 19 cases, non-numerical statements of probability or frequency were offered despite a lack of any empirical data. In the field of microscopic hair comparison, due to the lack of empirical data, the field adopted standards that the strongest statement of association that can be made by an analyst is that the hairs in question are "consistent" with the defendant's or "could have" come from the defendant.[46] All analyst testimony, therefore, stating that a crime scene hair was "highly likely" to have come, "very probably" came, or did come from the defendant violates the basic scientific criterion that expressions of probability must be supported by data. For example, in the Calvin Scott case, the analyst testified that the chance that another person could have similar hair was remote, explaining, "I would not give a figure. It would be quite large."[47] Use of such probability, frequency, or other individualizing statements was unsupported.

---

[44] See infra Subsection II.B.1.

[45] See infra Section II.B.

[46] To say that two items are "consistent" without being able to tell the jury that consistency is rare or common, renders the evidence potentially misleading and hence raises questions whether it is inadmissible as both irrelevant and unduly prejudicial. This study does not address evidentiary criteria—such as whether such testimony would be admissible under Federal Rules of Evidence 402 or 403—nor whether such testimony would satisfy *Daubert*. Other commentators have done so and courts should examine such questions carefully. See infra Section II.B. This study, however, is concerned only with the scientific validity of the testimony.

[47] See infra Subsection II.B.2.

6. Conclusion that Evidence Originated from Defendant

The sixth and final type of invalid testimony, present in 6 cases, claimed that the evidence did in fact come from the defendant and was unique to the defendant, despite no empirical data permitting such conclusions. For example, in Ray Krone's case, the analyst testified that the bite marks did in fact come from Krone's teeth, telling the jury, "that tooth caused that injury."[48] In two other cases the forensic odontologists (forensic dentists) were unequivocal that the defendants' teeth made the bite marks on the victim.[49] Forensic disciplines involving impression evidence, such as bite mark and shoe print comparison, have not developed any objective criteria at all by which to judge assertions about the likelihood that crime scene evidence came from a particular defendant.[50] Nor do any empirical data exist to support a claim that a bite mark is uniquely identifiable as belonging to a particular person.

* * *

These six types of invalid testimony may occur in other disciplines not reviewed here, and conversely, additional types of invalid forensic testimony may occur in cases not in the study set. As noted at the outset, this study cannot speak to questions concerning how often invalid forensic science testimony occurs in other types of more typical criminal cases. The study set is limited not only to DNA exonerees, but also to trials resulting chiefly in rape or rape and murder convictions in the 1980s. Perhaps such cases involving felonies in contentious cases that proceeded to trial were more likely to involve pressures on the state to overstate the evidence, including forensic evidence, making these cases unrepresentative of more common and less serious crimes. On the other hand, perhaps such cases did not involve such pressure to overstate forensic evidence. Perhaps there would be little pressure to overstate forensic evidence if the defense did not meaningfully contest forensic evidence. If so, these cases might be representative of a more

[48] See infra Section II.D.

[49] See infra Section II.D.

[50] See Michael J. Saks & Jonathan J. Koehler, The Individualization Fallacy in Forensic Science, 61 Vand. L. Rev. 199 (2008).

widespread problem. Those questions can not be answered by reviewing just the trials of DNA exonerees.

After all, the particular forms of forensic analysis reviewed reflect the make-up of the cases in the study set. Almost all exonerees in the study set were convicted of rape or rape and murder. This is not because the quality of forensic testimony is worse in rape cases, but rather because DNA testing could be later used to identify the source of the semen left by a rapist, which is usually dispositive of guilt in cases involving stranger-perpetrators in which the central issue is the identity of the assailant. Most of the exonerees were convicted of crimes involving strangers in which the identity of the perpetrator was in question at trial, and ultimately it was shown through post-conviction DNA testing that in fact the wrong person was convicted. Therefore, the study set disproportionately included evidence that one would expect in a rape case: serology analysis of material collected as part of a rape kit and microscopic hair comparison of hairs found at the crime scene, often from combings of the victim or the victim's clothes.

This explains why so many of the trials studied involved semen or hair evidence and also why there were few trials studied involving fiber analysis, bite marks, fingerprints, toolmark analysis, and other forensic disciplines. Such other forensic disciplines do not routinely examine evidence common in a sexual assault case.[51] Nor does evidence such as a latent fingerprint typically have preserved relevant biological evidence that can later be tested using DNA analysis.[52] These exonerees all had cases in which such evidence

[51] The one study to examine which types of forensic testing were conducted in different types of felony investigations supports this conclusion. See Joseph L. Peterson et al., Forensic Evidence and the Police, 1976–1980, Nat'l Archive of Crim. Just. Data, Inter-University Consortium for Political and Social Research, Study No. 8186 (1985). That study developed types of forensic analysis in over 1700 felony investigations conducted in four urban police departments from 1976 through 1980. The data sheets were analyzed with the always outstanding assistance of University of Virginia Reference Librarian Jon Ashley. The 183 rape cases in that set chiefly had serological analysis of blood (68 cases), semen (153), or hair evidence (87 with pubic hair and 55 with head hair). Id. Comparatively few rape cases had fiber analysis (5 cases), latent print analysis (46), bullet analysis (4), or impression analysis (5). Id. In contrast, as one would expect, few of the 223 murder cases had semen analysis (5 cases), but many involved analysis of bullets (142) or latent fingerprints (94).

[52] In the Peterson data set, the cases with biological evidence, such as serology evidence, were far more common in cases with hair evidence (86%) than in cases with,

was in fact collected at the crime scene, because each was exonerated when the DNA testing was later conducted on that material, typically from a rape kit or on certain hair evidence.[53] This study also does not include forensic analysis unrelated to the issue of identity introduced to show how a crime occurred or that it occurred, such as autopsy evidence.[54] Thus, the role of particular forensic disciplines as well as the role of invalid forensic science in the cases studied here would be different for other types of criminal cases, and even for other types of cases in which identity is at issue. For example, the study set cases did not typically involve analysis of bullets that one would expect in cases involving shootings, or tire tread analysis that one would expect in cases involving vehicular assault.

Forensic evidence in the vast majority of criminal cases that result in guilty pleas does not receive the scrutiny of a trial. However, the set of DNA exonerees in this study consists of persons convicted at a trial. The cases studied here not only involved trials, but they mostly involved trials in the 1980s. Today, issues of identity in sexual assault cases may often be resolved through DNA testing pre-trial, making it less likely that some of the invalid testimony observed regarding hair comparison or serology would occur.[55] However, other non-sexual assault cases are not as susceptible to DNA testing and may present some of the same issues implicated

---

for example, fingerprint evidence (28%). Id. Further, "although it is now possible, in the laboratory, to extract DNA from a fingerprint, this has not been done in the field, and it would certainly not be possible with a fingerprint that has aged in an evidence locker." Simon A. Cole, More Than Zero: Accounting for Error in Latent Fingerprint Identification, 95 J. Crim. L. & Criminology 985, 1026 (2005) (citation omitted).

[53] See Garrett, supra note 7, at 73 ("The 200 exonerees were charged and convicted chiefly of rape (71%), murder (6%), or both murder and rape (22%). This is not surprising; rape cases in particular often have relevant biological material for DNA testing.").

[54] False pathology evidence, for example, could lead to wrongful convictions where no murder in fact occurred but rather the death was due to natural causes. See, e.g., Mark Bonokoski, Editorial, Experts Must be Impartial, The Daily Observer, Feb. 7, 2008 (describing the work of Charles Smith, "an expert witness (supposedly) in forensic pathology who lied, invented, forgot, pretended, withheld, dismissed, neglected, guessed—and, as a result, sent many people to jail for crimes that never happened"). Cases in which no crime in fact occurred do not raise issues regarding the identity of the perpetrator for which post-conviction DNA testing would lead to exoneration. As a result, no such cases were present in the study set.

[55] See Brandon L. Garrett, Claiming Innocence, 92 Minn. L. Rev. 1629, 1634 (2008).

in these exonerees' cases. Latent fingerprint comparison, for example, is still in wide use; indeed it is used far more often than DNA testing.[56] New forensic techniques continue to be developed that involve the same sorts of subjective comparison not grounded in empirical data, which might then risk invalid testimony if analysts do not inform the jury that no probability or frequency can be supported. For example, the FBI is developing a new technology which it claims can identity unique characteristics of human voices.[57]

Further study is necessary to assess questions regarding incidence of invalid testimony in recent trials, in trials not involving sexual assaults, and in trials not involving wrongful convictions. The next Section describes what limited information is available concerning such questions.

C. Questioning the Incidence of Invalid Forensic Testimony

Senator Orrin Hatch, commenting on the need to provide new resources for forensic sciences, referring to the fraudulent work of Oklahoma City police department forensic analyst Joyce Gilchrist that contributed to several wrongful convictions, noted:

> [W]e are all troubled by allegations that mistakes by a police chemist in Oklahoma helped send innocent people to prison. This isolated situation should not be used unfairly to indict the thousands of forensic scientists who perform their work professionally and responsibly. It should, however, remind us that those who work in our criminal justice system have an obligation to be diligent, honest, and fair-minded.[58]

While not disagreeing with that statement, this study describes how the invalid testimony in DNA exoneration cases did not just

---

[56] See Integrated Automated Fingerprint Identification System, http://www.fbi.gov/hq/cjisd/iafis.htm [hereinafter IAFIS]; Peterson & Hickman, supra note 23, at 6–9.

[57] Dina Temple-Raston, Voice "Fingerprints" Change Crime-Solving, Nat'l Pub. Radio, Jan 18, 2008, http://www.npr.org/templates/story/story.php?storyId=18479095.

[58] See DNA Crime Labs: The Paul Coverdell National Forensic Sciences Improvement Act: Hearing Before the Comm. on the Judiciary, 107th Cong. 2–3 (2001) (statement of Sen. Orrin G. Hatch, Chairman, Comm. on the Judiciary), available at http://frwebgate.access.gpo.gov/cgi-bin/getdoc.cgi?dbname=107_senate_hearings&docid=f:78008.pdf.

involve a few "bad apples," like Gilchrist, who have been the subject of high profile investigations.[59] Several forensic analysts testified in more than one trial in the study set, including Pamela Fish (5 trials), Arnold Melnikoff (3), Joyce Gilchrist (3), and Fred Zain (6). However, 61 of the analysts who delivered invalid testimony did so in just one trial in the study set. The study set included invalid testimony by 72 forensic analysts called by the prosecution and employed by 52 laboratories, practices, or hospitals from 25 states.[60]

This study does not examine the state of mind of forensic analysts.[61] Invalid testimony could be explained not by intentional or reckless acts, but rather by inexperience, poor training, or inadequate supervision.[62] If these particular analysts lacked adequate training or supervision, then one wonders about their testimony in other cases as well as testimony by their colleagues. Most crime laboratories do not employ more than a dozen analysts; each one of these analysts could have testified in many cases each year.[63] Indeed, in many of the trials studied, the analysts, when describing their credentials, stated that they had testified on numerous occasions, sometimes even in hundreds of trials.[64]

---

[59] See William C. Thompson, Beyond Bad Apples: Analyzing the Role of Forensic Science in Wrongful Convictions, 37 Sw. U. L. Rev. 101, 112-119 (2009).

[60] The states are: Arizona, California, Connecticut, Florida, Georgia, Idaho, Illinois, Indiana, Kentucky, Louisiana, Maryland, Massachusetts, Mississippi, Missouri, Montana, Nevada, New Jersey, New York, Ohio, Oklahoma, Pennsylvania, South Carolina, Texas, Virginia, and West Virginia.

[61] Other studies develop questions regarding possible bias or observer effects, where forensic analysts are typically employed by law enforcement. D. Michael Risinger et al., The *Daubert/Kumho* Implications of Observer Effects in Forensic Science: Hidden Problems of Expectation and Suggestion, 90 Cal. L. Rev. 1, 5–6 (2002).

[62] Samuel R. Gross, Expert Evidence, 1991 Wis. L. Rev. 1113, 1178 ("The worst that can be said about an expert opinion is not that it is a lie—that criticism is often beside the point—but that it is unreasonable, that no competent expert in the field would hold it. Correspondingly, the most dangerous expert witness is not one who lies (although she may do that too), but one who is ignorant or irresponsible.").

[63] See Peterson & Hickman, supra note 23, at 3.

[64] For example, David Brody, in Neil Miller's case, testified that he had testified in court during his career "at least a thousand times." Trial Transcript at 190, Commonwealth v. Neil Miller, No. 92-P-612 (Okla. Dist. Ct. Dec. 14, 1990) [hereinafter Transcript, N. Miller]. Elmer Gist testified in the Honaker case that he had "testified in western Virginia, northern Virginia, the Tidewater area, approximately a hundred and eighty times total." Trial Transcript at 212, Commonwealth v. Edward William Honaker, No. 2159–75 (Va. Cir. Ct., April 10, 1985) [hereinafter Transcript, Honaker].

Nor does the prevalence of invalid forensic testimony in these trials speak to what "caused" these wrongful convictions. Though each case involved an erroneous outcome—an innocent person convicted—invalid forensic science testimony may not have been the deciding factor leading juries to convict. In addition to the forensic evidence, other evidence—particularly eyewitness identifications—supported most of these convictions. Forensic science testimony might not by itself "cause" a conviction where criminal trials typically involve multiple pieces of evidence and actions by several actors.[65] For example, Gary Dotson might still have been convicted even if the forensic analyst had correctly observed that any male could have been the semen donor. Among other evidence in the case, the victim had identified him as the rapist. The forensic analyst's invalid forensic testimony did serve some role in buttressing the false eyewitness identification, yet one cannot typically know how jurors weighed the evidence in reaching the decision to convict. As noted, several of these trials involved forensic evidence—in a few cases DNA evidence—that excluded the defendant, and yet the state still secured the conviction. However, courts and scholars have long recognized that jurors may place special trust in scientific evidence.[66] Studies also suggest that the manner in which the forensic evidence is presented to the jury impacts how jurors weigh that evidence.[67]

---

[65] See Garrett, supra note 7, § II.A.

[66] Daubert v. Merrell Dow Pharm., Inc., 509 U.S. 579, 595 (1993); see also United States v. Frazier, 387 F.3d 1244, 1263 (11th Cir. 2004) ("[E]xpert testimony may be assigned talismanic significance in the eyes of lay jurors, and, therefore, the district courts must take care to weigh the value of such evidence against its potential to mislead or confuse."); United States v. Hines, 55 F. Supp. 2d 62, 64 (D. Mass. 1999) ("[A] certain patina attaches to an expert's testimony unlike any other witness; this is 'science,' a professional's judgment, the jury may think, and give more credence to the testimony than it may deserve."); Gross, supra note 62, at 1179–81 (reviewing empirical research regarding degree to which juries rely upon and comprehend expert evidence).

[67] See, e.g., Jonathan J. Koehler, The Psychology of Numbers in The Courtroom: How to Make DNA-Match Statistics Seem Impressive or Insufficient, 74 S. Cal. L. Rev. 1275 (2001); Dawn McQuiston-Surrett & Michael J. Saks, Communicating Opinion Evidence in the Forensic Identification Sciences: Accuracy and Impact, 59 Hastings L.J. 1159 (2008); John Monahan et al., Communicating Violence Risk: Frequency Formats, Vivid Outcomes, and Forensic Settings, 1 Int'l J. Forensic Mental Health 121, 126 (2002).

Two of the exonerees' cases involved bench trials, providing information about how the fact-finder reasoned. In the Willie Davidson case, invalid forensic science appeared to have misled the fact-finder. The judge emphasized that the guilty verdict was "supported by that scientific evidence and [the victim's] identification."[68] But when explaining the scientific evidence, he appeared confused, understandably so, because improper testimony concerning the serology had ignored the problem of masking and quantification. The judge stated: "Then it had the type of a non-secretor. The defendant is a non-secretor. That by itself isn't totally conclusive. Forty-two percent are of that, so that doesn't nail it down."[69] Actually, 42% was not a proper statistic. No male could be excluded by the serological techniques used at the time. Where the victim was Type O and Type O material was observed, the blood group substances could have solely originated from the victim, and thus any person could have been the semen donor. Separately, in Nathaniel Hatchett's case, powerful exculpatory forensic evidence was disregarded. DNA testing conducted before trial on the semen evidence from a single-perpetrator rape had excluded Hatchett. Nevertheless, the judge in the bench ruling found the DNA results not dispositive where Hatchett had confessed, stating, "in light of the overwhelming evidence that the Court has . . . the Court does not find that the laboratory analysis is a fact which would lead to a verdict of acquittal."[70]

Again, this study's data do not support claims about the incidence of invalid forensic science testimony in cases outside of the 137 trials studied, but rather points to the need to investigate the nature of the problem. Some evidence from cases outside this study set also suggests that this problem deserves further attention, and that invalid forensic testimony may not be associated with wrongful convictions, but rather may be part of a different and larger problem. Studies have found high error rates in a series of forensic dis-

[68] Trial Transcript, Commonwealth of Virginia v. Willie Davidson, No. 919–81 (Va. Cir. Ct., May 27, 1981) (page numbers illegible).

[69] Id.

[70] Trial Transcript at 276–77, State of Michigan v. Nathaniel Maurice Hatchett, No. 97-1496-FC (Mich. Cir. Ct., Mar. 3, 1998).

ciplines.[71] Such studies may shed light on the reliability of the underlying method or its application by forensic practitioners, but they do not shed light on whether trial testimony comports with scientific standards. Indeed, few have studied testimony by forensic analysts. One of the purposes of this Article is to encourage future efforts to review and improve the quality of forensic science testimony.

One reason that compilations of more systemic data concerning the quality of forensic testimony during criminal trials are lacking is that crime laboratories do not routinely collect or review such testimony. Even after these DNA exonerations, not only have investigations into these individual cases often not occurred, but investigations regarding systemic problems in laboratories remain rare. When our system has investigated laboratories in response to these exonerations, systemic problems have been uncovered. Noteworthy examples include the Houston Police Department investigation led by Michael Bromwich that uncovered hundreds of cases involving invalid serology analysis beyond the two post-conviction DNA exonerations that sparked the investigation.[72] Similar audits have occurred in reaction to DNA exonerations at laboratories in Cleveland, Ohio, and Baltimore, Maryland, and laboratories in Oklahoma, West Virginia, and Virginia.[73]

What little information does exist regarding cases outside our study sample does not provide cause for optimism. Simon Cole has conducted a preliminary effort, examining 34 transcripts involving latent print testimony, finding "over-claiming," or expert testimony exaggerating its own probative value, prevalent in that group of cases.[74] Another example is the recent National Research Council report, which uncovered invalid testimony by FBI analysts who tes-

---

[71] See Joseph L. Peterson & Penelope N. Markham, Crime Laboratory Proficiency Testing Results, 1978–1991, II: Resolving Questions of Common Origin, 40 J. Forensic Sci. 1009, 1010 (1995); see also Evidence of Injustice, CBS News, Sept. 14, 2008, at http://www.cbsnews.com/stories/2007/11/16/60minutes/main3512453_page4.shtml (noting that a review of 100 FBI cases with trial testimony concerning bullet lead analysis uncovered that "almost half" involved flawed testimony).

[72] See Michael R. Bromwich, Executive Summary, *in* Fifth Report of the Independent Investigator for the Houston Police Department Crime Laboratory and Property Room 1–2 (May 11, 2006), available at http://www.hpdlabinvestigation.org.

[73] See Possley, supra note 9, § 1, at 1, 20–21.

[74] See Cole, supra note 16; see also Saks & Koehler, supra note 50, at 205–06.

tified for decades that bullets "must have come from the same box" without any empirical support.[75] The Detroit Police Department Crime Laboratory was recently closed based on a "systemic" failure to properly conduct firearms analysis, a type of analysis not studied here.[76] In several disciplines involving impression evidence, as developed below, the relevant disciplines provide guidelines regarding trial testimony that explicitly permit invalid testimony not based on empirical evidence.[77]

Nor is it difficult to find a host of reported appellate decisions describing invalid forensic science testimony similar to that in these exonerees' trials. Reported decisions regarding invalid serology, hair comparison, fingerprint comparison, and bite mark comparison testimony can readily be found on Westlaw, and numerous such cases are collected in treatises on scientific evidence.[78]

Our quite preliminary effort to test whether the testimony in these exonerees' trials is representative of testimony in similar trials suggests that invalid testimony was also common in trials in which there has been no DNA exoneration, involving similar rape and murder charges and from the same time period. To date, 30 trial transcripts in such "matched" cases have been collected from Missouri (10 transcripts), Texas (11), and Virginia (9). Almost two-thirds of those trials exhibited invalid forensic science testimony, including the same types observed in the exonerees' trials, and including testimony by some of the same analysts who testified in the

---

[75] Comm. on Scientific Assessment of Bullet Lead Elemental Composition Comparison, Nat'l Research Council, Forensic Analysis: Weighing Bullet Lead Evidence 90–94 (2004); see also Cole, supra note 16, at 820.

[76] See Nick Bunkley, Detroit Police Lab Is Closed After Audit Finds Serious Errors in Many Cases, N.Y. Times, Sept. 25, 2008, at A17.

[77] See Michael Bowers, The Scientific Status of Bitemark Comparisons, *in* Modern Scientific Evidence 538, 549–50 (David L. Faigman et al. eds., 2002); Iain A. Pretty, Reliability of Bitemark Evidence, *in* Bitemark Evidence 531, 543 (Robert B.J. Dorion ed., 2005); see also Cole, supra note 16, at 820–22 (discussing "institutionalized overclaiming").

[78] See infra notes 301–02; 2 Paul C. Giannelli & Edward L. Imwinkelried, Scientific Evidence § 24-3 (4th ed. 2007) (describing and citing to a "massive body of case law" admitting testimony regarding hair comparison, including testimony found here to be invalid, such as use of probabilistic statements); see also id. § 24-5 (describing reported cases reviewing fiber comparison testimony); 1 Paul C. Giannelli & Edward L. Imwinkelried, Scientific Evidence § 13-5 (4th ed. 2007) (digesting case law concerning bite mark comparison).

exonerees' trials.[79] Such matched cases likely do not involve innocent convicts, but rather guilty convicts who also had invalid forensic testimony presented at their trials.

Neither matched cases involving likely correct outcomes, nor most cases involving wrongful convictions, tell us about false negatives: cases in which invalid forensic analysis led to guilty persons going free. Studies of proficiency testing of forensic laboratories, however, suggest that false negatives are far more common than false positives, and also that error rates may be generally high across a wide range of forensic techniques, including those studied here.[80]

Thus, preliminary evidence suggests that even if most of the forensic science testimony in DNA exonerees' trials was invalid, such invalid testimony may not be associated with wrongful convictions. More troubling, it may be a phenomenon in serious criminal trials generally, at least during the time period in question. However, that question can not be definitively answered nor can the more difficult question of whether such testimony is common in more typical criminal cases. Future research should investigate the incidence of invalid forensic science testimony.

---

[79] See Garrett, supra note 7, § I.B (explaining the method for selecting such "matched" cases with similar characteristics to those of the exonerees). In short, a "matched" case involved the same type of conviction in the same state and with reported decisions in the same years, but in which no DNA testing was later conducted to exonerate the defendant. The transcripts collected from these states in non-exoneration cases have been made available online at the same webpage at which the exoneree materials have been posted. Twenty of the cases involved serology testimony, 10 hair comparison, 5 fingerprint comparison, 2 bite mark comparison, and 3 involved testimony concerning DNA testing. Nineteen cases involved invalid testimony and one more involved concealment of exculpatory information that was uncovered post-trial. Thus, 63% involved invalid forensic science testimony, approximately the same percentage as among the trials of exonerees who had forensic science testimony at trial. Special thanks to Kent Olson and the Texas Defender Service for their invaluable assistance in locating these materials.

[80] Peterson & Markham, supra note 71, at 1009–11 (summarizing study results finding a series of forensic disciplines with better than 10% correct identifications in proficiency tests, but other disciplines with error rates in the 10–20% range or even higher error rates).

D. Ethics and Forensic Science Testimony

Forensic science is uniquely concerned with the introduction of evidence in courtrooms, particularly in criminal courts where the stakes can be extremely high. Thus, "criminalistics . . . has as its primary objective a determination of physical facts which may be significant in legal cases."[81] An ethical forensic analyst has a professional obligation not to mislead the jury during testimony at trial and not to mislead the state and defense when preparing forensic reports.

To the extent that a prosecutor or defense attorney asks questions that are misleading or confusing, "[t]he expert witness's obligation . . . is to give a full and complete presentation of the opinion and the reasons for that opinion," Peter Barnett writes, adding that "[t]actics on the part of either the witness or the lawyer that tend to obscure the testimony, limit the full disclosure of the basis for the testimony, or confuse or obscure the implications of the testimony are inappropriate and, under some circumstances, may be unethical or illegal."[82]

While no single ethical code applies to all practicing criminalists, much less all forensic analysts in the United States, a series of professional entities have promulgated ethical codes that shed light on testimony discussed here, including the American Board of Criminalists ("ABC"), the American Academy of Forensic Sciences ("AAFS"), and the California Association of Criminalists ("CAC").[83] As a general matter, these codes counsel independent evaluation of the evidence and truthful and non-misleading testimony in court. The ABC Code of Ethics asks that all analysts ensure that opinions are rendered "only to the extent justified" by the evidence, and to ensure that their testimony is presented "in a clear, straightforward manner" that does not "extend themselves beyond their field of competence, phrasing their testimony in such a manner so that the results are not misinterpreted."[84]

---

[81] The Code of Ethics of the California Association of Criminalists Preamble (1957), *reprinted in* Peter D. Barnett, Ethics in Forensic Science: Professional Standards for the Practice of Criminalistics 125 (2001).

[82] Id. at 81.

[83] See id. at 7, 81.

[84] Am. Bd. of Criminalistics, Code of Ethics §§ 9–10, *reprinted in* Barnett, supra note 81, at 153. The Code also states that criminalists shall "[m]aintain an attitude of inde-

The AAFS Code simply forbids a "material misrepresentation of data upon which an expert opinion or conclusion is based."[85] The AAFS Guidelines also adopt "good forensic practice guidelines," which add that "[u]nlike attorneys, forensic scientists are not adversaries. They take an oath in court to tell the whole truth. They should make every effort to uphold that oath."[86] Further, when presenting their opinions, "[e]very reasonable effort should be made to ensure that others (including attorneys) do not distort the forensic scientist's opinions."[87]

The CAC Code does not apply to most of the analysts in this study set, but in contrast to the ABC and AAFS codes, it imposes far more rigorous requirements. The CAC Code states that "[i]n all respects, the criminalist will avoid the use of terms and opinions which will be assigned greater weight than are due them. Where an opinion requires qualification or explanation, it is not only proper but incumbent upon the witness to offer such qualification."[88] The CAC Code requires that the expert indicate when an opinion "may lack the certainty of other opinions he might offer," and will "leave no false impressions in the minds of the jurors."[89] The CAC Code adds that an expert "will not . . . assign greater significance to an interpretation than is justified by the available data."[90]

---

pendence and impartiality in order to ensure an unbiased analysis of the evidence." Id. § 14.

[85] Am. Acad. of Forensic Sci., Code of Ethics and Conduct, § 3, *reprinted in* Barnett, supra note 81, at 123.

[86] Am. Acad. of Forensic Sci., Good Forensic Practice Guidelines § 13, *reprinted in* Barnett, supra note 81, at 144.

[87] Id. § 9.

[88] Cal. Ass'n of Criminalists, Code of Ethics § III.E, *reprinted in* Barnett, supra note 81, at 128.

[89] Id. § III.D.

[90] Id. § III.C. This study does not examine cases in which analysts made a "material misrepresentation of education, training, experience, or area of expertise." Am. Acad. of Forensic Sci., supra note 85, § 2. There is evidence suggesting that this would be a useful area for future study. For example, Fred Zain, who testified in five cases in this study, had performed poorly in the basic FBI serology course in 1977. However, this was not included in his personnel file, and he was promoted to supervisor of the West Virginia State Police Crime Laboratory, Serology Division, shortly thereafter. See In re Renewed Investigation of the State Police Crime Lab., Serology Div., 438 S.E.2d 501, 514–20 (W. Va. 1993) (noting also that "Zain may have testified falsely concerning his academic credentials").

Those ethical rules do not provide guidance on the permissible scope of testimony within a particular discipline; they speak to the general norms of expert conduct. Thus, those rules do not provide any scientific standards governing courtroom testimony which are the focus of this study.

E. Legal Regulation of Forensic Testimony

Courts do not typically review the presentation of forensic science testimony during criminal trials. As noted, courts recognize that jurors place special trust in expert witnesses to explain applicable scientific principles. Courts therefore regulate the matters upon which experts may testify. Thus, while a police officer could identify a defendant as the person seen committing a crime, a forensic analyst may only testify regarding an identification of a defendant using forensic methods supported by sound science. The wrongful convictions in this study occurred chiefly in the 1980s, prior to the trilogy of Supreme Court decisions heightening reliability requirements for scientific and expert testimony.[91] Under the *Frye v. United States* test that governed in federal courts and most states at the time of these convictions (since replaced in most jurisdictions by the Supreme Court's decision in *Daubert v. Merrell Dow Pharmaceuticals*), courts would permit expert testimony based only on a methodology that was "'generally accepted' as reliable in the relevant scientific community."[92] Scholars have criticized "the stunning failure of judges to provide any sort of check" on unsupported forensic evidence, describing a failure to rigorously adhere to *Daubert*'s standards in criminal cases.[93] This study does

[91] Kumho Tire Co. v. Carmichael, 526 U.S. 137 (1999); Gen. Elec. Co. v. Joiner, 522 U.S. 136 (1997); Daubert v. Merrell Dow Pharm., Inc., 509 U.S. 579 (1993).

[92] *Daubert*, 509 U.S. at 584 (quoting Frye v. United States, 54 App. D.C. 46, 47 (1923)); Alice B. Lustre, Annotation, Post-Daubert Standards for Admissibility of Scientific and Other Expert Evidence in State Courts, 90 A.L.R.5th 453, 481 (2001) (describing that the *Frye* approach is now a minority approach).

[93] David L. Faigman, Anecdotal Forensics, Phrenology, and Other Abject Lessons From the History of Science, 59 Hastings L.J. 979, 991–92 (2008). An analysis of the application of *Daubert* in its first decade reveals that while it was used frequently to exclude questionable scientific evidence in civil cases, it almost never resulted in the exclusion of forensic evidence proffered by the prosecution in criminal cases. Peter J. Neufeld, The (Near) Irrelevance of *Daubert* to Criminal Justice and Some Suggestions for Reform, 95 Am. J. Pub. Health S107, S109 (2005); see also D. Michael Ris-

not address that debate, because as those scholars point out, *Daubert* has not been carefully applied to regulate the subject of this study, the trial testimony of forensic analysts.

At least in criminal cases, having found that the underlying discipline is satisfactory and the evidence admissible following the *Frye*—or now the *Daubert*—standard, courts do not typically examine conclusions experts reach on the stand regarding whether statistical claims or other inferences drawn from the data are supported by the evidence.[94] There is no screening of the case specific inferences and opinions before the jury hears them. Yet it is precisely while the expert testifies that, as Simon Cole puts it, "the rubber meets the road," and the jury hears claims about the actual evidence in the case.[95] In the few cases where the exonerees' defense counsel raised objections to invalid forensic testimony, judges rarely limited it. When appellate attorneys challenged faulty forensic testimony, courts rarely granted relief, often finding any error to be harmless.[96]

Thus, if an expert overstates the evidence or presents it in a misleading fashion, cross-examination is relied upon to test the evidence. Yet in a criminal case, the defense is typically an unarmed adversary that lacks expert assistance. Also of crucial importance, the presentation of forensic science during criminal trials is usually one-sided, provided only by analysts testifying for the prosecution. Most states do not routinely fund the provision of forensic experts for indigent defendants, though there are strong arguments that under *Ake v. Oklahoma* defendants should be entitled to expert assistance as a matter of due process, at least in some types of cases.[97] As a result, courts routinely deny indigent defendants the funds to

---

inger, Navigating Expert Reliability: Are Criminal Standards of Certainty Being Left on the Dock?, 64 Alb. L. Rev. 99, 149 (2000).

[94] See Cole, supra note 16, at 819 ("[J]udges assume that their work is done once they have ruled proffered evidence admissible or inadmissible.").

[95] Id. at 818.

[96] See infra Subsection III.A.3.

[97] See Ake v. Oklahoma, 470 U.S. 68, 83 (1985); 1 Giannelli & Imwinkelried, supra note 78, § 4-5; Paul C. Giannelli, *Ake v. Oklahoma*: The Right to Expert Assistance in a Post-*Daubert*, Post-DNA World, 89 Cornell L. Rev. 1305, 1338–41 (2004); Gross & Mnookin, supra note 18, at 189 ("In many criminal cases, there is only one side on expert issues: the prosecution. The result is a national scandal. We have seen case after case of systematic fraud and incompetence by prosecution experts and police crime laboratories, with no end in sight.").

hire their own forensic experts. Almost all of the analysts testifying in the 137 exonerees' trials worked for police laboratories; only 19 exonerees retained experts.[98] The fact-finders in most of these cases were jurors: non-experts who could be easily misled by invalid scientific testimony. Prosecutors not only elicited invalid forensic testimony, but sometimes further misrepresented the forensic science in their closing arguments, perhaps leading the jury to draw incorrect conclusions in cases where the analyst provided proper testimony.[99]

In addition to *Daubert*, a second legal rule applicable to state experts, the Supreme Court's decision in *Brady v. Maryland*, holds that the State violates the due process rights of a defendant by withholding material, exculpatory information from the defense.[100] Expert fabrication of evidence violates the Due Process Clause as well. For example, the Court unanimously held in *Miller v. Pate* that a conviction should be set aside where the State obtained a conviction based on testimony that certain stains on underwear owned by the defendant matched the victim's blood type but where it was later shown that the stains were paint.[101] By its nature, concealed evidence rarely comes to light and violations are rarely detected, much less remedied.

Where courts do not regulate the content of expert testimony, and defendants typically do not have experts with which to effectively counter State-proferred forensic testimony in criminal trials, the scientific standards within the forensic sciences are the most important source for regulating the content of forensic science testimony. This Article next develops a series of examples in which analysts did not adhere to valid scientific standards. The Article concludes that existing regulations are not adequate to prevent invalid forensic science testimony.

II. RESULTS: INVALID FORENSIC SCIENCE TESTIMONY

The cases of the exonerees whose trials had forensic science testimony chiefly involved serology analysis of material collected as

---

[98] See infra Subsection III.A.2.
[99] See infra Subsection III.A.1.
[100] 373 U.S. 83, 87 (1963).
[101] 386 U.S. 1, 5–7 (1967).

part of a rape kit and microscopic hair comparison of hairs found at the crime scene, often from combings of the victim or the victim's clothes. The Sections that follow first develop the use of serology and hair comparison in these exonerees' trials. Next, the Article discusses additional forensic disciplines employed in smaller numbers of these cases—namely, bite mark comparison, DNA testing, and fingerprint comparison. For each type of analysis, Sections below describe the types of invalid testimony present with illustrative examples of each.

A. Invalid Forensic Serology Testimony

Of the 137 trial transcripts in the study set, 100 had testimony regarding serology analysis. Of those, 57 involved invalid testimony, 46 of which involved "masking" and quantification problems, which will be described further below.

In the "serology era" prior to the advent of DNA testing technology, the most precise method for including or excluding an individual as the source of the biological evidence at a crime scene was conventional serology, which involves analysis of fluids for certain markers that are lifelong individual characteristics, chiefly based on water-soluble ABO blood group substances and the phosphoglucomutase ("PGM") enzyme genetic marker system. The ABO blood group substances are found on the surface of red blood cells. In addition, water-soluble ABO blood group substances are expressed by about 80% of the population in other body fluids, including saliva, semen, and vaginal fluid;[102] these individuals are called secretors.[103] Secretor status is a genetically determined trait. Analysts test fluids for the presence of the A, B, and H blood group substances using ABO typing, the first means developed for distinguishing individuals based on characteristics of their body fluids.[104] This conventional serology analysis cannot identify particular individuals; it can, however, exclude individuals or place individu-

[102] See 1 Giannelli & Imwinkelried, supra note 78, § 17-8; George F. Sensabaugh, Jan Bashinski & Edward T. Blake, The Laboratory's Role in Investigating Rape, Diagnostic Med., Mar. 1985, at 4.

[103] See 1 Giannelli & Imwinkelried, supra note 78, § 17-8.

[104] See, e.g., Thomas C. Meyers, The Role of the Forensic Serologist, *in* Forensic Science and Law 409, 409 (Cyril H. Wecht & John T. Rago eds., 2006).

als within a percentage of the population that possesses a given type and cannot be excluded as a source of the fluid.[105]

The ABO frequencies were derived from well-established, scientifically valid databases. Based on the analysis of more than 70,000 samples, it was observed that approximately 40% of the Caucasian population are Type A, 11% are Type B, 45% are Type O, and 4% are Type AB.[106] For Blacks, 23% are Type A, 22% are Type B, 51% are Type O, and 4% are Type AB.[107] The most crucial element of any conclusion concerning serology is the relevant population that is included by a finding of blood markers in a crime scene sample. Any testimony that misstates the relevant included population violates the scientific basis for conventional serology.

Serologists in these cases often violated that scientific criterion by misstating the included population in a manner that made their findings appear more probative. Most often they claimed the perpetrator was part of some subset of the population that included the defendant, when in fact no part of the population could be excluded. In other cases, they inaccurately narrowed the subset of the population including the defendant. In still additional cases, the serology excluded the defendant, but analysts argued that the results were non-probative or could somehow nevertheless include the defendant. In each of these examples of invalid testimony, the analyst misstated the statistics regarding the included population to make them seem smaller and therefore more inculpatory than they in fact were.

1. Ignoring the Problem of Masking and Quantification

Most of the DNA exonerations involved sexual assault convictions. During the criminal investigations in most of those cases, a rape kit was prepared, which would include swabs taken by doctors

---

[105] A few courts bar serology results including the defendant, fearing that jurors might misunderstand statistical evidence regarding the population included or deeming such results legally inconclusive. See 1 Giannelli & Imwinkelreid, supra note 78, § 17-9.

[106] Dale D. Dykes, The Use of Frequency Tables in Parentage Testing, *in* Probability of Inclusion in Paternity Testing: A Technical Workshop 15, 20, 29 (Herbert Silver ed., 1982).

[107] Id.

from the victim's body.[108] In addition, law enforcement might preserve other crime scene evidence, such as clothing, on which the assailant may have deposited fluids. Sexual assault cases typically involve mixed stains, in which the victim's own genetic markers may often be present and obscure the genetic markers from the assailant.

While modern DNA techniques allow analysts to isolate and amplify miniscule amounts of semen contained in a mixed stain, conventional serology was not capable of doing so. The proportion of semen in the sample could be so small that any material from the semen would not be detected; this is known as the problem of masking and quantification. The victim's own genetic markers could overwhelm—or "mask"—any genetic markers from the semen, making it impossible to detect the blood antigen type of the assailant absent the ability to quantify the semen content of the sample.[109] As Blake put it, because "[s]emen evidence is normally contaminated with vaginal material from the victim," the interpretation of such evidence "must take into consideration the possible contribution of the victim to the genetic marker pool."[110] This problem was well known in the 1980s, when most of the people later exonerated by DNA testing were convicted.[111] Quantification techniques later made it possible to reach additional conclusions re-

---

[108] Fed. Bureau of Investigation, Handbook of Forensic Sci. 35–36 (1984) [hereinafter 1984 FBI Handbook].

[109] In other words: "Conventional serology is further limited, in that analysis of mixed-fluid stains in which two or more contributors are involved can mask an individual donor." National Research Council, DNA Technology in Forensic Science 158 (1992).

[110] Affidavit of Edward T. Blake, supra note 3, at 15.

[111] The 1984 FBI Handbook explained that in a mixed stain, with both seminal and either urine or vaginal secretions, testing "is more difficult," and further the donor might be a "'weak' secretor" or the "amount of blood group factor present in the semen or saliva" may be "insufficient for reliable grouping tests." 1984 FBI Handbook, supra note 108, at 34; see also Forensic Sci. Research and Training Ctr., Proceedings of a Forensic Science Symposium on the Analysis of Sexual Assault Evidence 61 (1983) (describing the situation in which "the group of the victim masks or hides that of the assailant's semen"); Brian Wraxall & Thomas Fedor, Oklahoma City Police Department Crime Laboratory Serology Audit, Serological Research Institute, Report 4 (2001) (finding that where the analyst observed ABO typing results consistent with the victim, "the only conclusion that she should correctly draw is that the semen donor could be any male in the population. These interpretation rules were well known and established in 1986.").

garding mixed-fluid stains in which no antigens foreign to the victim were found. Such techniques were developed in the mid-1980s, but were not widely adopted by crime laboratories until later. None of the cases in the study set with invalid testimony involved the use of techniques for the quantification of semen.

In a case involving a mixed stain in which no ABO blood group substance or enzymes foreign to the victim are detected, any forensic serologist knows that, absent additional information based on quantification of the semen content of the sample, "no potential semen donor can be excluded."[112] Under such circumstances, the failure to inform the jury that 100% of the male population could be included and that none can be excluded is highly misleading. In David Sutherlin's case, Ronald Enzenauer, of the State of Minnesota, Bureau of Criminal Apprehension, properly explained this phenomenon:

> Q. So that [ABO typing] test—you can't tell anything about the donor because she masks all of those blood groupings?
>
> A. That is correct.
>
> Q. Would there be any blood grouping that that wouldn't mask?
>
> A. No.[113]

The problem of masking and quantification may assist the State to explain why seemingly unfavorable serology evidence is "neutral evidence at worst."[114] Masking and quantification can also be important to the defense, to show why seemingly inculpatory serology evidence is in fact non-probative. In 46 of the invalid serology testimony cases, the analysts described serological results from a mixed stain, yet failed to explain that nothing informative could be said about the potential semen donor population; the serological evidence included 100% of the population. Instead, the analysts testified that the perpetrator was included within some smaller

---

[112] Michael R. Bromwich, Fourth Report of the Independent Investigator for the Houston Police Department Crime Laboratory and Property Room 19 (Jan. 4, 2006), available at http://www.hpdlabinvestigation.org/reports/060104report.pdf.

[113] Trial Transcript at 252, 263, State v. Brian Sutherlin, No. 4181 (Minn. Dist. Ct. June 7, 1985).

[114] Driscoll v. Delo, 71 F.3d 701, 708 (8th Cir. 1995).

percentage of the population that could have produced the semen recovered from the rape victim.

State of California v. Herman Atkins

One example is the Herman Atkins case. There, the victim was a Type A secretor, PGM type 2+1+, as was Atkins. Similarly, the vaginal swab test results disclosed Type A secretor, PGM type 2+1+ material, which was consistent with both the victim and Atkins. James Hall, of the Riverside Crime Laboratory, California Department of Justice, testified as follows:

> Q. Based on the information that you received, what kind of conclusions could you tell me about the swab and the blood type of the two persons?
>
> A. Well, the blood type of the swab was the same blood type of the two individuals. That tells me that possibly the semen . . . could be of that blood type combination, or the activity that I detected could have come from the victim herself, or it could have come from a combination of individuals with those blood types.
>
> Q. Do your results exclude the person that you identified as Herman Atkins' blood?
>
> A. No, they don't.
>
> Q. Now, are there certain percentages of the population that have this ABO typing and the PM—what you discovered?
>
> A. PGM.
>
> Q. PGM that you discussed?
>
> A. Yes, there are.[115]

Hall then stepped down from the stand and made a chart in front of the jury. He wrote the figure 4.4% next to the word "black," and this testimony followed:

> Q. It would be 4.4% of the black population?
>
> A. That's right.[116]

---

[115] Trial Transcript at 233–34, People v. Herman Atkins, No. 28832 (Cal. Ct. App. Aug. 22–25, 1988).

[116] Id. at 234–36.

In the Atkins case, the serologist found nothing foreign to the victim and he made no assessment of the quantity of semen in the mixed body fluid sample. Rather than testifying that no conclusion could be drawn about the genetic profile of the semen donor, Hall presented a statistic: 4.4% of the black population are Type A secretors, PGM type 2+1+, thus excluding more than 95% of the population and including Mr. Atkins as a potential semen donor. One might argue in defense of Mr. Hall that all he did was accurately answer the prosecutor's irrelevant question. But since the analyst knew that the *only* frequency relevant in a sexual assault case is the combined frequency of the potential semen donors, by providing an irrelevant frequency in response to the prosecutor, the testimony misled the jury. National symposia on serology left no scientific doubt about these principles.[117]

Again, the only probative frequency statistic that is considered by a forensic scientist investigating a sexual assault is the combined frequency of all possible genotypes potentially possessed by the semen source. In those cases where all of the traits detected in the sample could originate from the female body fluids and there is no assessment of the semen dilution, the potential types for the semen source encompass the entire population and no one can be eliminated as a potential semen source.

State of Texas v. Kevin James Byrd

Another example is the case of Kevin James Byrd, in which James Bolding of the City of Houston Police Department Crime Laboratory found no antigens in an examination of a vaginal swab from the rape kit. The victim was a Type B non-secretor, and Byrd was a Type A non-secretor.[118] Bolding testified as follows:

> A. My conclusion would be that the individuals present or the samples present contained a non-secretor status, that is, we could not determine whether or not they had type A, B, or O blood factor.

---

[117] See supra note 111.

[118] Trial Transcript at 164–69, State v. Kevin Byrd, No. 424823 (Minn. Dist. Ct. Aug. 7, 1985).

> Q. Does that mean the victim in the case would have been a non-secretor?
>
> A. That would be the conclusion we would draw.
>
> Q. What would the conclusion you would draw be regarding the suspect or the attacker in this case?
>
> A. That would also mean that the semen donor would also be a non-secretor.[119]

Bolding testified that 15–20% of the population are non-secretors, and that the defendant was a non-secretor.[120] However, no quantification was conducted to assess the degree to which semen was present in the sample. As a result, the failure to observe any ABO blood group substances could be due to an inadequate concentration of semen in the extract employed to conduct the ABO typing. The proper statistic, therefore, was that 100% of the population could have been the source of the semen because there was no information to prove that the quantity of semen was adequate to detect blood group substances, even if the semen contained them.

Because the Type O secretor, in which the individual secretes only the H antigen, is the most common ABO type, many of the masking cases involved the common situation in which the victim and the defendant were both Type O secretors and the materials from the rape kit exhibited just the H antigen. In all such cases, absent quantification, 100% of the population could have been the semen donor. Yet in a series of cases, the examiner testified that the defendant, a Type O secretor, was included in some subset of the population that could have committed the crime.

Commonwealth of Massachusetts v. Neil Miller *and*
Commonwealth of Massachusetts v. Marvin Mitchell

In the Neil Miller case, David Brody of the Boston Police Department testified that "[t]he H blood group substance that I found had to be deposited by a Group O individual, a Group O secretor"; additionally, he stated that "[a]pproximately forty-five percent of

---

[119] Id. at 164–65.
[120] Id. at 165–66.

the population are Group O individuals."[121] Brody's testimony is particularly disturbing because in the Marvin Mitchell trial—also included here—he understood well the problem of masking and quantification. At Mitchell's trial, Brody testified regarding the phenomenon where it assisted the prosecution in attempting to explain why Mitchell, a Type A secretor, was not excluded by a test of the vaginal swab containing H antigens only, which was consistent with the victim, a Type O secretor. There Brody testified:

> A. Mr. Mitchell could not be excluded. *No secretor could be excluded from depositing that stain because the stain may have been too diluted or graded* [sic] *to pick up Mr. Mitchell's blood type.* So I cannot exclude him, but I cannot say that I found the A blood group type. In other words, again no secretor can be totally excluded from the stain.[122]

Even in the Mitchell case, Brody did not fully explain the phenomenon of masking; he erroneously implied that as a secretor, Marvin Mitchell was part of some subset of the population that could have been the donor. He never informed the jury that no individual, whether a secretor or non-secretor, could be excluded.

2. Invalid Analysis of Presumptively Exculpatory Serology

In a series of cases, traditional serology evidence strongly supported the defendant's innocence, but forensic analysts engaged in a series of unsupported and misleading speculations on the stand attempting to discount the evidence of innocence. Sometimes the testimony converted exculpatory serology findings into false inculpatory evidence. These cases typically involved analysts telling the jury that antigens can "selectively degrade" due to bacterial contamination, thus disregarding a result that excluded the defendant. The Supreme Court recently decided a case involving invalid testimony of this sort by an FBI analyst.[123]

---

[121] Transcript, N. Miller, supra note 64, at 198.

[122] Trial Transcript at 51–52, Commonwealth v. Marvin Mitchell, No. 072574 (Mass. Dist. Ct. Jan. 18, 1990) (emphasis added).

[123] See House v. Bell, 126 S. Ct. 2064, 2071–72 (2006); Brief for the Innocence Project, Inc. as Amicus Curiae in Supporting Petitioner at 25, House v. Bell, 126 S. Ct. 2064 (2006) (No. 04-8990) (noting that "if antigens could selectively 'vanish,'" and

State of Montana v. Paul Demetri Kordonowy

In the case of Paul Kordonowy, Julie Long of the Montana State Forensic Laboratory performed the testing on the victim's underpants that revealed A antigens, which neither the victim nor Kordonowy possessed; both were Type O secretors.[124] Rather than testify that the testing excluded Kordonowy, Long testified as follows: "[I]n this case there was a large amount of bacteria, which I noted, and it has been reported that a large amount of bacteria can give you an A substance reading in your analysis because your ABO substances are sugars, and bacteria also produce sugars."[125] As Edward Blake concluded in his report examining the case, this "bacteria" testimony lacks any scientific foundation:

> [T]here is no evidence whatsoever that bacteria produce water soluble ABO antigens of any sort much less ABO antigens of type "A." If this assertion were true, the ABO typing of sexual assault evidence would be inherently unreliable because no scientist could ever know whether or not the ABO antigens detected in vaginal or oral samples were from ubiquitous bacteria or the human being from whom the sample was collected or some other human being contributing a body fluid to the sample. Like the claim that bacteria preferentially destroy spermatozoa, the claim that bacteria preferentially secrete ABO "A" antigens is without scientific basis; and, if true, would undermine the entire scientific foundation for the ABO typing of body fluid evidence.[126]

Indeed, Long not only stated that based on her analysis, Kordonowy could not be excluded,[127] but went further by affirmatively including Kordonowy. Long stated that Kordonowy fell within the 29% of the population who are Type O secretors and could be the

---

blood types could mutate from one to another, "then serology would *never* have been a reliable method"); Garrett, supra note 55, at 1638.

[124] Trial Transcript at 355, 369–70, State v. Paul Demetri Kordonowy, No. DC-89-013 (Mont. Dist. Ct. Jan. 16–18, 1990) [hereinafter Transcript, Kordonowy].

[125] Id. at 371.

[126] See Lori E. Hutchinson, Report on State v. Paul Kordonowy 3–4 (2003).

[127] Transcript, Kordonowy, supra note 124, at 385.

donor.[128] That testimony was also invalid. Putting aside her failure to give correct attribution to the A antigen, Long failed to recognize that the H blood group substance was not foreign to the victim and hence cannot be used to limit the population of semen donors.

State of Illinois v. Gary E. Dotson

The case of Gary Dotson, the first person exonerated by post-conviction DNA testing, also included the same type of invalid testimony—in addition to the invalid testimony already described that ignored the problem of masking and quantification. Timothy Dixon of the Illinois Department of Law Enforcement told the jury not to reach any conclusions based on the Type A antigens found in stains in several places on the victim's clothes, antigens that could not have come from the victim or Dotson, who were both Type B. Dixon testified:

> The A stain—I can't say the A stain, I can't say that blood is A, I can't say that blood is B, all I can say is that material was blood, and a mixture of—it could be perspiration, could be other body fluids in combination of B and H activity.[129]

He added: "Unfortunately for us there are lots of materials; dust, wood, leather, certain kinds of clothes, different cloth materials, detergents in materials" that could somehow explain the presence of the Type A antigens.[130]

As Blake explained in his report, control testing could be used to test the area outside the stain to assess whether the result was due to contamination. If it were actually the case that contamination could never be ruled out, then "ABO typing of biological samples" would have always been an "inherently unreliable" type of analysis.[131]

---

[128] Id. at 386. Long agreed with the statement that "we cannot rule out Mr. Kordonowy because of the presence of the H Substance," and then added that 29% of the population are Type O secretors. Id. at 383, 385–86.

[129] Trial Transcript at 359, State of Illinois v. Gary E. Dotson, No. P.C. 4333 (Ill. Cir. Ct. July 25, 1985).

[130] Id. at 368.

[131] Affidavit of Edward T. Blake, supra note 3, at 23.

State of West Virginia v. Gerald Davis *and*
State of West Virginia v. Dewey Davis

Similarly, during the trials of Gerald and Dewey Davis, Fred Zain of the West Virginia Department of Public Safety, Criminal Investigation Bureau, gave misleading testimony. In Gerald Davis's trial, Zain explained how bacteria could account for the presence of Type A antigens where the victim and Gerald Davis were both Type O secretors, stating, "Bacterial contamination can give you what is called false positives and give you blood types separate and aside from what you're truly identifying."[132] Where the testing excluded Gerald Davis, Zain instead claimed that by ignoring the Type A finding, one should conclude that Davis was included in the group of "around the realm of 7 percent of the general population of West Virginia" who could have committed the crime.[133]

The American Society of Crime Laboratory Directors Laboratory Accreditation Board ("ASCLD/LAB") Investigation Report of Zain's work conducted in 1993 concluded that finding "an ABO type foreign to both the victim and defendant . . . would normally be interpreted as excluding defendant as the semen donor," and that there was "no satisfactory foundation" for the conclusion in the Gerald and Dewey Davis trials that resulted in the "ABO mismatch" being "dismissed as bacterial contamination."[134]

---

[132] Trial Transcript at 259, State v. Gerald Davis, No. 86-F-152 (W. Va. Cir. Ct. July 21, 1986) [hereinafter Transcript, G. Davis]. Or as Zain put it in Dewey Davis's trial: "[A]nytime a body fluid leaves an individual's body, there is an automatic contamination of whatever the body fluid might be up to and including the time that that stain becomes dried." Trial Transcript at 238, State v. Dewey Davis, No. 86-F-153 (W. Va. Cir. Ct. Mar. 9, 1987). Zain added:

> You have foods that once you open a can—I'm sure most of the ladies on the jury have done some canning at one time or another. Once you open that up, you have to either keep it in a refrigerator or you have to have it in sort of a preservative to keep bacterial activity from happening. . . . Blood and body fluids are the same thing.

Id. at 239.

[133] Transcript, G. Davis, supra note 132, at 249.

[134] See James J. McNamara, Ronald R. Linhart, ASCLD/LAB Investigation Report, West Virginia State Police Crime Laboratory, Serology Division 8 (1993); see also In re W. Va. Crime Lab, Serology Div, 190 W.Va. 321 (1993).

State of Indiana v. Jerry Watkins

Several other similar examples are included in the data set. In the case of Jerry Watkins, the victim was a Type A secretor, Watkins was a Type O secretor, and yet the swabs displayed the A, B, and H blood group substances.[135] Forensic analyst Carol Kohlman of the Indianapolis and Marion County Forensic Services Agency was asked, "Is there anything from your results that would allow you to exclude the possibility of Jerry Watkins being the semen donor?" Despite the presence of the B blood group substance, which was foreign to both the victim and Watkins, she answered, "No sir." She gave another similar explanation: "I was suspecting bacterial contamination as part of the problem or as a possible explanation . . . ."[136] The defense attorney did question Kohlman regarding this surprising opinion on cross-examination, asking whether bacteria of such a kind were observed, asking, "Did you do any cultures?" She responded, "No sir, we do not do cultures in our laboratory."[137] No effort was made to support the bald conjecture, nor was the valid result—that the defendant should have been excluded—ever presented.

3. Additional Invalid Use of Population Statistics

In addition to the large set of cases involving invalid testimony that ignores the problem of masking and quantification, several other examiners misstated or miscalculated population statistics. In a series of cases, serologists divided the relevant statistic in half, claiming that men constitute half of the population and only a male could have deposited semen at the scene. It is logically incorrect to divide a frequency in half when it is understood at the beginning that the relevant population is males, because only males produce semen. The population statistics regarding these blood group substances are identical for both sexes. For example, approximately 40% of both Caucasian men and women are ABO Type A. It is incorrect to divide that figure in half and claim that only 20% of men are Type A.

---

[135] Trial Transcript at 961, 977, 988–89, State v. Jerry Watkins, No. 30D01-8603-CF-005 (Ind. Super. Ct.) (date unknown).

[136] Id. at 999.

[137] Id. at 1025.

In the Mark Bravo case, Richard W. Catalani of the Los Angeles County Sheriff's Crime Lab stated that although 3% of the population was PGM type 2–1+, that figure should be divided in half to eliminate females. That testimony provided the jury with the incorrect figure that 1.5% of the male population could have been the semen donor.[138] In the Perry Mitchell case, John Barron of the South Carolina Law Enforcement Division, noting that H antigens were observed and that 35% of the population are Type O secretors, testified, "You would probably have to also cut that by another 50% because we're dealing with males." When the defense counsel questioned why population statistics are not the same for men and women, Barron responded, "[T]here is a difference in regard to semen."[139] Citing the same false rationale, Fred Zain similarly divided statistics in half in the Dewey Davis and Glen Woodall cases, as did Julie Long in the Chester Bauer case.

To summarize, in each of these examples of invalid forensic science testimony, the analyst failed to accurately provide the relevant statistic regarding the included population. These analysts instead offered invalid, reduced frequencies (a rarer event) that appeared to further inculpate the defendant.

B. Invalid Microscopic Hair Comparison Testimony

Sixty-five of the trials examined involved microscopic hair comparison analysis. Of those, 25—or 38%—had invalid hair comparison testimony. Most (18) of these cases involved invalid individualizing claims.

Forensic hair microscopy involves the side-by-side comparison under a microscope of head and pubic hairs found at a crime scene with dozens of head and pubic hairs plucked and combed from the scalp and pubis of the victim and suspect(s).[140] Hair examination has long been an important part of police investigations, because

---

[138] Trial Transcript at 267–68, People v. Mark Diaz Bravo, No. Va. 003313 (Cal. Super. Ct. Sept. 27, 1990).

[139] Trial Transcript at 225, State v. Perry Mitchell, 83-GS-32-479 (S.C. Cir. Ct. Jan. 19, 1984).

[140] Richard E. Bisbing, The Forensic Identification and Association of Human Hair, *in* 1 Forensic Science Handbook 390, 414–21 (Richard Saferstein ed., 2002).

hairs are so commonly and readily transferred to skin or clothing.[141] A head hair found on the sheets of a rape victim is a "questioned" hair; the numerous hairs plucked and combed from the victim and suspect are "reference exemplars."

Forensic hair evidence has increasingly been scrutinized due to studies indicating high error rates.[142] Examiners commonly distinguish human from animal hair, opine on the racial group from which the questioned hairs originated, determine from which part of the body the hair originated, and then testify that the hairs have microscopic characteristics that are similar or dissimilar to the exemplar samples.[143] This study is not concerned with the analyst's ability to distinguish hair of different species, races, or parts of the body. This study is concerned with testimony which attempts to declare a positive association between a questioned hair from a crime scene with a set of exemplars from a suspect.

Not only was forensic human hair comparison frequently relied upon in criminal cases at the time relevant to this study because hairs are easily left at a crime scene, but also because there is considerable variation in the microscopic characteristics of hairs coming from different people. There exists significant intrapersonal variation among the hairs from a single individual's head, and as a result, the competent examiner will collect perhaps 100 hairs taken from different regions of the scalp and then select a representative subset of perhaps 20 hairs to be compared with the questioned head hair. There are many different general or "class" characteristics of hair that can be microscopically examined and compared along the length of the hair. Some of the general characteristics are color, structure, and cuticular traits. Although no single class characteristic is very unusual, 20 representative hairs possessing a range

---

[141] Edward J. Imwinkelried, Forensic Hair Analysis: The Case Against the Underemployment of Scientific Evidence, 39 Wash. & Lee L. Rev. 41, 42–43 (1982).

[142] Id. at 44–45 (describing a Law Enforcement Assistance Administration study); D. Michael Risinger & Michael J. Saks, Rationality, Research and Leviathan: Law Enforcement-Sponsored Research and the Criminal Process, 2003 Mich. St. L. Rev. 1023, 1049–50 (deriving an 11.5% error rate for visual hair comparison from FBI study data); see also Clive Stafford Smith & Patrick D. Goodman, Forensic Hair Comparison Analysis: Nineteenth Century Science or Twentieth Century Snake Oil?, 27 Colum. Hum. Rts. L. Rev. 227, 242–45 (1996).

[143] Bisbing, supra note 140, at 418–19.

of characteristics that are similar to the questioned hair would be forensically significant.[144]

As the FBI noted in its 1984 handbook, microscopic hair examination is "[n]ot positive evidence."[145] The National Research Council has explained:

> Although hair examiners can associate a hair with racial characteristics and body source (trunk, head, or pubic area) the variations among hairs on a given person make definitive association of a single hair with an individual problematic. The microscopic comparison of hairs is also subjective and can lead to differences of opinion among equally qualified experts.[146]

Where even qualified examiners may disagree about whether hairs could come from the same source, hair evidence cannot be individualized based on microscopic analysis.[147] Consequently, the field adopted standards that the strongest statement of association that can be made is a statement that hairs in question are "consistent" with the defendant's or "could have" come from the defendant. The 1985 International Symposium on Forensic Hair Comparisons ("1985 Symposium") adopted these standards.[148]

Thus, hair examiners following those standards may testify that two samples are visually or microscopically similar. Though "valid" in the sense used here—because such testimony does not misstate empirical data or incorrectly claim empirical support—such a conclusion is not highly probative. As Richard Bisbing has put it, such

---

[144] Moreover, in some cases, there may be an artificial treatment which can be of value. For example, if a suspect had brown hair, dyed it green, and then let it grow out, and the length that was brown versus the length that was green would be approximately the same between the questioned hair and the exemplars; the similarity would be significant. Id. at 410.

[145] 1984 FBI Handbook, supra note 108, at 37; see Bisbing, supra note 140, at 419 ("[O]ne can never say with absolute certainty that a particular hair originated from one individual to the exclusion of all others.").

[146] Comm. on DNA Tech. in Forensic Sci. Nat'l Research Council, DNA Technology in Forensic Science 158 (1992).

[147] This is an important difference between serology and hair microscopy. In the former, it is expected that all competent analysts will agree on whether two samples have the same blood group substances.

[148] See FBI, Lab. Div., Proceedings of the International Symposium on Forensic Hair Comparisons 108 (1985) [hereinafter 1985 Symposium]; discussion infra Subsection II.B.1.

evidence is "by necessity, equivocal, that is 'could be evidence.'"[149] Evidentiary questions that are not addressed here are raised by such testimony. Courts should ask whether a finding that hairs are "similar" or "consistent" has sufficient probative value to be admissible, or causes undue prejudice due to the ambiguity of the terms "similar" and "consistent." Important questions remain whether such subjective analysis and such imprecise conclusions would today satisfy *Daubert* requirements for admissibility of expert testimony.[150] The NAS report was emphatic that "[n]o scientifically accepted statistics exist about the frequency with which particular characteristics of hair are distributed in the population."[151] The report added that, "[t]here appear to be no uniform standards on the numbers of features on which hairs must agree before an examiner may declare a 'match.'"[152] Linking these defects with the problem of invalid terminology used in reports and testimony, the NAS report explained that "[t]his illustrates not only the imprecision of microscopic hair analyses, but also the problem with using imprecise reporting terminology such as 'associated with,' which is not clearly defined and which can be misunderstood to imply individualization."[153] A range of other forensic disciplines lacking empirical data—such as bite mark comparison and handwriting comparison—raise these questions and also may not survive proper *Daubert* scrutiny.[154] However, for the purposes of this study, which does not reach such questions, testimony solely observing a similarity, while imprecise and potentially misleading, is deemed valid, because it does not render a conclusion contrary to

[149] Richard E. Bisbing, Forensic Hair Comparisons: Guidelines, Standards, Protocols, Quality Assurance and Enforcement, Presentation to NAS, April 24, 2007, http://www7.nationalacademies.org/stl/April%20Forensic%20Bisbing.pdf.

[150] See, e.g., 2 Giannelli & Imwinkelried, supra note 78, § 24-3; Imwinkelried, supra note 141; Smith & Goodman, supra note 142, at 283–90. But see Houck et al., The Science of Forensic Hair Comparisons and the Admissibility of Hair Comparison Evidence: Frye and Daubert Considered, Mar. 2004, http://www.modernmicroscopy.com/main.asp?article=36,12Mar2004.

[151] See Strengthening Forensic Science, supra note 12, at 5-25.

[152] Id.

[153] Id.

[154] See, e.g., Saks & Koehler, supra note 50, at 218.

underlying empirical data or claiming support based on nonexistent data.[155]

DNA testing of the mitochondria, or when the hair roots are present, of the nucleus, has now supplanted microscopic hair comparison in many cases. In six exonerees' cases, for example, the analyst identified hairs as consistent with the defendant at trial, but mitochondrial or other DNA analysis later determined that those same hairs originated from a person other than the convicted defendant.[156]

1. Invalid Probability Testimony

Where an examiner can only opine whether hairs are similar or consistent, forensic hair comparison experts also advise that an examiner should not make "any statements about the probability that a specific hair could have come from someone other than the person to which it was associated."[157] The 1985 Symposium noted the possibility of a "coincidental match," and therefore cautioned that there was a need for "further research" on frequency before probability statements can be used when describing a hair compari-

---

[155] See id. at 216 (recommending that until empirical research is done to support certain forensic disciplines, for the present, "criminalists should report [a] finding with the appropriate clarity and restraint. For example, they could explain that a conclusion that two patterns are consistent (or a match) does not require a conclusion that the patterns share a common source" and noting that "there is no scientific justification" for speculation regarding the likelihood that the patterns share a common source).

[156] The cases are those of Richard Alexander, William Dedge, George Rodriguez, Drew Whitley, Stephen Avery, and Anthony Hicks. See Innocence Project, http://www.innocenceproject.org/know/Browse-Profiles.php.

[157] See Max M. Houck et al., The Science of Forensic Hair Comparisons and the Admissibility of Hair Comparison Evidence: Frye and Daubert Considered, Mod. Microscopy J. 5 (Mar. 2, 2004), available at http://www.modernmicroscopy.com/main.asp?article=36&page=5&searchkeys=Houck ("All of these numbers notwithstanding, to attempt to derive a population frequency of traits or to determine how likely it may be to encounter a given hair in a given population is fraught with complexity. Most experts . . . do not feel comfortable with any statements about the probability that a specific hair could have come from someone other than the person to which it was associated. The authors agree with that approach. The justification for that reluctance is based on the complexity of the probability question, difficulty choosing a population to which to assign the probability, the lack of sufficient data where that question was addressed, and court decisions excluding such statements of probability in the past." (citations omitted)).

son.[158] No such systematic efforts to research the frequency with which particular microscopic features occur in any population have been conducted. Thus, there is not and never has been any statistical basis for hair comparison.[159]

After all, examiners lack empirical data on the frequency of any of the general or "class" characteristics, and lack data as to the extent to which there is a correlation between different class characteristics. Without frequencies (in contrast to the ABO blood group system), hair examiners are also precluded from expressing to the jury a probability that hairs recovered at the crime scene came from the defendant. A probability can ordinarily be determined in one of two ways: theoretically or empirically. A theoretical probability requires a well understood situation under demonstrable assumptions. If one rolls a six-sided die, assuming that each of the six sides is equally likely to show up on top, the theoretical probability of any particular side showing up in a single roll of the die is 1/6. Given what is known about hair, hair analysis itself does not lend itself to theoretical probabilities. Empirical probabilities, on the other hand, are gained from a large set of data expressed as: total number of times the outcome occurred divided by total number of instances examined. Since there are no adequate sets of data for the occurrence of general hair characteristics, the examiner cannot present an empirical probability.

---

[158] 1985 Symposium, supra note 148, at 110; see also James Robertson & Colin Graeme Girdwood Aitken, The Value of Microscopic Features in the Examination of Human Head Hairs: Analysis of Comments Contained in Questionnaire Returns, 31 J. Forensic Sci. 563, 568 (1986) ("There is a clear need for an extensive research program to evaluate the microscopic features of hair from a forensic science standpoint . . . ."). The "general opinion" among experts in the mid-1980s, for example, was that "Gaudette's study is only an initial step toward determining the significance of hair comparison evidence," and that "[f]orensic experts, including those employed by the FBI, still recommend that hair examiners use cautious statements when asked to give an opinion on whether the origin of an unknown hair and of a representative sample is the same." Dalva Moellenberg, Splitting Hairs in Criminal Trials: Admissibility of Hair Comparison Probability Estimates, 1984 Ariz. St. L.J. 521, 536.

[159] See Richard E. Bisbing et al., Peer Review Report: Montana v. Jimmy Ray Bromgard 2 [hereinafter Peer Review Report] ("[T]here is not—and never was—a well established probability theory for hair comparison.").

State of Montana v. Jimmy Ray Bromgard

Nevertheless, several analysts in these exonerees' cases violated that criterion and bolstered their testimony by making invalid statements regarding supposed probabilities. Arnold Melnikoff, Director of the Montana State Crime Laboratory, testified regarding probabilities in a series of cases. Though there is not, and never has been, any statistical basis for hair comparison, he would simply fabricate frequencies and probabilities based on his own undocumented estimates, rather than any reliable empirical data. He would then go even farther and multiply his made up probabilities by the number of hairs "matched" from different parts of the body, as if each represented independent events. In the case of Jimmy Ray Bromgard, Melnikoff testified that "[t]he hair from the blanket on the left matches all the characteristics of the known pubic hair from James Bromgard on the right, and they almost look like one hair."[160] He then explained the probability of such a "match":

> Well there are actually two mutually exclusive events because they come from different areas of the body, and their characteristics are not necessarily the same. So if you find both head and pubic hair there you have one chance in a hundred for the head hair matching a particular individual and one chance in a hundred for the pubic hair. If you find both it's a multiplying effect, it would be one chance in 10,000, it's the same as two dice, if you throw one dice with a one, one chance out of six; if you throw another dice with a one, it's one chance out of six, you multiply the odds together. You do the same in this case so, it's one times one hundred, times one, times one hundred, and you get one in 10,000.[161]

State of Montana v. Chester Bauer

Arnold Melnikoff testified similarly in Chester Bauer's case, stating:

---

[160] Trial Transcript at 231, State v. Jimmy Ray Bromgard, No. 88108 (Minn. Dist. Ct. Nov. 16, 1987) [hereinafter Transcript, Bromgard].

[161] Id. at 237–38.

> To have them both match, it would be the multiplication of both factors so as an approximately [sic] using that 1 out of 100, you come out with a number like 1 chance in 10,000. Multiply 100 x 100. It becomes a very highly improbable situation that both events would occur, that you could not distinguish the head hair and the pubic hair from two individuals.[162]

And he testified the same way in Paul Kordonowy's case, again claiming that hairs from different parts of the body are "independent events" to which he attached numerical probabilities which he then multiplied.[163] Not only did Mr. Melnikoff not have any data to support the number 1/100, he also had no information to conclude that the color and class characteristics of head hair and pubic hair are independent so that their frequencies can be multiplied.

The probabilities of two events can only be multiplied if the events are statistically independent; that is to say that the outcome of one event does not influence the outcome of the other event. Melnikoff multiplied the two probabilities without it first being scientifically proven that the events are statistically independent. There is no published research on the question of statistical independence for head and pubic hair. Indeed, on the contrary, a person with dark-colored scalp hair may be more likely than a person chosen at random to have dark colored pubic hair.

State of Oklahoma v. Timothy Edward Durham

In Timothy Durham's case, Carol English Cox of the Tulsa Police Laboratory opined that the particular reddish-yellow hue of his hair and the questioned hair were only found in "about 5 percent of the population."[164] Cox did not provide scientific support for that statistic, nor could she do so.

---

[162] Trial Transcript at 250, State v. Chester Bauer, No. 83-CR-27 (Mont. Dist. Ct. July 16, 1983).

[163] Id. at 251.

[164] Trial Transcript at 385, State v. Timothy Edward Durham, No. CF-91-4922 (Okla. Dist. Ct. Mar. 9, 1993).

2. Exaggeration of Probative Value of Hair Association Without Using Numbers

The 1985 Symposium established standards regarding conclusions that analysts may reach regarding association of questioned and exemplar hairs. Statements of association may consist of conclusions that a hair "could have" come from an individual or "is consistent" with an individual's hair, or "could not have" come from an individual and "is not consistent" with an individual's hair; "is consistent" is the strongest statement of association permitted.[165] Many of the experts testifying in the trials studied here described and adhered to those standards with care. Nevertheless, these trials are also replete with examples in which analysts expressed far greater certainty that hairs came from a particular defendant. Testimony used to convey strong association ranged from "more likely than not" that questioned hairs came from the defendant, to naming a "quite large" probability that the questioned hairs came from the defendant, to stating that hairs in fact did come from the defendant.[166] Each of these terms implies a known probability, which, as explained above, does not—and, in the absence of empirical data, cannot—exist. Such testimony or analysis violates the scientific criterion that statements of frequency or probability must be supported by empirical population data.

---

[165] See 1985 Symposium, supra note 144, at 108–09.

[166] Some examiners in the cases studied declared a "match" between the questioned and known hairs. On its face, the word "match" may not seem invalid because it need not imply individualization. For instance, the co-authors "match" in appearance at the most basic level—we each have two arms and two legs. It all depends upon what is meant by "match." In criminal jury trials, "match" was commonly used to mean individualization. The most frequent use of "match" is when an analyst opines that a crime scene fingerprint "matches" the suspect's; there the examiner is attempting to communicate uniqueness. For that reason, many hair examiners would shy away from using the word "match" with reference to hair, particularly since there is rarely a one-to-one correspondence between the questioned hair and a single known hair. Rather, in most cases of positive hair association, the characteristics exhibited in the questioned hair fit within the *range* of characteristics reflected in the set of hair exemplars. The word "match" misleadingly implies to jurors that the analyst observed a perfect and complete consistency between only two hairs. Yet the 1985 Symposium did not take a clear position on "match" and some analysts still use the word. Thus, if use of "match" was the only transgression in an expert's testimony, this study did not conclude the case involved invalid testimony. In the cases deemed invalid, the examiners went further to amplify the probative value of the evidence.

State of Illinois v. Willie L. Rainge, Kenneth E. Adams, and Dennis Williams

One example of an expression of an invalid degree of non-numerical associative certainty is in the "Ford Heights Four" case. In the trial of Kenneth Adams, William Rainge, and Dennis Williams, Michael Podlecki of the Illinois Department of Law Enforcement Crime Lab testified, while displaying side-by-side photomicrograph comparisons,[167] regarding two hairs found in the trunk of Dennis Williams's car: "I couldn't distinguish if I was looking almost at two hairs. They looked just like one."[168] He added:

> What I saw was a hair like this where I could see all the pieces like a puzzle. Like in the previous hair. A line would just fit in. In other words it wasn't a type of situation of sitting down and looking for it. Just like if you dropped two dollar bills and you see two dollar bills on the floor. You see two one dollar bills. It's obvious. And that's how it looked there.[169]

Yet when asked to state his ultimate conclusions regarding the exhibits containing the hair evidence, Podlecki opined not that the hairs were identical, but that they were "similar in color and characteristics."[170] Where the evidence supported only a conclusion that questioned hairs exhibited a range of characteristics found in the exemplar hairs, it was quite misleading to describe the hairs as identical, and then use a further misleading visual display to convey identity to the jury.

---

[167] Using such visual displays was itself considered improper because they could easily mislead the jury. See 1985 Symposium, supra note 144, at 112 ("The Subcommittee is strongly opposed to showing side by side photomicrographs .... It can be highly inflammatory to a jury.").

[168] Trial Transcript at 2223, People v. Willie L. Rainge, No. 78-I6-5186 (Ill. Cir. Ct. Oct. 10, 1978) [hereinafter Transcript, Rainge].

[169] Id. at 2226.

[170] Id. at 2227. Podlecki had earlier explained that he had received sets of exemplar hairs, and that in order to conclude that hairs were "similar," he would have to find a less than 0.1% difference between them. Id. at 2208, 2211–12.

State of Oklahoma v. Calvin Scott

Another example is from Calvin Scott's case, in which Claud Berry of the Oklahoma State Bureau of Investigation testified as follows:

> Q. Do you know whether or not, Mr. Berry, there have ever been any studies done as to the probabilities of finding another person with hair like ours, or—
>
> A. Well there is one gentleman out of Canada, his name is B.D. Goday [sic], he made a study. He's the only one that has made a study that's been published, and he has found that head hair, one person in forty-five hundred would have a chance of—in other words, identification of one hair to—I mean, one person in forty-five hundred may have features of hair comparison in head hair. Now one in eight hundred in pubic hairs. That's his results. That's the only one I have been able to find who has ever come up with any results with figures. Others have made statements on theory, but they haven't made any practice, or made any study.
>
> Q. Would he have given, or would there be any number type odds to the probability of the hair found on May Ann Fulsom's bottom sheet and the hair, unknown hair found in her pubic combings, both belonging to anyone other than the defendant, Calvin Scott?
>
> A. His hair, I would say this: his studies were made on caucasian hair, I believe. In this case having two hairs identified, two hairs of different kind, I mean, head hair from one person would be quite large, I would say, I would not give a figure. It would be quite large.[171]

Going beyond expressing a high likelihood of association, Oklahoma City Police Department analyst Joyce Gilchrist expressed a definitive association in the Curtis McCarty case. Gilchrist concluded her direct examination at McCarty's first trial by stating her opinion "[t]hat he was in fact there."[172] Similarly, in the Larry Pe-

---

[171] Trial Transcript at 47–48, State v. Calvin Lee Scott, No. CRF 83-74 (Okla. Dist. Ct. Sept. 19, 1983).

[172] Trial Transcript at 177, State v. Curtis Edward McCarty, No. CRF-85-02637 (Okla. Dist. Ct. Mar. 24, 1986); see also infra Section II.F.

terson case, Gail Tighe of the New Jersey State Police Laboratory agreed that all of the questioned hairs examined were identified as either "belonging" to the victim or to Peterson.[173]

Some examiners did not provide numerical statements regarding probability, but made other efforts to describe the probability of finding comparable hairs, or to describe a particular characteristic as "rare" or "uncommon" without the requisite reliable database from which to draw such inferences.

Commonwealth of Kentucky v. William Gregory

Analysts made conclusions regarding probability based on claims that the hairs had supposedly unusual or unique features. Such claims are unsupported where empirical data regarding the frequency of microscopic features in hair is lacking. During William Gregory's trial, Dawn Katz of the Kentucky State Police Jefferson Regional Crime Lab testified that the hairs "more than likely" belonged to Gregory.[174] In part, this was based on a finding of "ovoid bodies" in the hairs, which she called "kind of an unusual characteristic."[175] Katz explained:

> A. I told you, there is no statistics [sic] on this. I can tell you this is the first time I have ever had a negroid origin hair that has not had a medulla in it.
>
> Q. What percentage of people have ovoid bodies in them?
>
> A. This is probably the first time I have ever seen an ovoid body in a human hair. I have seen them in cattle hair before.[176]

Katz added:

> I mean, you have to compare hairs from brothers and sisters that have the same genetic background and carried a lot of the same genetic characteristics from the same parents, you might run into very similar characteristics in two people. But, in general, you

[173] Trial Transcript at 152, State v. Larry L. Peterson, A-3034-89T4 (N.J. Super. Ct. Mar. 6, 1989).

[174] Trial Transcript at 221, 246, Commonwealth v. William Gregory (Ky. Cir. Ct. Aug. 1993).

[175] Id. at 233.

[176] Id. at 251.

> wouldn't see that kind of an overlap in two people you would just pick off the street.[177]

This testimony was invalid. In addition to the invalid use of probability—claiming that the hair "more than likely" originated from Gregory—Katz testified that a characteristic was "unusual" based on no extant population data. Indeed, she admitted that "there is no statistics [sic] on this." Katz further embellished that she had never seen such an "unusual" feature before except in "cattle hair." Finally, Katz implied that only siblings would share such characteristics, again without any data to support such a claim.

The testimony in the Gregory case not only included invalid statements concerning probability, but the analyst made additional claims about particular characteristics based on undocumented experience. Several other examiners buttressed their conclusions by claiming that never in their careers had they seen either special characteristics or more generally, exemplars from any two different people that "matched." Michael Blair's case involved similar testimony by analyst Charles Linch, then a consultant and formerly of the Southwestern Institute of Forensic Sciences in Dallas.[178] He testified that he had "never seen a Caucasian or Mongoloid hair that was opaque like that," referring to opaque features he described within the hairs, and added, "I haven't seen a hair like that before. Not a human hair."[179]

In his deposition for the Wilton Dedge case, David Jernigan of the Florida Department of Law Enforcement stated: "Out of all the pubic hairs that I have examined in the laboratory, I have never found two samples, two known samples to match in their microscopic characteristics."[180]

In the trial of Kevin Richardson and Kharey Wise in the so-called "Central Park Jogger" case, retired New York City Police Department Detective Nicholas Petraco was asked whether it was possible that the hairs found did not come from Richardson, but

---

[177] Id. at 249.

[178] Trial Transcript at 691–92, State v. Michael Nawee Blair, No. CRB 21,1152 (Tex. Dist. Ct. Sept. 12, 1994).

[179] Id. at 750–51.

[180] Deposition Transcript at 43, State v. Wilton Allen Dedge, No. 82-135-CF-A (Fla. Cir. Ct. June 7, 1984).

rather some unknown person. He responded that although it was possible "[i]n a sense," it was unlikely:

> I've looked at thousands of hair standards over the course of my work and I haven't seen any that have the same range of physical characteristics yet. But I really haven't looked at them in the sense of exclude one from the other. But I have in fact looked at thousands of standards and haven't seen two that matched exactly.[181]

Such testimony exaggerates the probative value of the hair association. As Detective Petraco, who did explain that hairs are not unique, indicated in passing, one would not expect an examiner to have difficulty distinguishing between thousands of standards known to have been obtained from different people. When an analyst claims that in a thousand cases he has never had two reference samples that were not microscopically distinguishable, the rareness of that event is not transferable to estimating the rareness of the association between a questioned hair and a set of exemplars from a known suspect.[182] In a typical sexual assault case the victim is female and the perpetrator is male. Thus, the two sets of reference hairs come from a female and a male. Moreover, the analyst knows in advance that the two sets of samples came from two different people—a rape victim and a suspect. Not only would an analyst be predisposed to differentiate the two sets from one another, but hair length alone can often distinguish such sets of reference samples. Since the analyst's experience in comparing reference samples to other reference samples answers a very different question than the one that is material to a criminal case, such testimony is misleading.

---

[181] Trial Transcript at 2838, People v. Kharey Wise, No. 4762/89 (N.Y. Sup. Ct. Nov. 13, 1990).

[182] See Saks & Koehler, supra note 46, at 212–13. Distinguishing one set of exemplars from another when the analyst knows a priori that they originate from two people also introduces context bias. It is a much easier task than comparing a single hair of unknown origin with a collection of hairs taken from a suspect. Moreover, a majority of the hair cases are sexual assaults where in which the victim is a female and the perpetrator is a male. Since one characteristic used to compare hair is length, the question needs to be asked whether in general, hair is more easily distinguishable between men and women. Finally, the hair analyst was limiting his experience to comparing standards in a single case. The analyst was not making inter-case comparisons.

Commonwealth of Pennsylvania v. Drew Whitley

Several cases also involved comparisons made where the analyst admitted that the questioned hairs were in fact unsuitable for comparison. For example, in the Drew Whitley case, Dorothy Menges of the Allegheny County Crime Laboratory testified, "Because these hair fragments were so small, I could not make the statement that they were microscopically consistent, but I did see so many overlapping characteristics within the questioned hairs and the standard hairs that I want to make some kind of statement as to their similarities."[183]

Menges then reversed course and testified, "I found no inconsistencies. Based on what I am basing my comparing on, yes, they are consistent."[184] After making those statements, she stated: "I wouldn't go that far to say they were microscopically consistent."[185] Those equivocations were deeply misleading and unsupported where the fragments were conceded as unsuitable for comparison.

Each of the above examples also suggests a related question: under what circumstances will an examiner conclude that the hair evidence excludes a criminal defendant? Earlier this Part described serology cases in which the analyst failed to exclude. In hair comparison cases, several examiners noted differences but nevertheless concluded that in their estimation, they were not sufficiently "material" so as not to find the hairs to be "similar."

Some of these cases involved testimony in which experts admitted that the samples possessed manifest differences, but invalidly told the jury that it would be impossible to exclude any person. An example is the case of James Waller, in which Timothy C. Fallon of the Dallas County Institute of Forensic Sciences testified that he "found three dark hair fragments with negroid characteristics that were different from the head and pubic hair of James Douglas Waller."[186]

To conclude that the hair possessed characteristics that could be described as "different," however, was not sufficient to induce

---

[183] Trial Transcript at 898–99, Commonwealth v. Drew Whitley, No. CC-8902462 (Pa. Ct. Com. Pl. July 21, 1989) [hereinafter Transcript, Whitley].

[184] Id. at 935.

[185] Id.

[186] Trial Transcript at 190, State v. James Douglas Waller, No. F-82-91294-TK (Tex. Crim. Dist. Ct. Jan. 31, 1983).

Fallon to tell the jury that Waller could not have been the source of that hair. Fallon relied upon speculation, stating that the hair

> could have come from another area of the body that we did not take a sample from. It could be that the sample that was the known standard sample that we took from the Defendant was not a large enough random sample to take in all the different characteristics that his hair possessed.[187]

Fallon then explained, "if you wanted to say that this hair did not come from this individual, you would have to check it against every hair to be positive that it did not come from that individual."[188] Fallon told the jury, in effect, that it would be an impossible task to conclude with certainty that a defendant was not the source of a specific hair. He agreed that one would "practically have to denude a person to make a proper comparison."[189] There was no suggestion that a similar effort should be made for a "proper comparison" permitting an analyst to say that a hair could have come from a defendant.

Similarly, in the case of Habib Abdal (named Vincent Jenkins at the time of trial), Michael R. Krajewski of the Central Police Services Laboratory in Erie County, New York, testified that he "could not make a positive comparison."[190] By that, he meant that "the hairs—hair samples were distinctively different," and he explained several key differences.[191] He added, "It's not unusual to have different hairs come from the same person."[192] On redirect, he explained that even if the exemplar hairs from Abdal did not match the questioned hairs, other hairs of his might match. His hair could have changed over time, or other hairs on his body might

---

[187] Id. at 194.

[188] Id. at 195.

[189] Id. at 194.

[190] Trial Transcript at 26, People v. Vincent Jenkins, No. 82-1320-001 (N.Y. Sup. Ct. June 2, 1983).

[191] Id. at 27 ("In order for a comparison to be made, two strands of hair would have to be exactly identical, and in this particular case, it varied in the diameter, which is measured under a microscope. The hair taken off the victim had been cut, the end had been cut, the hair taken from Mr. Jenkins had a tapered end, which meant that it was not cut. And the hair taken from Mr. Jenkins had a different medulla, which is the center part of the hair. And, in general, I cannot possibly say that the two hairs were similar.").

[192] Id.

match.[193] Krajewski testified there was a statistical possibility that other hairs on Abdal's body might match by citing to a study in an invalid way:

> The study shows it would not be unusual to have to look at 4,500 strands of hair from the head in order to get a match with any one particular hair. And, from the pubic hair, one may have to look at as much as 800 hairs, and it can be from the same person. That gives an idea of how much a hair can vary just within one single person.[194]

Again, there is no empirical data for the frequency of different characteristics that hair microscopists compare. Krajewski's testimony—even if reliance on the study referred to were appropriate—would suggest a statistical basis *not* to rely on the forensic method of hair comparison, which is based on selected exemplar hairs rather than on review of hundreds of hairs from a given person.

No hair comparison, resulting in either inclusion or exclusion of an individual, could be reliable if it were true that human hairs exhibit such variation. No studies have addressed that crucial question. Nor have any analysts in these cases suggested such reasons to doubt the methodology used when they readily concluded that hairs were similar. Instead, in these exonerees' trials, hair examiners made a range of invalid claims purporting to individualize hairs based on probabilities, supposedly "unusual" characteristics, or the examiner's undocumented experience. None of these statements were supported by empirical data.

C. Invalid DNA Testimony

Three of the 11 exoneree trial transcripts obtained that had testimony concerning DNA testing contained invalid testimony concerning the DNA testing. Of the other 8 cases, 1 involved a gross error in analysis; the last 7 involved DNA that excluded the defendant at the time of the criminal trial, in 3 of which the defense

---

[193] Id. at 36–37.

[194] Id. at 37–38. That study and the consensus that it should not be relied on to suggest a statistical basis for microscopic hair comparison is discussed supra note 158. In addition, the testimony misrepresents the erroneous conclusions of that study.

called the only expert witness.[195] The first invalid DNA testimony case, that of Gilbert Alejandro, involved egregious testimony by Fred Zain, who testified that he had conducted DNA testing and obtained results inculpating Alejandro. He told the jury, "the banding patterns that were identified from these items that you mentioned were identical to the banding patterns of Mr. Alejandro. As I stated in the report, they could only have originated from him."[196] He gave no random match criteria for this supposed DNA inclusion, but falsely told the jury that "DNA typing is a hundred percent identity as to whether a blood or body fluid may have originated from a particular donor or not."[197] A subsequent internal inquiry concluded that Zain had at best compared only partial banding pattern results visually; later tests excluded Alejandro.[198]

State of Texas v. Josiah Sutton

In the Josiah Sutton case, the victim had been raped by two men in the back seat of her car. Semen was present in the vaginal swab and on the stain removed from the back seat where the rape occurred. The official report authored by the Houston Police Department Crime Laboratory and the trial testimony of laboratory analyst Christy Kim presented invalid DNA results.[199] The raw data and the analyst's bench notes indicate that whereas the vaginal sample reflected a mixture of the victim's DNA and DNA from two male donors, the semen stain on the car seat suggested it came from one man only and that the lone male could not be Sutton. Although Sutton was excluded as the source of the car seat semen stain, that conclusion was not mentioned in the official report nor

[195] Those 7 cases are those of R. Alexander, J. Deskovic, C. Elkins, N. Hatchett, T. Hayes, E. Karage, and R. Mathews.

[196] Trial Transcript at 149, State v. Gilbert Alejandro, No. 90-09-8445-CR (Tex. Dist. Ct. Dec. 11, 1990).

[197] Id. at 146.

[198] See Evaluation of Out-of-County Serology Cases, Memorandum from Lonnie D. Ginsberg, Chief Forensic Serologist, Bexar County Forensic Science Center Criminal Investigation Laboratory, to Vincent DiMaio, Chief Medical Examiner (June 28, 1994); The Innocence Project, Profile: Gilbert Alejandro, http://www.innocenceproject.org/Content/47.php (last visited Nov. 10, 2008).

[199] See Trial Transcript at 168–230, State v. Josiah Sutton, No. 800450 (Tex. Dist. Ct. 1999), available at http://www.scientific.org/archive/Christy%20Kim%20Testimony.pdf [hereinafter Transcript, Sutton].

in the analyst's testimony at trial. Most importantly, if the back seat stain was considered in conjunction with the vaginal swab data, Sutton could probably have been ruled out as one of the rapists.[200]

Instead, the report erroneously concluded that the DNA profile on the seat stain was consistent with a mixture from Sutton, the victim, and another man. The report then states: "The DNA type of J. Sutton can be expected to occur in 1 out of 694,00[0] people among the black population."[201] But as explained in the section on serology, this frequency is irrelevant and misleading. The only relevant statistic is the combined frequency of all potential semen donors. Since the sample was supposedly a mixture of two male donors and it was impossible in this case to distinguish primary and secondary donors, the correct statistic for characterizing the value of the evidence is the sum of the frequencies of all possible donors. The final result would have revealed a potential donor population far larger and an event far more common than reported.[202]

At trial, Kim presented no statistics. However, she gave testimony that implied uniqueness for each DNA pattern and said that Mr. Sutton's DNA pattern was detected in the evidentiary samples.[203] Kim testified, for example, "If it came from one person, it should have a same exact DNA pattern. No other two persons will

---

[200] See William Thompson, Review of DNA Evidence in *State of Texas v. Josiah Sutton* (2003), available at http://www.scientific.org/archive/Thompson%20Report.PDF; see also Thompson, supra note 59, at 107–19. Since the victim had denied recent sex and indicated that the only event that could have produced the semen on the back seat was the rape, in all likelihood the single male profile on the stain from the seat—which excluded Sutton—was deposited by one of the two rapists. The profile from the seat stain is also consistent with one of the two male profiles contained in the vaginal swab. By a process of elimination, the genetic profile of the second rapist can be inferred. That second profile is also inconsistent with Mr. Sutton. In other words, the DNA evidence taken as a whole provides strong evidence of Sutton's innocence. The jury knew nothing about this exculpatory evidence. Cf. Michael R. Bromwhich, Final Report of the Independent Investigator for the Houston Police Department Crime Laboratory and Property Room 210 (2007), available at http://www.hpdlabinvestigation.org/reports/070613report.pdf (agreeing with Thompson's conclusion that test results on Sample #1 were not properly reported, but suggesting that they should have been reported as an exclusion as to that sample, but with an inconclusive result due to "Ms. Kim performing poor DNA testing on a potentially probative sample").

[201] Thompson, supra note 200, at 7.

[202] Id. (calculating the frequency of possible contributors as 1 in 15, or 1 in 8 for black men in Texas).

[203] Transcript, Sutton, supra note 200, at 181, 184–85.

have [the] same DNA except in the case of—of identical twins."[204] The jury was left with the mistaken impression that the DNA evidence uniquely identified Sutton as the rapist.[205]

State of Florida v. Chad Richard Heins

In Chad Heins's case, examiners at the Florida Department of Law Enforcement conducted then-available first-generation DNA testing. When they testified, the examiners failed to report the percentage of the population that could have contributed to the mixed samples found in a sink drain plug and washcloth, and instead left the impression that Chad Heins, who they explained had a genetic marker shared by only 8.5% of the population, could have contributed to the sample.[206] This testimony did not play a dispositive role in the case, however, because Chad Heins lived in the house from which the samples were taken, and could have been expected to have used the washcloth and sink. Ultimately, more sophisticated STR testing would show that the same unknown person's DNA profile was found in the sink drain, washcloth, and hairs.

A final case, that of Timothy Durham, involved not faulty testimony concerning DNA analysis (though the hair comparison testimony was invalid), but rather gross error in conducting the testing and interpreting the results.[207]

---

[204] Id. at 176.

[205] See Thompson, supra note 200, at 8.

[206] See Trial Transcript at 1158–59, State v. Chad Richard Heins, No. 94-3965-CF (Fla. Cir. Ct. Dec. 13, 1994). The DQ Alpha testing used only typed a single genetic marker, generating frequencies like in this case of 1 in 12. In contrast, modern STR DNA testing used today types 13 genetic markers and is capable of generating frequencies of one in many billions. See, e.g., Nat'l Comm'n on the Future of DNA Evidence, U.S. Dep't of Justice, The Future of DNA Testing 19 (2000).

[207] See William C. Thompson et al., How the Probability of a False Positive Affects the Value of DNA Evidence, 48 J. Forensic Sci. 47, 48 (2003) ("The initial DNA test result that helped convict Durham was proven to have been a false positive. The error arose from misinterpretation. The laboratory had failed to completely separate male from female DNA during differential extraction of the semen stain. The victim's alleles, when combined with those of the true rapist, produced an apparent genotype that matched Durham's. The laboratory mistook this mixed profile for a single source result, and thereby falsely incriminated an innocent man. Durham was released from prison in 1997.").

D. Invalid Bite Mark Testimony

Forensic odontology includes two very different disciplines. The older discipline involves the identification of a decedent by matching a well-preserved and three-dimensional set of teeth to dental records. Dentists perform a vital service in distinguishing among the dead in mass disasters such as plane and train crashes. X-rays collected from the deceased's dental records can be readily matched to the fillings and crowns observed in the mouth of the deceased. The second, and certainly more controversial, application involves the interpretation of lacerations, abrasions, and bruises of questionable origin on skin and, in particular, on decomposing skin. Although the small forensic dental community refers to the discipline as "bite mark" analysis, often the most challenging conclusion is the threshold question of whether the marks are due to a human bite as opposed to some other post mortem artifact. Unlike the wax mold a dentist makes in her office, skin, given its elasticity and potential for distortion, is a poor medium for accurately registering the bite marks.

Such bite mark work is "based on the assumption that every person's dentition is unique," though this assumption has not been tested.[208] Indeed, the NAS report noted that not only do "no scientific studies support" the assumption "that bite marks can demonstrate sufficient detail for positive identification," but that "[s]ome research is warranted in order to identify the circumstances within which the methods of forensic odontology can provide probative value."[209] After all, even if the assumption of dental uniqueness were established as true, that uniqueness would be far easier to identify from pristine wax molds made in a dentist's office than from the few distorted impressions left on the skin during a very dynamic biting situation. Nevertheless, courts permitted forensic

---

[208] 1 Giannelli & Imwinkelried, supra note 78, § 13-2; see also C. Michael Bowers, The Scientific Status of Bitemark Comparisons, *in* 3 Modern Scientific Evidence 538, 549–50 (David L. Faigman et al. eds., 2002) (reviewing "less than persuasive" literature and concluding that "[t]he demonstration of uniqueness is a blend of art and opinion"); Paul C. Giannelli, Bite Mark Analysis, 43 Crim. L. Bull. 930, 931–36 (2007).

[209] See Strengthening Forensic Science, supra note 12, at 5-37.

odontology testimony in the 1980s and continue to do so; "no reported case has rejected bite mark evidence."[210]

While "there is no quantitative base for bitemarks analysis,"[211] the guidelines promulgated by the American Board of Forensic Odontology ("ABFO") permit its members to render conclusions expressing near certainty—they may conclude that a bite mark matches a criminal defendant to a "reasonable medical certainty" and "high degree of certainty," explaining that the intended connotation is a "virtual certainty; no reasonable or practical possibility that someone else did it."[212] The guidelines counsel that, while experts may not convey "unconditional certainty," they may express "reasonable medical certainty"; moreover, "It is . . . acceptable to state that there is 'no doubt in my mind' or 'in my opinion, the suspect is the biter' when such statements are prompted in testimony."[213] No scientific criteria exist for what observations and analysis permit an expert to draw each type of conclusion.[214] Indeed, analysts conclude that variation between the bite mark and the defendant's teeth need not disturb a finding that the bite marks are consistent, and no guidelines explain which points of comparison are required for a positive identification (an ABFO effort in 1984 to adopt a scoring system was abandoned).[215]

---

[210] Paul C. Giannelli, Bite Mark Evidence, GP Solo (Sept. 2007), available at http://www.abanet.org/genpractice/magazine/2007/sep/trialpractice-bitemark.html; see also Giannelli, supra note 208, at 933 n.15 ("The overall 'uniqueness' of dental characteristics is a common statement used in court and in literature. This conclusion is generally accepted but is subject to considerable criticism. The reason it is criticized is that it has never been proven." (quoting C. Michael Bowers, Forensic Dental Evidence: An Investigator's Handbook 197 (2004))).

[211] Iain A. Pretty, Reliability of Bitemark Evidence, *in* Bitemark Evidence 531, 543 (Robert B.J. Dorion ed., 2005).

[212] ABFO Bitemark Methodology Guidelines, http://www.abfo.org/bitemark.htm (last visited Nov. 10, 2008).

[213] Id.

[214] See Bowers, supra note 208, at 565 ("The center point of disagreement amongst odontologists is the issue, 'what is necessary to support a positive identification from a bitemark?' The odontological literature is silent on the sufficiency of evidence necessary to accomplish this task, yet this positive opinion is permitted to any dentist.").

[215] See 1 Giannelli & Imwinkelried, supra note 78, § 13-2, -4. A 1999 ABFO study involving matching of four bite marks to seven sets of teeth produced a 63.5% false positive rate, and found "less than optimal accuracy." Kristopher L. Arheart & Iain A. Pretty, Results of the 4th ABFO Bitemark Workshop—1999, 124 Forensic Sci. Int'l 104 (2001) (noting that the study used only medium-to-good forensic quality bite marks); Bowers, supra note 208, at 545 (calculating false positive rate not presented in

Six cases involved bite mark comparison, and trial transcripts were located for all 6: the cases of Kennedy Brewer, Roy Brown, Ray Krone, Willie Jackson, James O'Donnell, and Calvin Washington. In all but James O'Donnell's and Calvin Washington's cases (in which the odontologist merely observed a consistency), the odontologists provided invalid testimony. In the Brewer, Krone, and Jackson cases, the odontologists testified they were certain that the defendant left the bite marks.

People of the State of New York v. Roy Brown

In Roy Brown's case, the defense presented an expert, Homer Campbell, who concluded that the bite marks were inconsistent with Brown's teeth because, among other manifest differences, one showed impressions of six teeth from the upper bite where Roy Brown had only four.[216] The prosecution never disclosed to the defense that the State Police forensic dentist, Lowell A. Levine, had previously opined to the prosecutor that the bite marks "excluded" Brown.[217] Instead, the prosecutor presented the testimony of Edward Mofson, who found the bite marks similar to "a reasonable degree of dental certainty" and called the differences "[i]nconsistent but explainably so in [his] opinion."[218]

State of Arizona v. Ray Milton Krone

Ray Krone's case was particularly troubling, for the bite mark evidence played a "critical" role: two experts concluded that the defendant made the bite mark on the victim, and Krone was then sentenced to death.[219] Experienced forensic odontologist Raymond Rawson presented the bite mark evidence at trial, along with John Piakis, a dentist who was inexperienced and just beginning to serve

---

the ABFO results and noting, "If this reflects their performance in actual cases, then inculpatory opinions by forensic dentists are more likely to be wrong than right").

[216] Trial Transcript at 951, 953, People v. Roy Brown, No. 91-2099 (N.Y. County Ct. Jan. 13–23, 1992) [hereinafter Transcript, R. Brown].

[217] Fernanda Santos, With DNA From Exhumed Body, Man Finally Wins Freedom, N.Y. Times, Jan. 24, 2007, at B5.

[218] Trial Transcript, R. Brown, supra note 216, at 740, 774.

[219] State v. Krone, 182 Ariz. 319, 322 (1995) ("The bite marks on the victim were critical to the State's case. Without them, there likely would have been no jury submissible case against Krone.").

as the police department's odontologist. Rawson presented a highly inflammatory and unusual video with images of Piakis holding molds of Krone's teeth to the marks on the deceased victim's body. Rawson attempted to quantify the probability of a tooth association:

> And it turns out that on average a tooth can be in about 150 different positions, each one of which is easily recognizable. And if you are looking at a tooth in that kind of detail, then you can see that very quickly. Just having two teeth, the possibilities of two teeth being in the same position, it would be 150 times 150, whatever that is. Maybe 1200 or something like that.[220]

Rawson then told the jury in no uncertain terms that Krone had left the bite marks:

> A. That's as nice a match as we—as we really ever see in a bite mark case.
>
> Q. By "nice" do you mean accurate?
>
> A. Yes. That was a nonscientific term. This is really an excellent match, and would be held in high regard by forensic odontologists.
>
> Now there's a wiping action just to show the same thing. Again, high correlation. I mean, that is—that tooth caused that injury.[221]

He concluded his testimony agreeing that "it was Ray Krone's teeth."[222] Piakis similarly testified, "I say that there is a match. Okay? I'm saying there's a definite match."[223]

The defense never learned that, before trial, police had initially consulted FBI odontologist Skip Sperber, who after examining the bite marks concluded, "It could not have been clearer.... Ray Krone had two higher teeth than his incisors that would have marked when he bit. Those weren't there in the evidence."[224] The

---

[220] Trial Transcript at 15, State v. Ray Milton Krone, No. CR 92-00212 (Ariz. Super. Ct. Aug. 4, 1992).

[221] Id. at 39.

[222] Id. at 57.

[223] Id. at 91.

[224] Robert Nelson, About Face, Phoenix New Times, Apr. 21, 2005, http://www.phoenixnewtimes.com/2005-04-21/news/about-face/.

Chicago Tribune later reported, "The discrediting of Rawson's testimony in the Krone case is one of numerous instances in which leading practitioners of bite-mark comparison have erred."[225]

E. Additional Forensic Techniques

1. Shoe Print Analysis

Three cases in the study set involved shoe print testimony. Two of the cases involved shoe prints that either excluded the defendant or were non-probative. The third is the case of Charles Fain.

State of Idaho v. Charles I. Fain

In Charles Fain's case, Robert B. Hallett of the FBI testified that the make of the shoe print was consistent with Fain's, and that "[i]t was possible that this shoe made this impression."[226] Not satisfied with his initial cautious conclusion, Hallett added that, although it was a common type of boat shoe sole, the wear patterns on the shoe individualized the print:

> Q. Okay, you also, if I understand correctly, that you said if another shoe made the impression, it would have to have the same characteristics as the actual left shoe that we have here?
>
> A. That's correct, sir.
>
> Q. What are those characteristics?
>
> A. The same size, the same design, and having general wear in exactly the same locations.
>
> Q. Now, did you indicate that the wear characteristics are put there by a gait of a particular individual?
>
> A. You would have to have the same characteristic walk as the individual who owned those shoes.

Indeed, Hallett also testified so as to imply that other examiners might have reached even stronger conclusions:

---

[225] Flynn McRoberts, Bite-Mark Verdict Faces New Scrutiny, Chi. Trib., Nov. 29, 2004, at 1.

[226] Trial Transcript at 281, 294, State v. Charles I. Fain, No. C-5448 (Idaho Dist. Ct. Oct. 14, 1983).

> I found, therefore, that the shoe which made this impression, and this left shoe had sustained wear in the same area. To a—a shoe print examiner, this would indicate that the individual who walked with these shoes has the same walking gait.
>
> Some examiners believe, I have not quite gone that far myself, but that could be a positive identifying characteristic. They believe we all walk differently.
>
> That wear corresponded exactly.[227]

This practitioner suggested that the effect of gait on the sole of a shoe is unique. No data supports such an opinion. Adding a clever embellishment, he testifies that other examiners would go even further to say that wear patterns on shoes "correspond[] exactly." Unfortunately, that is the case: other examiners may indeed go further in their testimony, on the recommendation of the Scientific Working Group on Shoeprint and Tire Tread Evidence ("SWGTREAD"), which offers the guideline that an examiner can find an "[i]dentification (definite conclusion of identity)."[228] The guideline explains that "this is the highest degree of association expressed in footwear and tire impression examinations. This opinion means that the particular shoe or tire made the impression to the exclusion of all other shoes or tires."[229] No scientific criteria are provided regarding when an expert may render any of those conclusions.[230]

2. Fingerprint Analysis

Fingerprint comparisons were conducted in 14 exonerees' cases.[231] Trial transcripts were located for 13 of these cases. Two in-

---

[227] Id. at 298.

[228] Scientific Working Group on Shoeprint and Tire Tread Evidence, Int'l Ass'n for Identification, Standard Terminology for Expressing Conclusions of Forensic Footwear and Tire Impression Examinations 1 (Mar. 2006), http://www.theiai.org/guidelines/swgtread/terminology_final.pdf.

[229] Id.

[230] See Yaron Shor & Sarena Weisner, A Survey on the Conclusions Drawn on the Same Footwear Marks Obtained in Actual Cases by Several Experts Throughout the World, 44 J. Forensic Sci. 380, 383 (1999) (finding a wide range of variability in criteria experts use to draw conclusions in shoe print cases).

[231] In several additional cases, examiners did not conduct comparisons because latent prints were unsuitable for comparison.

volved troubling testimony or analysis; the others all involved fingerprint exclusions at trial (and in one the only expert testified for the defense).[232] In the first troubling case, that of Gene Bibbins, as discussed further below, the examiner testified that the comparison between his fingerprints and latent prints found on the window fan in the victim's room was non-probative, when in fact the Louisiana State Crime Lab had excluded Bibbins and documented its contrary finding in a report not disclosed to the defense.[233]

Commonwealth of Massachusetts v. Stephan Cowans

In the case of Stephan Cowans, a Boston police officer was shot by a civilian. In the course of the assailant's escape, the assailant picked up a glass mug, drank from it, and put it back down. The crime scene unit promptly vouchered the mug and lifted two latent prints from it. After Cowans became a suspect, Boston Police latent expert Dennis LeBlanc compared Cowans's known ink thumb print to one of the latent prints and declared a match. The second print was favorably compared to the woman who owned the mug. After Cowans was exonerated by post-conviction DNA testing, the District Attorney asked the Massachusetts State Police to reexamine the thumb print. The State Police declared that Cowans was clearly excluded.[234]

The Boston Police then hired an external auditor, Ron Smith & Associates, to conduct an independent investigation into, among other things, Officer LeBlanc's conduct in the case. The audit team had four members, all experts in fingerprint comparison.[235] The auditors reached the unanimous conclusion that Officer LeBlanc realized at some point prior to trial that Cowans was excluded, but

---

[232] The other cases were those of Antonio Beaver, Michael Cromedy, Frederick Daye, James Giles, Dennis Halstead, Anthony Hicks, Larry Mayes, John Restivo, Samuel Scott, James Tillman, and Ron Williamson. Cromedy, Daye, Giles, Halstead, Hicks, Mayes, Restivo, and Tillman's cases involved fingerprint exclusions. Scott's case involved a conclusion that a fingerprint belonged to Scott, a conclusion that was not terribly probative since it was found on his glass in his house. Curtis McCarty's case also involved a matched latent print; that portion of his trial transcript, however, has not been located.

[233] See discussion infra Subsection II.F.1.

[234] See Possley, supra note 9.

[235] Ron Smith & Associates, Inc., Reference: Request for Latent Print Consultation Services 6 (Mar. 8, 2004).

that he nevertheless concealed that fact in his trial testimony.[236] Instead, Officer LeBlanc misrepresented to the jury that the latent print matched Cowans's. The auditors' conclusion was based on facts including: Cowans's exclusion was clear to every member of the review team; Officer LeBlanc had made correct associations and exclusions routinely in more difficult cases over the preceding four years; he made efforts to conceal other errors made in the same case; there were numerous inconsistencies in his testimony; and he intentionally used a method of showcasing the erroneous Cowans match evidence to the jury that not only made it more difficult for the jury to follow but was contrary to the preferred methods of fingerprint examiners and contrary to what Officer LeBlanc did with the other latent print in the same case.[237]

In the other fingerprint cases, the evidence played little role. For example, in the cases of James Giles and Michael Cromedy, the examiners testified that the prints excluded the defendants; similarly, in the Dennis Halstead and John Restivo cases, the fingerprints did not match any known person.

3. Voice Analysis

One final case, that of David Shawn Pope, involved voice comparison using a spectrograph, an instrument that generates a visual pattern depicting an audio recording using lines that represent the frequency and intensity of the sound wave over time.[238] Although the National Academy of Sciences ("NAS") issued a report in 1979 concluding that the use of voice spectrograph analysis to identify individuals "is not adequately supported by scientific data" (after which the FBI stopped permitting court testimony concerning such analysis), a series of courts continue to permit testimony concerning voice spectrography, apparently acting "as if the NAS Report did not exist."[239]

In Pope's case, the victim of a 1985 rape in Garland, Texas, received several messages on her answering machine shortly after the crime. The Dallas County police arrested Pope after the victim

[236] Id. at 26–27.
[237] Id.
[238] See The Scientific Basis of Expert Testimony on Talker Identification, *in* 5 Modern Scientific Evidence § 37-1 (David L. Faigman et al. eds., 2007).
[239] Id.

identified him in a lineup, and later conducted voice spectrograph analysis comparing a recording of Pope's voice with the messages on the victim's answering machine. The State retained Henry Truby as an expert. He testified, based on finding "10–15 similar patterns" shared by the recording of Pope and the recording from the victim's answering machine, that "the original producer of [the recordings] was the same individual."[240] Truby explained:

> A. I found a sufficient number [of identical patterns] to serve as an identification to convince me, and then take a few more just to reinforce it, that no matter how much you do of these samples, you would continue to get points of similarity every now and then.
>
> Q. All right. Let me ask you then, so that it is clear, are you saying the known tape and the unknown tape were made by one and the same person?
>
> A. I do so state.

Truby testified at the end of his direct examination:

> Q. The bottom line analysis on the known voice and the unknown voice in this situation were only made by one single person in the whole wide world?
>
> A. Exactly.
>
> Q. Just like fingerprints, it is unique?
>
> A. Exactly.[241]

In addition to voice spectrography being found unreliable by the NAS panel and barred from use in court by the FBI, no study has suggested that an analyst can conclude that only one person in the world could produce a particular pattern exhibiting certain similarities with an unknown pattern. Indeed, the defense retained Stuart I. Ritterman, a professor of communicology at the University of South Florida, who testified that studies show that spectrography

---

[240] Trial Transcript at 290, State v. David Shawn Pope, No. F85-98755-NQ (Tex. Dist. Ct. Feb. 4, 1986).

[241] Id. at 295.

"is totally unsuitable as a tool for indentifying voices with any degree of accuracy."[242]

F. Forensic Science Misconduct Beyond Trial Testimony

1. Withholding Forensic Evidence

Thirteen cases were identified as involving either a failure to disclose exculpatory data or analysis, or outright fabrication of evidence.[243] Examples included withholding laboratory reports, analysis, or the existence of evidence. Other cases involved fabrication, including falsifying or altering lab reports. Putting aside the examples of fabrication, this study does not opine whether evidence was withheld due to deliberate, reckless, or negligent acts.[244] The known failures to disclose forensic evidence helpful to the defense remain only a subset; other evidence of innocence may remain undisclosed to this day. From the trial transcripts alone it is impossible to know whether material was concealed. Even with the benefit of bench notes or laboratory reports, one may not be able to ascertain whether experts falsified or concealed test results.

Commonwealth of Virginia v. Earl Washington, Jr.

Data or analysis was withheld in cases involving a number of forensic disciplines. Several cases involved serology. One example is in the case of Earl Washington, a man sentenced to death in Virginia for a rape and murder that he did not commit. The victim was a Caucasian woman who stated before she died from severe knife wounds that the attacker was African-American. The state never disclosed to the defense a serology report, conducted early in the investigation and dated August 19, 1982, which found that stains on a central piece of evidence, a light blue baby's blanket on the murdered victim's bed, were not only ABO Type A, PGM Type 1, but

---

[242] Id. at 321, 329.

[243] The cases are those of G. Bibbins, R. Brown, S. Cowans, W. Gregory, R. Krone, C. McCarty, N. Miller, J. Sutton, E. Washington, and K. Waters.

[244] State of mind is not relevant to the inquiry under *Brady v. Maryland*, 373 U.S. 83, 87 (1963). In these cases the evidence did not surface until after post-conviction DNA testing, post-exoneration investigations, or civil suits.

also "Tf CD."[245] As a subsequent October 22, 1982, police report noted, Transferrin CD is a fairly uncommon plasma protein that is most often found in African-Americans. Virginia Bureau of Forensic Science examiner Deanne Dabbs told the Virginia State Police that "the Tf CD is found in 10% of Negroes but to her knowledge has never been found in Caucasians. In order to be sure of this determination, she stated she ran a second test with the same results."[246]

This highly probative serology finding regarding the Tf CD was never disclosed to the defense. Instead, the initial report was altered, but only after Earl Washington was identified as a suspect in 1983. Earl Washington did not possess the unusual "Tf CD" characteristic. Rather than report to the defense that Washington lacked the Tf CD that had been observed on the stains, the state created a second "amended" report, dated August 23, 1983, and provided it to the defense. That second report, issued without having conducted any further testimony undermining the original results, nevertheless stated that "[t]he results of Tf typing were inconclusive."[247] Thus, law enforcement concealed probative information regarding the blood type of the perpetrator.

Other cases involved concealment of exculpatory information regarding hair comparison. In William Gregory's case, Dawn Katz concealed the fact that she determined that at least one hair was not consistent with Gregory's hair.[248] Joyce Gilchrist concealed and altered laboratory reports in which she had initially excluded Curtis McCarty, which led a court to grant a new trial because of her *Brady* violations.[249]

---

[245] Certificate of Analysis, Commonwealth of Virginia Bureau of Forensic Science (Aug. 19, 1982).

[246] SA C.R. Wilmore, Virginia State Police Report (Oct. 22, 1982) (emphasis omitted).

[247] Amended Copy of Certificate of Analysis dated August 19, 1982, Commonwealth of Virginia Bureau of Forensic Science (Aug. 26, 1983).

[248] See Gregory v. City of Louisville, 444 F.3d 725, 732 (6th Cir. 2006) (affirming denial of absolute or qualified immunity to Katz, stating that "Katz's deposition for this instant action revealed that Katz had actually found 7 negroid head hairs on the pantyhose, only 5 of which she found similar to Plaintiff's hair").

[249] McCarty v. State, 114 P.3d 1089, 1092, 1095 (Okla. Crim. App. 2005); see also The Innocence Project, Curtis McCarty, http://www.innocenceproject.org/Content/576.php (last visited Sept. 14, 2008).

In two bite mark comparison cases already noted—those of Roy Brown and Ray Krone—the state concealed that bite marks had been shown to other odontologists who excluded the defendant. The Josiah Sutton case discussed earlier involved gross overstatement of DNA results and a "failure to report aspects of the DNA evidence that appear[ed] to exonerate Josiah Sutton."[250] Two fingerprint cases, including the Stephan Cowans case described above,[251] involved the withholding of exculpatory evidence.

State of Louisiana v. Gene Bibbins

In Gene Bibbins's case, Annie Michelli of the Baton Rouge City Police had testified at trial that any comparison between Bibbins's fingerprints and a latent print found on the window at the crime scene was inconclusive, explaining that "[t]he latents are unidentifiable. You can't—they aren't any—there aren't any prints on there that we can use."[252] When asked, "Did you double-check your conclusion with the state crime lab?" and, "Did they have the same results?" she answered to both questions, "Yes, Ma'am."[253] That testimony was false; the state crime lab's finding and report had excluded Bibbins. The district court in a civil rights lawsuit filed by Bibbins denied Michelli's motion for summary judgment on Bibbins's *Brady* claim, stating:

> Michelli's testimony at Bibbins'[s] trial was that she was unable to identify Bibbins as a match to the fingerprint sample. Michelli also testified that she double checked her results with the Louisiana state crime lab and that the state crime lab reached the same results. However, it is undisputed that a Louisiana state police crime lab report made by Sybil Guidry showed a contrary result. Guidry's findings excluded Bibbins as a match.[254]

Additional cases involving the withholding of exculpatory evidence have been documented, but the authors have not yet obtained transcripts; thus, those cases were not included in this

---

[250] See Thompson, supra note 200, at 2.

[251] See supra Section II.E.

[252] Trial Transcript at 83, State v. Gene C. Bibbins, No. 2-87-979 (La. Dist. Ct. Mar. 25, 1987).

[253] Id. at 83, 84.

[254] Bibbins v. City of Baton Rouge, 489 F. Supp. 2d 562, 572 (M.D. La. 2007).

study.[255] The case of Armand Villasana also demonstrates the importance of the discovery of underlying laboratory notes. In that case, Villasana, another DNA exoneree (not included here because the trial testimony was never transcribed) later brought a civil suit because the forensic analyst never provided laboratory notes showing that material existed with which to perform DNA testing at the time of trial.[256]

2. Gross Error in Analysis

In a few cases, reanalysis of the evidence conducted post-conviction uncovered gross errors that were not apparent at the time of trial. Again, these cases represent only those in which errors came to light due to the rare use of post-conviction reevaluation or retesting. Few exonerees have had the forensic analysis evidence in their cases retested or reexamined. The gross errors that were uncovered include mistyping serological evidence, failing to notice abundant spermatozoa, erroneously linking large numbers of hairs, failing to note material differences during comparisons, and failing to use equipment properly. It is not known how many other exonerees, much less non-exonerated convicts, were convicted based on grossly erroneous forensic testing or analysis.

For example, in several serology cases, evidence was reported as non-probative and not more elaborately tested due to a supposed failure to observe spermatozoa; subsequent examiners who thereafter reviewed the same evidence found abundant spermatozoa. In

---

[255] In a series of cases, forensic expert Pamela Fish notoriously concealed evidence. None of those transcripts were included in the above figures because they have not yet been obtained—though the transcripts would shed little light on the matter, as the issue is precisely that exculpatory evidence was not disclosed to the defendants' trial counsel and did not arise at trial. For example, in John Willis's case, Fish testified that her tests were "inconclusive," but when DNA testing exonerated him years later, her lab notes surfaced showing Willis was a Type A secretor, whereas the material tested indicated a Type B semen donor. See Paul C. Giannelli, Bench Notes & Lab Reports, 22 Crim. Just. 50, 50 (2007).

[256] See Villasana v. Wilhoit, 368 F.3d 976, 979 (8th Cir. 2004) (holding forensic technician's notes underlying disclosed lab reports on DNA testing were not exculpatory, even though they led the defense to perform additional testing, because *Brady* applies only to evidence a reasonable prosecutor would identify at the time as material); cf. ABA Criminal Justice Standards for DNA Evidence Standard 16-4.1 (3d ed. 2007) (recommending that not only laboratory reports but underlying laboratory case file and notes be maintained and disclosed to the defense).

the Larry Peterson case, "[a]lthough the New Jersey State Police Laboratory had reported that there was no semen in the victim's rape kit," the Serological Research Institute, before conducting its post-conviction DNA testing, "identified sperm on her oral, vaginal, and anal swabs."[257]

The Ulysses Charles case provides another example of this type of error in conducting presumptive testing. Charles was a Type B secretor, while the victims were Type O secretors and the stains on a robe and sheet contained only H antigens consistent with that O type.[258] The prosecution called two experts to explain why no B antigens consistent with Charles were observed. Stanley Bogdan of the Boston Police Department Crime Laboratory explained that though acid phosphatase was detected through assays, P-30 testing did not indicate the presence of sperm. A second expert, John Cope Abbott, also explained the P-30 test results and noted that no spermatozoa were observed.[259] We now know this was gross error, for when Cellmark later performed DNA testing, spermatozoa were readily observed on the same robe and sheet.[260] The technology for identifying the presence of sperm, a conventional microscope, has remained unchanged for decades. There was no technological reason why the spermatozoa could not have been observed in 1984 when Charles was tried.

The serology cases involving conjectural theories of contamination all involve gross error. As described above, in the case of Gary Dotson, Edward Blake found gross error not only because the phenomenon of masking was not explained, but also because the analyst did not attempt to use control testing to assess whether contamination was a proper explanation for the finding of a blood type that was inconsistent with both Dotson and the victim.[261] Other serology cases involved mistyping. For example, in the Ford Heights Four case, Chicago Police Department examiner Michael

---

[257] The Innocence Project, Larry Peterson, http://www.innocenceproject.org/Content/148.php (last visited Nov. 10, 2008).

[258] Trial Transcript at 7-28 to -32, Commonwealth v. Rodriguez U. Charles, No. 035492-45 (Mass. Super. Ct. Feb. 9, 1984).

[259] Id. at 7-49 to -50.

[260] Letter from Gina Pineda, Reliagene Technologies, Inc., to Myeti Gametchu (Sept. 26, 2007) ("Cellmark microscopically observed sperm cells in each of the sperm f[r]action samples . . . .").

[261] Affidavit of Edward T. Blake, supra note 3 at 23–25.

Podlecki reported that Dennis Williams was a Type A secretor,[262] when Edward Blake found post-conviction that in fact he was a Type A non-secretor.[263]

Gross error in several bite mark comparison cases has just been discussed, as well as in the DNA testing conducted in the Timothy Durham case. In several hair comparison cases, reexamination of the evidence was conducted as part of a post-exoneration investigation, and reports found the comparison at the time of trial to have been erroneous. For example, in the Jimmy Ray Bromgard case discussed earlier, the FBI's reexamination revealed that the head and pubic hairs that Arnold Melnikoff had found microscopically indistinguishable were in fact "microscopically dissimilar" to Bromgard's, and that the head hair was in fact similar to the victim's.[264]

Additional examples in which hair evidence matched the actual perpetrator are discussed in the next Section. Additional egregious examples of error in hair comparison cases include cases where the examiner compared large numbers of hairs and still erroneously linked all of them to an innocent man. As the Bromgard Peer Review Report concluded, "While an experienced hair examiner might erroneously associate a single head or pubic hair, it is highly unlikely that a competent hair examiner would incorrectly associate both head and pubic hairs."[265] In quite a few cases, many more than one or two hairs were incorrectly associated with the defendant. For example, in the Curtis McCarty case, Joyce Gilchrist not only altered lab reports, but found dozens of hairs to have been consistent with McCarty.

3. Failures to Conduct Elimination Testing or Comparison

Related to the problem of gross error, forensic analysts in several cases stated that they failed to conduct serology testing on relevant potential sources. For instance, if the semen excludes the

---

[262] See Transcript, Rainge, supra note 168, at 2281.

[263] See Center for Wrongful Convictions, Dennis Williams, http://www.law.northwestern.edu/wrongfulconvictions/exonerations/ilWilliamsChart.pdf (last visited Dec. 1, 2008).

[264] See Adam Liptak, 2 States to Review Lab Work of Expert Who Erred on ID, N.Y. Times, Dec. 19, 2002, at A24.

[265] See Peer Review Report, supra note 159, at 2.

defendant, one would want to perform testing on a potential consensual donor such as a husband or boyfriend. However, neither prosecutors nor analysts have a legal duty to pursue exculpatory evidence; they need only disclose exculpatory evidence that they uncover.

Commonwealth of Massachusetts v. Neil Miller

In the Neil Miller case, a semen stain on the sheet where the victim was raped produced both B and H blood group substances, but neither the victim nor Miller possessed the B antigen. At trial, the prosecution implied in the opening arguments that the Type B semen came from the boyfriend of the victim's roommate: "A sample of this defendant's blood and saliva was later obtained, and it was determined that the semen from the sheet was not the defendant's semen. [The victim] will testify that her roommate did have a boyfriend who sometimes stayed overnight."[266] Although this hypothesis could easily have been tested, it never was. Post-conviction DNA testing revealed that the semen stain on the sheet had in fact been deposited by the rapist—not by Miller nor by the boyfriend.[267]

State of Oklahoma v. Ronald Keith Williamson

Several exonerees' cases involved not only false positives, but also false negatives. In several cases, elimination testing was not done on a known suspect that subsequent DNA testing revealed to have been the actual perpetrator. For example, in Ronald Williamson's case, Melvin R. Hett, a supervisor at the Oklahoma State Bureau of Investigation, Northwest Regional Laboratory, testified unequivocally that he had compared the hairs of the state's star witness, Glen Gore, with those at the crime scene. He testified, "I did, direct comparison with the unknown hairs," and when asked if any of Gore's hairs were microscopically consistent with the questioned hairs, he testified, "No, sir."[268] Later, during Williamson's appeal, the Court of Appeals for the Tenth Circuit reviewed Hett's

---

[266] Transcript, N. Miller, supra note 64, at 1-169.

[267] See Frontline, Burden of Innocence, http://www.pbs.org/wgbh/pages/frontline/shows/burden/profiles/miller.html (last visited Nov. 10, 2008).

[268] Trial Transcript at 736–37, State v. Ronald Keith Williamson, No. CRF 87-90 (Okla. Dist. Ct. Apr. 21–28, 1988).

lab report, which revealed that Hett's testimony was false. "In fact, the hair expert compared samples from Mr. Gore with hairs already determined to be consistent with those of the victim, [co-defendant] Mr. Fritz, and Mr. Williamson, but [did] not compare Mr. Gore's samples with unidentified hairs."[269] Indeed, Hett also opined on the significance of a "match" of 17 hairs, including both scalp and public hairs. It was later determined that none of the hairs belonged to Ronald Williamson or his co-defendant Dennis Fritz. Glen Gore—whose hair exemplars were, contrary to Hett's false testimony, never compared to the "unidentified hairs"—was shown by post-conviction DNA testing to have been the actual perpetrator.[270]

State of Oklahoma v. Robert Miller

In the Robert Miller case, Joyce Gilchrist excluded as the originator of the questioned hairs a suspect who was later identified by post-conviction DNA testing and indicted.[271] Post-conviction analysis by Microtrace strongly disagreed with Gilchrist's findings, finding no similarities and highly varied reference samples.

These failures to rule out other possible suspects or assess whether material could have come from a partner and not the perpetrator may have occurred more often than in the known cases. Rarely did the forensic expert mention during the trial testimony whether elimination analysis was conducted. Also troubling is the fact that some experts made clear during their trial testimony that they only performed testing as requested by police or prosecutors, rather than on all materials that could be probative in the case. Failures to conduct testing occurred even as to highly probative materials and tests. For example, at least 7 exonerees were tried at a time when DNA testing was technologically available but not used.[272]

---

[269] Williamson v. Ward, 110 F.3d 1508, 1522 n.15 (10th Cir. 1997).

[270] See Barry Scheck et al., Actual Innocence: Five Days to Execution and Other Dispatches from the Wrongly Convicted 165 (2000).

[271] The Innocence Project, Robert Miller, http://www.innocenceproject.org/Content/219.php (last visited Nov. 10, 2008) ("The 1996 DNA tests ultimately led to Miller's exoneration and release in 1998. The tests also identified the true perpetrator, Ronald Lott, a man whose samples had been included in all rounds of testing.").

[272] For a discussion of those cases, see Garrett, supra note 55.

To conclude this Part, invalid science testimony in these exonerations cases raised a related set of problems. Analysts failed to adhere to scientific standards and instead exaggerated the probative value of evidence to make it appear more inculpatory than it in fact was. Aside from serology and DNA testing, the relevant disciplines lacked scientific standards defining the permissible scope of conclusions reached regarding comparison. Invalid testimony should come as no surprise given the lack of such standards.

III. REFORMING FORENSIC SCIENCE

This Part steps back to look at the roles of criminal justice actors and the possibilities for systemic reform originating from the scientific community. The first Section examines the existing regulation of forensic science, beginning with the roles of other criminal justice actors in these 137 trials, specifically: (1) prosecutors, who often misrepresented forensic evidence during closing arguments; (2) defense attorneys, who rarely received their own experts and rarely effectively cross-examined forensic analysts concerning invalid science testimony; and (3) judges, who when called on to rule regarding invalid forensic science testimony at trial or post-conviction rarely provided relief. Where each of those criminal justice actors failed to correct these miscarriages of justice, this Article concludes by developing a framework for national regulation and oversight of the forensic sciences. The renewed scrutiny of forensic science error may finally provide the impetus for federal legislation to ensure a sound scientific foundation for forensic sciences, including the disciplines that are the focus of this study.

A. The Roles of Judges, Prosecutors, and Defense Attorneys

1. Invalid Prosecution Use of Forensic Science

Although the trial transcripts cannot tell us whether prosecutors sought out invalid forensic testimony, they certainly display prosecutors, knowingly or not, developing and relying on such testimony.[273] The Josiah Sutton case provides an example where, ac-

[273] See William C. Thompson, A Sociological Perspective on the Science of Forensic DNA Testing, 30 U.C. Davis L. Rev. 1113, 1115 (1997); William C. Thompson & Edward L. Schumann, Interpretation of Statistical Evidence in Criminal Trials: The

cording to William Thompson's investigation, the prosecutor diverted the forensic expert's testimony from the fact that there was a second unknown DNA profile, because the presence of two profiles meant that Sutton was excluded—not included as the examiner had testified.[274]

Opening arguments by prosecutors, when they included descriptions of the forensic evidence that overstated its probative value, may indicate that the prosecutor had met with the forensic expert and knew about the claims being advanced (of course, whether the prosecutor knew the claims were invalid cannot be ascertained). For example, in Jimmy Ray Bromgard's case, Deputy County Attorney David W. Hoefer anticipated Arnold Melnikoff's invalid testimony, telling the jury in his opening that

> the experts at the State Lab out of Missoula will come and testify, and they will tell you that that hair has the same range of microscopic characteristics as that of the defendant, and they will tell you the percentage of the population that would have that kind of hair, first for the head hair, secondly for the pubic hair, and then for the two combined.[275]

In a number of cases, the prosecutor exaggerated the testimony of the forensic analyst in closings, making claims that the forensic scientist clearly did not make. Twelve were in cases with valid testimony by all forensic analysts (an additional 6 cases included both invalid and valid analyses extended by the prosecutor). Convictions should not necessarily have been reversed on those grounds, nor did these prosecutors necessarily engage in misconduct or ethical lapses. The ethical and criminal procedure rules regarding closing statements "offer[] few standards for proper prosecutorial argument," and though prosecutors may not misrepresent facts in evidence, they may make arguments concerning inferences to be drawn from the facts.[276] There may be a fine line between properly

---

Prosecutor's Fallacy and the Defense Attorney's Fallacy, 11 Law & Hum. Behav. 167 (1987).

[274] Thompson, supra note 200, at 1; Thompson, supra note 59, at 119–21.

[275] Transcript, Bromgard, supra note 160, at 18.

[276] Fred C. Zacharias, Structuring the Ethics of Prosecutorial Trial Practice: Can Prosecutors Do Justice?, 44 Vand. L. Rev. 45, 96 (1991); see also Ann. Model Rules of Prof'l Conduct R. 3.4(e) (6th ed. 2007) (providing that a lawyer shall not "in trial,

drawing inferences and misstating facts. Even if prosecutors draw invalid inferences or make false statements to the jury, they may not face any consequences. In reviewing such claims of prosecutorial misconduct, courts examine the severity of the alleged misconduct, the strength of the State's case, and whether curative measures were taken. In doing so, appellate courts often find any error to be harmless. Federal courts limit relief to egregious cases in which the conduct "so infected the trial with unfairness as to make the resulting conviction a denial of due process."[277]

This Article thus does not address whether any particular statements constituted misconduct, but rather emphasizes that a series of closing statements did not comport with the science and left the jury with a faulty impression of what the forensic science had actually shown. In the case of Stephen Linscott, the Illinois courts reversed the conviction based on a finding of egregious prosecutorial misconduct concerning the forensic evidence as presented during closing arguments. The Appellate Court of Illinois explained, as to the serology:

> No one testified that "[the victim] was raped by a non-secretor" or that the seminal material "came from a non-secretor." The prosecutor simply made-up that piece of "evidence." The made-up evidence was doubly devastating because not only was it false, but it reduced the pool of possible assailants from a substantial percentage of the male population, or even from the entire population, to just the males in twenty percent of the population.[278]

In affirming the vacatur, the Illinois Supreme Court noted: "A prosecutor must confine his arguments to the evidence and to 'reasonable inferences' that follow from it. We believe that the prosecutor in the instant case contravened this fundamental rule."[279] However, illustrating the difficulty of prevailing on such claims, the Illinois Supreme Court, though reversing based on statements con-

---

allude to any matter that the lawyer does not reasonably believe is relevant or that will not be supported by admissible evidence").

[277] Darden v. Wainwright, 477 U.S. 168, 181 (1986); see also, e.g., State v. Graves, 668 N.W.2d 860, 877 (Iowa 2003).

[278] People v. Linscott, 511 N.E.2d 1303, 1306 (Ill. App. Ct. 1987).

[279] People v. Linscott, 566 N.E.2d 1355, 1362–63 (Ill. 1991) (internal citations omitted).

cerning the serology, found that the statements made by the prosecutor concerning the hair evidence alone did not warrant relief: "The prosecutor's 'match' comments were improper, but we do not find, as did the appellate court, that they were so 'egregious' as to deny defendant a fair trial."[280]

In other cases, courts did not provide any remedy at all. In the Larry Ollins case, the prosecutor—after stating that Ollins's blood and hair shared characteristics with evidence found at the crime scene—referred to Ollins as a man "whose semen matches what's found at the scene. Whose hair is found on the seat of the car."[281] When the issue was raised on appeal, the court denied relief, emphasizing that "[a] prosecutor is permitted a wide range of expression in comments made during closing argument. Reversal will not be warranted even if the prosecutive comments could be viewed as improper, unless the defendant is substantially prejudiced."[282]

In the Drew Whitley case, the criminalist, Dorothy Menges, the Forensic Serology Section Laboratory Manager at the Allegheny County Crime Laboratory, examined a number of very short shaved or cut hair fragments found on a stocking apparently worn by the perpetrator as a mask and found at the crime scene. Menges, though finding similarities despite admitting the hair fragments were unsuitable for comparison, was clear: "Because these hair fragments were so small, I could not make the statement that they were microscopically consistent."[283] During his closing argument, Assistant District Attorney Nicholas Radoycis extended Menges's already invalid testimony, stating, "She said all the characteristics overlap, came from the same individual."[284] Radoycis

---

[280] Id. at 1361.

[281] Trial Transcript at 62, People v. Larry Ollins, No. 02 L 000749 (Ill. Cir. Ct. June 20, 1988).

[282] People v. Ollins, 601 N.E.2d 922, 925 (Ill. App. Dist. 1992) (citation omitted). The Brian Piszczek case provides another example regarding serology evidence. The analyst properly explained that any male could have been the perpetrator where the stains exhibited the same A type as the victim and Piszczek was a non-secretor. Trial Transcript at 158, State v. Brian Piszczek, No. 257813 (Ohio Ct. Com. Pl. June 26, 1991). Yet the prosecutor argued in his closing argument, "What it is consistent with, ladies and gentlemen, that the person who did it would have been a non-secretor and who was a non-secretor? This man." Id. at 260.

[283] Transcript, Whitley, supra note 183, at 898. Menges never concluded that there was "no doubt" the hairs came from Whitley, as the prosecutor claimed.

[284] Id. at 43.

then added, "Let's see what the Crime Lab says. Dorothy Menges, do you remember the last thing she said on that witness stand Friday afternoon? I have no doubt those hairs came from Drew Whitley."[285] In fact, she had specifically rejected that conclusion. When the court asked her at the close of her testimony, "You can't say it belongs to the defendant," she had answered, "That is correct."[286] Despite stating the opposite of her actual testimony, Radoycis embellished further, telling the jury: "But it's only when the scientists come in and say, hey, we have a standard, we know this hair to be of Drew Whitley and they compare it all microscopically. Exact. No doubt about it. (Pointing.) Him."[287] In response, the defense attorney brought a motion for mistrial:

> The District Attorney stated to this jury that Dorothy Menges testified under oath that these hairs that she was comparing came from Drew Whitley. That is absolutely, positively not the evidence; and that is the most vital part of this whole case; and for him to say that constitutes prosecutorial misconduct.... She never said that they came from my client, Your Honor.[288]

Judge Walter R. Little equivocated, stating, "I do recall that she answered my question as she couldn't say exactly who those hairs belonged to.... I don't know if she did say it. I don't recall." When the prosecutor claimed he did hear such a statement and asserted that "[i]t's the jury's responsibility to remember things," Judge Little provided a curative instruction that told the jury to resolve any discrepancy themselves.[289]

Each of these examples suggests the importance not just of accurate forensic testimony at trial, but the importance of the defense objecting and the court providing curative instructions should the

---

[285] Id. at 45.
[286] Id. at 950.
[287] Id. at 50.
[288] Id. at 51–52.
[289] Id. at 53–55. The judge instructed:

> I recall asking Mrs. Menges whether or not if she could say who the hairs in the stocking cap belonged to. It is my recollection she said that she could not. There has been some discrepancy as to whether or not Miss Menges could identify who the hairs in the stocking cap belonged to. Again, I want to caution you it will be your recollection which will prevail as to what Dr. Menges' testimony is along with all the other testimony in this particular case.

Id. at 56–57.

science be presented in a misleading manner during closing arguments. Post-conviction courts are unlikely to provide a legal remedy for such conduct.

2. Failures of Defense Counsel

Defense 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. Defense counsel often also failed to address forensic science during their closing arguments. Defense experts testified in only 19 trials amongst the transcripts reviewed. For example, Gerald Davis presented an expert who contradicted Fred Zain's invalid testimony: "It is an absolute medical certainty that this seminal fluid did not come from Gerald Wayne Davis."[290] That expert, Echols A. Hansbarger, Jr., noted that if he dismissed ABO exclusions as false results, "I'm afraid that I wouldn't be in my job very long."[291] He was paid $100 for his work in the case.[292]

Stephen Linscott presented an expert who, on cross-examination, rejected any notion that probability of association can be used to evaluate hair evidence. The prosecutor nevertheless asserted that the defense expert endorsed such probability testimony, which the appellate court in reversing the conviction found to be a "calculated, rank misrepresentation."[293] Curtis McCarty also presented an expert, John Wilson, the Chief Forensic Chemist of the Regional Crime Lab in Kansas City. Wilson only reviewed Joyce Gilchrist's report and did not conduct any independent analysis, but he corrected Gilchrist's faulty use of serology population statistics.[294]

Perhaps defense attorneys cannot be expected to understand scientific evidence and effectively cross-examine state experts, much less test the accuracy of the underlying data, without access to defense experts. Nevertheless, courts frequently deny the defense

---

[290] Transcript, G. Davis, supra note 132, at 326.
[291] Id. at 331.
[292] Id. at 334.
[293] People v. Linscott, 511 N.E.2d 1303, 1309 (Ill. App. Ct. 1987).
[294] McCarty v. State, 765 P.2d 1215, 1218 (Okla. Crim. App. 1988) ("Mr. Wilson testified that Gilchrist's forensic report reflected that none of the pubic hairs found on the victim were consistent with appellant.").

funding for experts in criminal cases in which forensic evidence plays a central role.[295] The presentation of forensic science testimony is typically one-sided in the majority of states that do not routinely fund the provision of forensic experts for indigent defendants. Moreover, in cases where defendants are able to present expert testimony, the experts are sometimes inexperienced or ineffective, and they may not have access to the underlying forensic evidence. Thus, it should come as no surprise that, despite the stakes, the defense does not often meaningfully challenge invalid forensic science testimony.

3. Judicial Rulings on Forensic Science

Courts policed the introduction of forensic testimony in these trials in a highly deferential manner, typically trusting the jury to assess the expert testimony.[296] Defense attorneys did not often raise trial motions or objections regarding forensic testimony. In the Glen Woodall case, the defense moved to exclude Zain's "erroneous" serology chart because Zain had included a false statistic, dividing his figure in half and supposedly limiting his numbers just to men. The Court denied the motion: "That's in dispute. That's something the jury will have to determine."[297] In Edward Honaker's case, the court denied a new trial motion, stating, "In the opinion of the court the evidence against you was overwhelming. You couldn't get around the scientific evidence that one of your hairs was found on her person."[298]

Similarly, fairly few of these exonerees challenged the forensic evidence during their appeals or post-conviction. Few among the first 200 people exonerated by post-conviction DNA testing whose

---

[295] See 1 Giannelli & Imwinkelried, supra note 78, §§ 4.01–4.05, 13.07; Jack B. Weinstein, Speech, Science, and the Challenge of Expert Testimony in the Courtroom, 77 Or. L. Rev. 1005, 1008 (1998) ("Courts, as gatekeepers, must be aware of how difficult it can be for some parties—particularly indigent criminal defendants—to obtain an expert to testify.").

[296] See Gross & Mnookin, supra note 18, at 169 ("Once a witness has been permitted to testify as an expert under Rule 702, judges usually leave the task of correcting and explaining their instructional statements to the opposing parties and the expert witnesses they call.").

[297] Trial Transcript at 1168, State v. Glen Dale Woodall, No. 87-F-46 (W. Va. Cir. Ct. July 7, 1987).

[298] Transcript, Honaker, supra note 64, at 29.

cases were examined in the first author's *Judging Innocence* study had asserted claims challenging the forensic evidence during their appeals or post-conviction, though 6 exonerees had such claims granted.[299]

Among those exonerated since that study was completed, Kennedy Brewer and Curtis McCarty brought claims related to the forensic evidence. In Brewer's case, the Mississippi Supreme Court twice rejected his claims challenging Michael West's erroneous bite mark testimony.[300]

In contrast, McCarty had his conviction reversed twice based on challenges to Joyce Gilchrist's testimony. First, McCarty had his 1986 conviction and death sentence reversed in 1988 due to Joyce Gilchrist's misconduct concerning the forensic analysis at his criminal trial. Regarding her agreement with the statement that "McCarty was physically present during the time violence was done to [the victim]," the court noted, "We find it inconceivable why Ms. Gilchrist would give such an improper opinion, which she admitted she was not qualified to give."[301] McCarty was convicted again in 1989, and Gilchrist's testimony at his second trial was found not to be grounds for reversal on appeal.[302] When his conviction was ultimately vacated again in 2005 based on Gilchrist's alteration and fabrication of laboratory reports, the court emphasized that "Ms. Gilchrist's actions alone warrant a new trial."[303] In 2007, after serving twenty-two years in prison, McCarty was finally exonerated by post-conviction DNA testing.[304]

Courts denied relief to exonerees who asserted claims regarding some of the most flagrantly invalid forensic testimony reviewed in this study. Courts typically deferentially review any trial court evidentiary rulings as to the relevance of the proffered testimony, and

---

[299] See Garrett, supra note 7, at 73–94.

[300] See Brewer v. State, 819 So. 2d 1169, 1176 (Miss. 2002); Brewer v. State, 725 So. 2d 106, 134 (Miss. 1998).

[301] McCarty v. State, 765 P.2d 1215, 1218 (Okla. Crim. App. 1988).

[302] McCarty v. State, 904 P.2d 110, 125–29 (Okla. Crim. App. 1995) (denying relief on claims regarding Gilchrist's hair comparison testimony, but remanding for new capital sentencing hearing).

[303] McCarty v. State, 114 P.3d 1089, 1094 (Okla. Crim. App. 2005) (affirming grant of new trial based on Gilchrist's *Brady v. Maryland* violations).

[304] Cheryl Camp, Convicted Murderer Is Freed in Wake of Tainted Evidence, N.Y. Times, May 22, 2007, at A16; The Innocence Project, Curtis McCarty, http://www.innocenceproject.org/Content/576.php (last visited Nov. 10, 2008).

harmless error doctrines may further lead a court to excuse invalid admission of such testimony. Examples are easily found in reported cases outside this study set: for instance, though some courts have ruled that probability estimates by hair examiners are inadmissible,[305] other courts have affirmed their use even in cases involving wildly fabricated numbers like a one in fifteen billion chance of a random hair match.[306] One example from a case studied here is the testimony of Fred Zain in the Gerald Davis case, in which Zain testified that bacteria could account for serological results that excluded Davis. The Supreme Court of Appeals of West Virginia noted that Zain explained the presence of characteristics foreign to Davis as "the result of a false reading due to bacterial contamination."[307] Given a forgiving sufficiency standard in which the evidence is viewed in the light most favorable to the prosecution, the court found that no "injustice ha[d] been done," and that "the scientific evidence d[id] not conclusively demonstrate the appellant's innocence."[308]

Similarly, in the Larry Ollins case, the Appellate Court of Illinois denied relief despite the testimony of Pamela Fish, an expert from the Chicago Police Department Crime Laboratory, who falsely asserted that "the defendant's blood sample matched semen found in the victim." The court observed that "the test results were

---

[305] See, e.g., United States v. Massey, 594 F.2d 676, 680 (8th Cir. 1979) ("There is no foundation to show the factual circumstances surrounding each of [the expert's] examinations and certainly there is no statistical probability which could be drawn from his experience to show that there was only 'one chance in a 1,000' that hair comparisons could be in error."); State v. Carlson, 267 N.W.2d 170, 175–76 (Minn. 1978) (finding hair comparison testimony to be "improperly received" where it provided a "suggestion of mathematical precision"; Barry Gaudette had testified that, based on his study, he found a "1-in-4,500 chance that the head hairs did not belong to the accused"); Butler v. State, 108 S.W.3d 18, 21–27 (Mo. Ct. App. 2003); State v. Scarlett, 426 A.2d 25, 28–29 (N.H. 1981); State v. Faircloth, 394 S.E.2d 198, 202–03 (N.C. Ct. App. 1990).

[306] See, e.g., Lopez v. State, 643 S.W.2d 431, 433 (Tex. App. 1982) ("This expert testified that the chances of the resemblance he found between the hair samples occurring at random was one in 1.5×10^{10} (1 in 15,000,000,000)."); State v. Bridges, 421 S.E.2d 806, 808 (N.C. Ct. App. 1992) (finding harmless error where expert testified that the "likelihood of two Caucasian individuals having indistinguishable head hair . . . is very low. A conservative estimate for that probability would be . . . approximately one in a thousand.").

[307] State v. Davis, 376 S.E.2d 563, 567 (W. Va. 1988).

[308] Id. at 568.

corroborated by an eyewitness account of the crime. Additionally, the State points out that the jury was instructed as to the specific limitations of the test results in both the opening and closing arguments of the prosecution." The court concluded, "Because the test results were corroborated by substantial independent evidence, we find that Fish's testimony was properly admitted into evidence by the trial court."[309]

The Montana Supreme Court in the Chester Bauer case found no reversible error in analyst Julie Long's testimony. Long had ignored the problem of masking and quantification and falsely divided the statistic offered in half, claiming to rule out females. Presented with a claim regarding this invalid testimony, the court denied relief, stating that "the fact remains that Bauer is a secretor, and that should be relevant."[310]

B. A National Framework for Reform of Forensic Science

In 1989, just as DNA technology arrived, Eric Lander commented, "At present, forensic science is virtually unregulated—with the paradoxical result that clinical laboratories must meet higher standards to be allowed to diagnose strep throat than forensic labs must meet to put a defendant on death row."[311] Two decades later, that state of affairs has changed little, making the invalid testimony prevalent in these wrongful conviction cases unsurprising. No legislation or oversight mechanisms regulate the quality of forensic science reports or testimony.

DNA exonerations have provided some impetus for state and local forensic science reform. Several states have enacted legisla-

---

[309] People v. Ollins, 601 N.E.2d 922, 924–25 (Ill. App. Ct. 1992); see also People v. Saunders, 603 N.E.2d 32, 36 (Ill. App. Ct. 1992) ("Fish's testimony is corroborative of the defendant's own admission that he was present . . . even though the semen samples taken from the victim excluded the defendant.").

[310] State v. Bauer, 683 P.2d 946, 951 (Mont. 1984). Compounding the problem, the court supported its ruling with reference to additional invalid science testimony:

> [T]here is other independent evidence of Bauer's guilt. . . . Arnold Melnikoff, Bureau Chief of the State Crime Laboratory, testified that pubic hair and head hair found at the crime scene were similar to Bauer's pubic and head hair. Melnikoff estimated that the chances of another person having the same type of pubic and head hair were one in ten thousand.

Id.

[311] Eric S. Lander, DNA Fingerprinting on Trial, 339 Nature 501, 505 (1989).

tion creating independent bodies to review their crime laboratories in response to misconduct.[312] However, very few exonerations have resulted in scrutiny and audits of the invalid science that supported the wrongful conviction. Regarding DNA laboratories, Congress, in 1993, with the establishment of a national DNA databank, created a DNA Advisory Board to provide quality assurance standards.[313] Many state and local crime labs voluntarily participate in ASCLD/LAB.[314] But voluntary programs run by the crime laboratory directors, although a positive step, fail to address the needs identified in this study. None deal with the claims made by forensic analysts in their reports or in their trial testimony.

For disciplines other than DNA analysis, the federal government in 2004 required states to create an entity and process for external independent audits to respond to allegations of misconduct or serious negligence in laboratories receiving federal funds.[315] That legislation, however, has not been enforced, according to the Inspector General, and many states are not in compliance.[316] Moreover, the audits are limited to misconduct in the government laboratories and thus fail to cover the serious misconduct of unaffiliated forensic analysts hired by prosecutors, including forensic odontologists and the employees of police departments that conduct analyses of ballistics and fingerprint evidence.

Despite these faint stirrings of reform, crime laboratories and forensic analysts remain remarkably free from oversight and still lack basic scientific standards to govern their conclusions. No federal

---

[312] See, e.g., Minn. Stat. § 299C.156 (2007) (establishing Forensic Laboratory Advisory Board); N.Y. Exec. Law §§ 995-a to -b (McKinney 2003) (establishing forensic science commission and requiring accreditation); Okla. Stat. Ann. tit. 74, § 150.37 (West 2007) (requiring accreditation); Tex. Code Crim. Proc. Ann. art. 38.35(d) (Vernon 2005) (requiring accreditation by the Texas Department of Public Safety); Va. Code Ann. § 9.1-1101 (2006) (creating Department of Forensic Science and oversight committee).

[313] 42 U.S.C. § 14132(a), (c) (2000).

[314] See ASCLD/LAB Home Page, http://www.ascld-lab.org/ (last visited Nov. 10, 2008).

[315] See 42 U.S.C. § 3797(k)(4) (Supp. IV 2007) (requiring that laboratories receiving federal grants create mechanisms for external independent investigations).

[316] See Oversight of the Department of Justice's Forensic Grant Programs: Hearing Before the S. Comm. on the Judiciary, 110th Cong. (2008) (statement of Glenn A. Fine, Inspector General, U.S. Department of Justice); Oversight of the Justice For All Act: Hearing Before the S. Comm. on the Judiciary, 110th Cong. (2008) (statement of Peter Neufeld on Behalf of The Innocence Project).

legislation regulates the quality of non-DNA forensic disciplines or the content of reports or testimony, which is significant because the overwhelming majority of crime lab work involves techniques other than DNA analysis.

Nationally enforceable standards should govern the interpretation of forensic data within scientifically acceptable parameters. The authors also appreciate the need for a federal infusion of capital for both basic and applied research and to ensure that only validated and reliable methods and assays are used to analyze evidence of crime. But since these are needs that are not derived solely from this study set, they will not be addressed here.[317]

National scientific standards should address the use of forensic science: both the methodology and, importantly, the way that conclusions are drawn from evidence. Even in areas such as ABO blood typing, in which there is consensus on the reliability of the testing methods, invalid testimony can result from a failure to ensure adherence to scientific standards when drawing potentially unsound conclusions from sound testing methods. In contrast, invalid science testimony was unsurprising in disciplines where there was simply no consensus on the boundaries of permissible trial testimony. Even in disciplines that provided non-binding guidelines, no criteria were provided for reaching conclusions. Indeed, bite mark and shoe print guidelines explicitly permit conclusions unsupported by science. Furthermore, the forensic disciplines have created no means to enforce any scientific standards.[318] Forensic laboratories have also not created effective mechanisms for review of analysts' work and have typically not responded even after invalid forensic testimony and analysis was uncovered. Outside intervention is urgently needed.

The NAS Committee report provided the long overdue recommendation that we meaningfully regulate the presentation of forensic science. In particular, the NAS report recommended that

---

[317] See, e.g., Giannelli, supra note 32, at 72–76, 87–89 (discussing proficiency testing, accreditation of crime laboratories, and other avenues of reform); Henry C. Lee, Forensic Science and the Law, 25 Conn. L. Rev. 1117, 1124 (1993) ("Perhaps the most important issue in forensic science is the establishment of professional standards.").

[318] See Gross, supra note 62, at 1178 ("Unfortunately, what an expert says in court is generally invisible and inaudible in her own professional world. If expert witnesses were accountable to their colleagues, even informally, they might fear the consequences of irresponsible testimony far more than they do.").

Congress establish an independent federal agency, a "National Institute of Forensic Science," which would promulgate "standard terminology" for report writing and testimony.[319] The history of the development of standards for DNA analysis provides a model. The National Institutes of Health funded basic early research, as did universities and other institutions. The National Institute of Standards and Technology promulgated quality assurance standards, including match criteria. A National Institute of Forensic Science, led by independent scientists—not those employed by crime laboratories—could take on the important task of developing scientific criteria for interpreting data in many of the non-DNA disciplines, particularly those that attempt to identify a person or object in connection with evidence recovered from a crime scene. The forensic disciplines would benefit from scientific criteria to promote consistency and best practices. Invalid written reports and expert testimony damage the credibility of the entire forensic science community. National standards would reduce the number of wrongful convictions and enhance the likelihood that forensic science could help identify the real perpetrator.

The NAS Committee's report and the responses of those in the scientific community will contribute to a national discussion regarding the future of the forensic sciences. That discussion will hopefully lead to the type of legislation and oversight proposed. Forensic sciences urgently require a far more rigorous scientific foundation. Specifically, there should be a sound foundation for the process by which analysts reach the conclusions in their reports and trial testimony. Should reformers focus only on methodology and not also on the actual practice of forensic science in the courtroom, invalid testimony and miscarriages of justice will continue to tax our criminal justice system and society.

CONCLUSION

This exploration of the types of invalid forensic science testimony that have contributed to convictions of the innocent provides one window into failures of our criminal justice system to adequately review the use of forensic science. That system still has not responded with a full investigation into most of these known mis-

[319] See Strengthening Forensic Science, supra note 12, at S-14, 6-3–6-5.

carriages of justice, much less looked into other potentially affected cases or routinely conducted investigations to ensure accurate dispositions. These cases suggest that the adversary system cannot be depended upon as an adequate safeguard. The defense bar lacked the expertise and resources to detect and address invalid forensic science effectively in most of these cases, and judges did not remedy most errors brought to their attention.

Finally, the invalid science testimony described here ranges from cases in the 1980s involving conventional forensic disciplines employing visual comparison, to serology analysis employing clear population statistics, to the use of modern DNA technology in the 1990s. Though the technology has changed over time, the sources of human error, misinterpretation, and misconduct have not. This body of cases in which innocent people were convicted based on invalid forensic testimony demonstrates the importance of efforts to study the validity of forensic science as presented in reports and testimony at trial and encourages us to rethink how our criminal system handles forensic science in the laboratory and in the courtroom.

The evidence from these wrongful conviction cases supports efforts to create and enforce national standards for interpreting data in the various forensic disciplines. The scientific community can take the lead in reform efforts. Detailed scientific standards are needed to establish permissible parameters both for report writing—particularly important since so many cases are resolved by plea bargains—and for trial testimony. As the criminal trials of these innocent people demonstrate, if reforms are not implemented to regulate the practice of forensic science during criminal investigations long before the adversary process begins, the opportunity to avert miscarriages of justice may be lost.